Child Development, March/April 2007, Volume 78, Number 2, Pages 558–580

Teachers' Education, Classroom Quality, and Young Children's Academic Skills: Results From Seven Studies of Preschool Programs

Diane M. Early, Kelly L. Maxwell, and Margaret Burchinal
University of North Carolina at Chapel Hill

Soumya Alva
Westat Inc.

Randall H. Bender
RTI International

Donna Bryant, Karen Cai, and Richard M. Clifford
University of North Carolina at Chapel Hill

Caroline Ebanks
Institute of Education Sciences, U.S. Department of Education

James A. Griffin
National Institute of Child Health and Human Development, NIH

Gary T. Henry
University of North Carolina at Chapel Hill

Carollee Howes
University of California at Los Angeles

Jeniffer Iriondo-Perez
RTI International

Hyun-Joo Jeon
University of California at Los Angeles

Andrew J. Mashburn
University of Virginia

Ellen Peisner-Feinberg
University of North Carolina at Chapel Hill

Robert C. Pianta
University of Virginia

Nathan Vandergrift
University of North Carolina at Chapel Hill

Nicholas Zill
Westat Inc.

In an effort to provide high-quality preschool education, policymakers are increasingly requiring public preschool teachers to have at least a Bachelor's degree, preferably in early childhood education. Seven major studies of early care and education were used to predict classroom quality and children's academic outcomes from the educational attainment and major of teachers of 4-year-olds. The findings indicate largely null or contradictory associations, indicating that policies focused solely on increasing teachers' education will not suffice for improving classroom quality or maximizing children's academic gains. Instead, raising the effectiveness of early childhood education likely will require a broad range of professional development activities and supports targeted toward teachers' interactions with children.

In the United States, 4-year-olds increasingly are being served in programs specifically designed to improve their school-readiness skills. This increasing focus on early learning skills for 4-year-olds is due in part to research in two areas. First, evidence ranging from studies of brain development to evaluations of preschool programs points to the importance of high-quality early childhood experiences in providing the foundation for later school success (NICHD Early Child Care Research Network, 2005; National Research Council and Institute of Medicine, 2000; Peisner-Feinberg et al., 2001). Exposure to high-quality care appears especially important for at-risk children's later school success (Burchinal et al., 2000; Campbell, Ramey, Pungello, Sparling, & Miller-

After the first three authors, all authors contributed equally to this project and are listed in alphabetical order.

The national Early Head Start Research and Evaluation Project was funded by the Administration for Children and Families

Johnson, 2002; Peisner-Feinberg et al., 2001; Reynolds, Temple, Robertson, & Mann, 2002; Schweinhart, 2004). Second, research has shown that children from low-income families and children of color enter school with significantly fewer skills than their more advantaged and White peers (Lee & Burkam, 2002; McLanahan, 2005; Phillips, Brooks-Gunn, Duncan, Klebanov, & Crane, 1998; Stipek & Ryan, 1997).

In response to both these trends—the emphasis on the early years, especially for children's later school success, and recognition that all children do not start school on an equal footing—states are increasingly funding programs to provide an early educational experience for 4-year-olds. For instance, in 2004–2005, over 800,000 children, mostly 4-year-olds, were enrolled in state-funded prekindergarten, representing 17% of the nation's 4-year-olds and a 20% increase from 2001–2002. In 2004–2005 states spent 2.84 billion dollars on prekindergarten initiatives (Barnett, Hustedt, Robin, & Schulman, 2005). Furthermore, the federal government serves approximately 500,000 4-year-olds through Head Start (U.S. Department of Health and Human Services, 2006). State and federal programs for 4-year-olds aim to provide an educational experience at a level of quality high enough to change children's developmental trajectory meaningfully. Policymakers and administrators establish program standards, such as teacher qualifications, that are intended to ensure high-quality experiences for participating children.

Teacher qualifications have been identified as an important correlate of classroom quality (Burchinal, Cryer, Clifford, & Howes, 2002; de Kruif, McWilliam, Ridley, & Wakely, 2000; Howes, Whitebook, & Phillips, 1992; NICHD ECCRN, 2002; Scarr, Eisenberg, & Deater-Deckard, 1994). Policymakers and program advocates sometimes prioritize teacher qualifications as a key strategy for ensuring that programs positively affect children's skills. Policies that mandate certain levels of educational attainment are controversial because they are expensive to the public. Teacher salaries are one of the largest expenditures for any education program (Cost, Quality, & Child Outcomes Study Team, 1995; U.S. Department of Education, IES, 2004; Table 161) and salaries are closely linked to educational requirements. Thus, policymakers face the difficult task of identifying and setting teacher qualification standards high enough to produce high-quality classrooms with the desired child outcomes and yet not so high that programs cannot afford to pay the needed salaries or cannot recruit enough teachers who meet the standards.

Teachers' Education and Classroom Quality

Increasingly, early childhood advocates are calling for all teachers of 3- and 4-year-olds to have at least a

(ACF), U.S. Department of Health and Human Services, under contract 105-95-1936 to Mathematica Policy Research, Princeton, NJ, and Columbia University's Center for Children and Families, Teachers College, in conjunction with the Early Head Start Research Consortium. The Consortium consists of representatives from 17 programs participating in the evaluation, 15 local research teams, the evaluation contractors, and ACF.

The Head Start Family and Child Experiences Survey (FACES) was funded by the ACF, U.S. Department of Health and Human Services, under task order 03Y00318101D to contract GS23F8144H to Westat, Rockville, MD.

The Georgia Early Childhood Study (GECS) was funded in part by the UPS Foundation, National Institute for Early Education Research (NIEER), which is generously funded by the Pew Charitable Trusts, and Bright from the Start: Georgia Department of Early Care and Learning.

The Evaluation of the North Carolina More at Four Prekindergarten Program was funded by the North Carolina Department of Health and Human Services under contract #2090002872.

The Multi-State Study of Pre-Kindergarten was supported by the Educational Research and Development Center Program, PR/Award Number R307A60004, administered by the Institute of Education Sciences (IES), U.S. Department of Education. The State-Wide Early Education Programs (SWEEP) Study was supported by the NIEER, The Pew Charitable Trusts, and the Foundation for Child Development.

The National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (SECCYD) is directed by a Steering Committee and supported by the NICHD through a cooperative agreement (U10), which calls for scientific collaboration between the grantees and the NICHD staff. No official support or endorsement by NICHD, NIH, or the U.S. Department of Health and Human Services is intended or should be inferred.

Support for the analysis of data from the PCER was provided by the IES, U.S. Department of Education, through a contract (ED-01-CO-0052/0004) to RTI International. The views expressed in this paper are those of the authors and do not represent the views of the U.S. Department of Education. James Griffin worked for the IES, U.S. Department of Education, and served as the Preschool Curriculum Evaluation Research (PCER) Program Project Officer when the data reported here were collected and analyzed. This article was coauthored by James Griffin in his private capacity.

Finally, the authors would like to thank the Foundation for Child Development for partial support of the preparation of this manuscript, including a meeting to discuss preliminary findings.

The authors are grateful for the help of the many children, parents, teachers, administrators, and field staff who participated in these studies, as well as several colleagues who reviewed an earlier version of this paper. The views expressed in this paper are those of the authors and do not represent the views of the funding agencies.

Correspondence concerning this article should be addressed to Diane M. Early, FPG Child Development Institute, CB 8040, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599. Electronic mail may be sent to diane_early@unc.edu

Bachelor's degree, often also including a major in early childhood education or state certification to teach this age group (Barnett, 2003; Barnett et al., 2005; National Research Council, 2001; Trust for Early Education, 2004). Many public prekindergarten programs are following this recommendation. According to the National Institute for Early Education Research (NIEER; Barnett et al., 2005), 17 of 38 states with prekindergarten programs require that all lead teachers hold a Bachelor's degree, and another 12 states require a Bachelor's degree for at least some of their prekindergarten teachers. Furthermore, 27 states require that the teachers have specialized training in early childhood education.

This push for every preschool teacher to have a Bachelor's degree in early childhood education is based on two lines of past research: (1) studies, generally from community-based child-care settings; linking teachers' education to classroom quality, and (2) research linking early care and education classroom quality to children's academic gains. Logically, if programs are interested in children's academic gains, improving teachers' education seems to be a reasonable place to start.

In the child-care literature, most research indicates that higher levels of teachers' education are linked to higher global quality in center-based care (Burchinal et al., 2002; de Kruif, et al., 2000; Howes et al., 1992; NICHD ECCRN, 2002; Scarr et al., 1994). However, one study of state-funded prekindergarten, using some of the data included in the current project, found largely null associations (Early et al., 2006) and other research is not entirely conclusive. For instance, using data from the Cost, Quality, and Outcomes study, both Blau (2000) and Phillipsen, Burchinal, Howes, and Cryer (1997) found that level of teachers' education was associated with quality in uncontrolled models; however, once a host of parent- and center-level variables known to be linked to quality were added to the model, the association with teachers' education disappeared. Likewise, Phillips, Mekos, Scarr, McCartney, and Abbott-Shim (2001) found significant, positive zero-order correlations between teachers' level of education and classroom quality and between teacher training in early childhood education and classroom quality. However, these effects disappeared for preschoolers when hierarchical multiple-regression models controlled for regulatory stringency, regulatory compliance, group size, and ratio. Thus, the relation between teachers' education and quality classrooms is sometimes evident only when simple analysis techniques are utilized, and the relationship weakens when a more complex model is used. This is possibly due to the fact that teachers' education tends to be correlated with these other important predictors of classroom quality, making it difficult to tease them apart. It may be that some type of selection is taking place, possibly in the form of more highly educated teachers choosing to work in higher quality settings (Hamre & Bridges, 2004).

Tout, Zaslow, and Berry (2005) recently completed a review of the research examining links between early childhood teachers' education and classroom quality. They concluded that higher levels of teacher education, especially education that focuses on early childhood development, are generally linked to higher quality, but that there is insufficient research addressing "thresholds" to support a specific cut point. In other words, whereas the existing literature generally indicates that more education may be beneficial, there is no conclusive evidence that a teacher with a Bachelor's degree or any other specific level of education will produce or ensure a high-quality classroom or children's learning.

The lack of common definitions of education and training, coupled with the use of different controls in different studies, severely limit our ability to draw straightforward conclusions from the existing early childhood literature (Maxwell, Feild, & Clifford, 2005; Tout et al., 2005). The current study aims to address this problem by asking the same set of questions, using the same set of definitions and controls, across a number of large data sets. Using this strategy, we can be confident that any differences in findings are not due to differences in definition or methodology. The study approach allows us to test the direct effects of teacher education on classroom quality and children's skills using a value-added specification.

Purpose

The goal of the current project was to consider the links between teachers' education, specifically educational degree and major, and two important outcomes—classroom quality as well as children's academic skills in the year before kindergarten entry. To answer the research questions with the greatest degree of confidence, we conducted a series of common analyses using comparable data from seven major studies. Thus, unlike past research in this area that has used different definitions, different methodologies, and different statistical approaches, the answers derived from the current analyses are directly comparable to one another.

Research Questions

In light of the mixed evidence from past research and rapidly expanding public early education programs, it is important to answer some basic questions with regard to teachers' education and major. Policymakers want to know what set of policies regarding teachers' education and major are most likely to lead to high-quality classrooms where children make meaningful academic gains. Policymakers are not in a position to interpret the subtleties of various types of education and training. For example, whereas researchers may suspect that degrees from some colleges and universities are more valuable than others, policymakers are not likely to require that all teachers have degrees from a limited pool of schools. Likewise, whereas recent changes in teacher preparation programs may have increased the value of some degrees, policymakers cannot require that all teachers be recent graduates. Highly nuanced information about the precise types of instruction, experiences, and skills teachers should have would be of great interest to the early childhood field and should be examined to improve quality and outcomes. However, policymakers are not likely to request or use this detailed information in promoting child development through early education programs. The goal of this project, then, is to answer questions that policymakers who are setting standards for early childhood programs might ask.

This project uses data from multiple studies to answer three questions:

(1) Does the educational degree of lead teachers relate to observed classroom quality and children's academic skills using a value-added specification? And, more specifically, do teachers with a Bachelor's degree or higher have classrooms of higher quality or children who learn more during the prekindergarten year?
(2) Among lead teachers whose highest degree was in early childhood education or child development, does the level of the highest degree predict classroom quality and/or children's academic skills using a value-added specification?
(3) Among teachers whose highest degree is a Bachelor's, does a major in early childhood education or child development predict better quality or greater academic skills than a major in another field of education or a noneducation major, controlling for baseline skill levels using a value-added specification?

Originally, we intended to test the value of an early childhood major among Associate's degree-level teachers as well (parallel to question 3). However, no study had enough variance in teacher major among the Associate's level teachers to conduct meaningful analyses.

Method

Participating Studies

The goal of this project was to analyze several large data sets using similarly defined variables and equivalent model specification to answer a common set of questions—an analysis strategy we call ''replicated secondary data analysis.'' This technique involves selecting studies that contain similar information, gathered in the same way, and using common analysis protocols across data sets so that any differences in relations among the variables are attributable to the sample or study circumstances, rather than to different data collection strategies, variable operationalizations, or analysis techniques. All participating studies had to meet three criteria: (1) contain data about teachers' education, (2) contain observed classroom quality data, and (3) contain direct assessments of children's academic skills during the 4-year-old year, plus pretest data that could be used to control for prior child functioning. In addition, five of seven studies are statistically representative because each is based on a sample that was randomly selected to represent a known population. The first three authors identified eight studies that met these criteria; seven of the eight participated. These studies were not designed or conducted jointly, and therefore some differences exist in available variables; nonetheless, these seven are sufficiently similar that we believe any differences in findings must be attributed to differences in samples, rather than methodology or analysis strategy.

Several safeguards were put in place to ensure that the data were analyzed consistently across studies. The first and second authors reviewed each study's questionnaire and interview protocols and selected the appropriate questions for specifying the teacher education and major variables. Each study was provided with a set of specific instructions for creating the needed analysis and control variables, handling missing data, and specifying the models. Furthermore, each study was provided with a sample SAS code that could be modified to ensure that analyses were conducted identically. Analyses for three of the studies (More at Four [MAF], National Center for Early Development and Learning

[NCEDL], and NICHD) were conducted by the same individual. Data analysts for the other four studies were in regular contact with the project's lead data analyst (the manuscript's third author). After initial analyses were completed, each study's data analysts and principal investigators participated in a 2-day face-to-face meeting to discuss preliminary findings, analysis strategies, appropriate interpretation of the findings, and next steps. The first author carefully reviewed all data tables, checking for internal consistency. Analyses were rerun whenever inconsistencies were discovered.

Study Descriptions

Table 1 briefly describes the seven participating studies.

Early Head Start (EHS) Follow-Up

The primary purpose of the EHS study was to assess the effectiveness of the EHS program. Low-income families and children who were eligible for EHS were selected from 17 sites, when their infants were <12 months old. Participating children were born between July 1996 and September 1998. These families and children were randomly assigned to receive EHS services or to a control condition. The children were followed every year from birth to 3 years of age and then follow-up data were collected again immediately before beginning kindergarten. Data for the current analyses came from the follow-up wave, just before kindergarten entry (controlling for assessments at 3 years of age). At that time, the children were no longer in their randomly assigned setting and were attending a variety of early care and education settings (including 45% Head Start). The current analyses included children from both the EHS and control conditions. The two groups were combined because participation in the experimental treatment (EHS vs. control) had ended more than a year earlier, the teachers participating in the current wave of the study were not affiliated with EHS, and the role of earlier experiences is not a question of interest for the current analyses. For more details about the EHS methodology, see Administration for Children and Families (2002).

Head Start Family and Child Experiences Survey (FACES 2003)

The primary purpose of the FACES 2003 study was to describe the quality of Head Start programs in a nationally representative sample, toward the goal of implementing a system of program performance measures and improving accountability for Head Start programs. The sample comprising this longitudinal data set consisted of 63 randomly selected Head Start programs, stratified by census region, percent minority, and urbanicity. As in earlier FACES cohorts from 1997 and 2000, FACES 2003 featured four phases of data collection and followed 3- and 4-year-old Head Start children from program entry through the spring of kindergarten. The current analyses are restricted to children who were 4 years old in Spring 2004 and include children's assessments from the fall (2003) and spring (2004) of their first year of Head Start. For more information on this study and its methodology, see Zill and Resnick (2005) and Zill et al. (2003).

Georgia Early Care Study (GECS)

This study sought to examine the experiences and development of children attending public and private preschool in Georgia. The sample includes 128 classrooms and 630 children, representing all full-day, full-year preschools in Georgia, including Head Start, Georgia Pre-K, and private programs. To select the sample, counties were stratified by the number of 4-year-olds and four to eight counties were selected per strata. Georgia Pre-K, Head Start, and private preschools were then selected within each county from complete lists obtained from the three agencies that administered each program. Within each selected site, one classroom was selected at random. Five children in each participating classroom were randomly selected for participation from the children whose parents consented to have their child participate. The current analyses include child assessments from the fall and spring of the preschool year. For more details on this study, see Henry et al. (2003, 2004).

More at for (MAF) Evaluation

The primary purpose of the MAF Evaluation study was to examine the quality of the program and outcomes for children participating in MAF, a statewide prekindergarten initiative for at-risk 4-year-olds in North Carolina. The classroom quality sample included 233 classrooms, randomly selected over two program years from all those in operation across the state. A subsample of 98 of these classrooms was randomly selected for child assessments, and those children participating in the MAF program within these classrooms were recruited for the study. A total sample of 785 children was included across the 2 years

Table 1
Study Descriptions

	EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Classroom sample	887	310	128	233	721	639	76
Child sample	887	1041	630	785	2966	639	667
Year quality and posttest collected	2001–2003 (three cohorts)	2004	2002	2002–2003 and 2003–2004	2001 (Multi-State Study of Pre-K) 2004 (SWEEP)	1995–1996	2002–2003
Population analysis sample represents	Children who had been in EHS Evalution Study as infants or toddlers	4-year-old children enrolled in their first year of Head Start in Fall 2003	Full-day, full-year preschool participants in Georgia including Head Start, Georgia Pre-K and private programs	MAF (North Carolina's state-funded prekindergarten program) participants	State funded prekindergarten classrooms and children in 11 participating states	Children at the 10 locations across the United States who were in center-based care the year before kindergarten.	Cohort 1 control group children at seven locations across the United States
Percent Head Start Classrooms (%)	45	100	22	15	16	9	47
Percent in state prekindergarten (%)	NA	NA	54	100	100	8	46
Percent housed in public schools (%)	NA	NA	18	59	61	NA	50
Mean child (*SD*) age (months) at Time 1	37.10 (1.41)	47.98 (41.64)	55.57 (3.49)	54.00 (3.45)	55.44 (3.84)	36 (NA)	56.17 (4.07)
Mean child (*SD*) age (months) at Time 2	62.36 (3.84)	53.50 (40.98)	62.47 (3.49)	60.96 (3.42)	60.60 (3.84)	55.53 (1.03)	60.63 (3.96)

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NA = information not available; NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
Head Start, state prekindergarten, and housed in public schools are not mutually exclusive due to blended funding streams and variations in state initiatives.

from these 98 classrooms. The current analyses include child assessment data from the fall and spring of their MAF prekindergarten year and MAF classroom quality data that were gathered mid-year. For further details about the methodology for these studies, see Peisner-Feinberg and Maris (2005a, 2005b).

National Center for Early Development and Learning (NCEDL)

NCEDL conducted two studies of state-funded prekindergarten: the Multi-State Study of Pre-Kindergarten and Study of State-Wide Early Education Programs (SWEEP). The goals and methodologies of the two studies were largely similar; therefore, they have been combined for the current analyses. Both studies sought to describe state-funded prekindergarten programs in states that had large, well-established programs. In all, 11 states participated. In each state, sites were randomly selected from lists provided by the states of sites providing state-funded prekindergarten. Then, in each site, one classroom serving primarily 4-year-olds was selected at random. Within each classroom, four children who were old enough to attend kindergarten the following year were selected. The current analyses include children's assessments from the fall and spring of the prekindergarten year. For further details about the methodologies for these studies, see Early et al. (2005).

Study of Early Child Care and Yolk Development (NICHD SECCYD)

This study was designed to examine the relationship between child-care experiences and characteristics and children's developmental outcomes. The participating children were a conditional random sample selected shortly after birth during hospital visits at 10 locations across the United States. Data were collected in whatever care and education setting the children attended. The current analyses include child assessments from 36 and 54 months. For more information on this study's methodology, see NICHD SECC (n.d.).

Preschool Curriculum Evaluation Research Program (PCER) Program

The PCER Program was designed to conduct small-scale efficacy evaluations of available preschool curricula that had not been rigorously evaluated. PCER began in 2002 when the IES in the U.S. Department of Education awarded grants to seven researchers to implement several widely used preschool curricula, with Research Triangle Institute (RTI) International serving as the national evaluation coordinator. The evaluations were conducted using a common assessment protocol and a randomized experimental design. Participating classrooms or schools were randomly assigned to intervention or control conditions. Only children assigned to the control group classrooms for the pilot study (2002–2003) are represented in these analyses. For more information regarding this project, see PCER Project (n.d.).

Variable Specification

Each study created a series of identical variables to be included in analyses.

Teachers' Education and Training

Distributions for all teachers' education and training variables for each study are presented in Table 2.

Highest level of education. A four-level variable was created to specify the highest degree attained by the lead teacher. No study contained enough teachers with less than a high school degree to create a separate category; hence, those teachers were excluded from the analyses. The four levels were specified as follows: (1) High school degree or general education diploma (GED). This category includes people who have taken some college or technical courses but have not received a postsecondary degree. (2) Associate's degree. (3) Bachelor's degree. This category includes people who have taken some graduate coursework but have not received a graduate degree. (4) Graduate degree. Graduate is defined as any postbaccalaureate degree such as MA, MEd, EdD, or PhD.

Bachelor's degree. A two-level variable (Bachelor's vs. no Bachelor's) was created in which the Bachelor's group includes all teachers with a Bachelor's or Graduate degree as their highest level of education. The no-Bachelor's group includes all teachers with a high school diploma/GED or Associate's degree.

Major. A three-level variable was created to describe the lead teacher's major when she/he received her/his highest degree. The three categories were: (1) early childhood education or child development (ECE/CD), (2) any other education major, such as elementary or special education, and (3) non-education major, including any field outside of education such as psychology, sociology, biology, or business.

Table 2
Teacher Education and Training: Ns (% of Nonmissing, Non-NA)

	EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Highest level of education							
High school or GED	257 (29)	91 (29)	20 (21)	9 (4)	99 (14)	50 (8)	17 (22)
Associate's	139 (16)	114 (37)	26 (27)	22 (9)	104 (15)	221 (34)	8 (11)
Bachelor's	335 (38)	81 (26)	36 (37)	172 (74)	340 (49)	196 (31)	38 (50)
Graduate	156 (18)	24 (8)	15 (15)	29 (13)	152 (22)	172 (27)	13 (17)
Missing	0	0	31	1	26	0	0
Bachelor's degree							
No	396 (45)	205 (66)	46 (47)	31 (13)	203 (29)	271 (42)	25 (33)
Yes	491 (55)	105 (34)	51 (53)	201 (87)	492 (71)	368 (58)	51 (67)
Missing	0	0	31	1	26	0	0
Major							
ECE/CD	609 (97)	151 (71)	52 (72)	125 (58)	303 (50)	178 (35)	24 (40)
Other education	NA	35 (16)	14 (20)	58 (27)	211 (36)	296 (60)	35 (60)
Other major	NA	26 (12)	6 (8)	34 (16)	82 (14)	23 (5)	0 (0)
NA (no degree)	257	91	20	9	99	50	17
Missing	21	7	36	7	26	92	0
Degree among ECE/CD majors							
Associate's in ECE/CD	132 (21)	103 (68)	19 (37)	20 (16)	92 (30)	178 (47)	5 (21)
Bachelor's in ECE/CD	323 (53)	32 (21)	25 (48)	93 (74)	138 (46)	144 (38)	16 (67)
Graduate in ECE/CD	154 (25)	16 (11)	8 (15)	12 (10)	73 (24)	53 (14)	3 (13)
NA (no degree or degree not ECE/CD)	278	159	70	107	392	255	52
Missing	0	0	6	1	26	9	0
Major among teachers with Bachelor's as final degree							
Bachelor's and ECE/CD	NA	32 (42)	25 (71)	93 (55)	137 (41)	89 (45)	16 (42)
Bachelor's and other education	NA	28 (37)	5 (14)	46 (27)	136 (40)	46 (23)	13 (34)
Bachelor's and other major	NA	16 (21)	5 (14)	30 (18)	65 (19)	61 (31)	9 (24)
NA (Bachelor's not final degree)	NA	229	92	61	381	443	38
Missing	NA	5	1	3	2	0	0
Major among teachers with Associate's as final degree							
Associate's and ECE/CD	NA	103 (92)	19 (71)	20 (100)	92 (90)	65 (29)	5 (63)
Associate's and other education	NA	1 (1)	3 (29)	0 (0)	3 (3)	91 (41)	1 (13)
Associate's and other major	NA	8 (7)	1 (0)	0 (0)	7 (7)	65 (29)	2 (25)
NA (Associate's not final degree)	NA	196	102	211	617	418	68
Missing	NA	2	3	2	2	0	0

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; GED = general education diploma; MAF = More at Four; NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research;
GECS included eight teachers with an Associate's degree and 18 teachers with a 1-year technical degree. These two groups have been combined into a single group (Associate's) to be consistent with the other studies and because the Associate's group was too small to analyze alone. For NICHD, when teachers were asked about their education, Associate's and "some college" were combined as a single option. The Associate's value for NICHD represents all teachers who selected that option.

Classroom Quality

Table 3 provides descriptive information about classroom quality from each study. All studies except one (NICHD) used the Early Childhood Environment Rating Scale–Revised (ECERS–R; Harms, Clifford, & Cryer, 1998) to measure classroom quality. The ECERS–R is a widely used measure of global classroom quality, specifically designed for use in classrooms serving children between $2\frac{1}{2}$ and 5 years of age. Scores on the ECERS–R range from 1 to 7 with 1 indicating "*inadequate*" quality, 3 indicating "*minimal*" quality, 5 indicating "*good*" quality, and 7 indicating "*excellent*" quality. The scale's authors report a total scale internal consistency of .92. For the current analyses, the ECERS–R score is the mean of all items scored. We do not recommend comparing ECERS–R scores across studies because data collectors for the various studies were not trained together, may have followed slightly different coding

Table 3
Descriptive Information About Outcome Variables

	EHS		FACES		GECS		MAF		NCEDL		NICHD		PCER	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ECERS–R Total	5.29	1.13	4.20	0.84	4.32	0.93	5.02	0.74	3.80	0.81	NA		3.42	1.06
Receptive language standard scores														
Time 1	82.91	16.09	84.48	11.01	92.38	15.90	85.42	17.91	94.00	15.01	NA		89.83	15.61
Time 2	92.31	14.44	86.19	11.68	96.54	14.50	89.57	16.20	96.29	14.31	101.20	19.79	93.78	14.42
Prereading standard scores (Woodcock–Johnson Letter–Word Identification)														
Time 1	NA		94.73	16.50	102.28	15.43	NA		101.24	16.05	NA		101.58	14.98
Time 2	90.38	15.03	99.86	15.62	103.77	13.37	NA		102.92	14.08	100.59	13.39	101.73	14.30
Math standard scores (Woodcock–Johnson Applied Problems)														
Time 1	NA		88.95	18.07	96.51	14.43	93.02	14.50	98.42	13.66	NA		94.50	15.29
Time 2	88.31	17.85	92.58	14.16	98.30	13.31	93.98	13.24	99.11	12.85	105.06	15.22	97.32	13.80

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
NICHD used the ORCE quality composite ($M = 12.17$, $SD = 2.18$) rather than the ECERS–R to measure classroom quality. NICHD used the Preschool Language Survey (PLS) as its measure of receptive language and the Reynell Developmental Language Scale ($M = 100.13$, $SD = 15.68$) as the Time 1 control for the receptive language test. At Time 1, NICHD did not administer the Woodcock–Johnson; the Bracken Basic Concepts Scale ($M = 9.47$, $SD = 2.78$) was used as the Time 1 control for the prereading test. At Time 1, EHS did not administer the Woodcock–Johnson; the 36-month PPVT ($M = 82.91$, $SD = 16.09$) was used as the Time 1 control for the prereading test and the Bayley Mental Index at 36 months ($M = 91.12$, $SD = 12.20$) was used as the Time 1 control for math. For NCEDL, the WJ Letter–Word was only included in the SWEEP study.

rules, and use of the scale in some settings involves omitting some items.

The NICHD Study used the Observational Record of the Caregiving Environment (ORCE). The ORCE was designed specifically for the NICHD SECCYD to assess the quality of caregiver–child interaction experienced by individual children. It measures care that is attentive and appropriately responsive, expresses positive affect and affection, is not excessively restrictive or intrusive, and offers activities believed to promote children's cognitive and social development. Observations of child-care quality were made during a single half-day when the child was 54 months of age. The quality measure used for the current project is the positive caregiving rating composite, the mean of 4-point ratings of caregivers' sensitivity/responsivity, stimulation of cognitive development, intrusiveness (reflected), and detachment (reflected). Detailed descriptions of the ORCE assessments can be found in NICHD SECC (n.d.), including coding definitions, training procedures, internal consistency, and interobserver agreement.

Children's Academic Skills

Table 3 provides descriptive information for each measure of children's academic skills for each study.

Receptive language. All studies included a measure of receptive vocabulary. Six studies used the Peabody Picture Vocabulary Test, 3rd ed. (PPVT; Dunn & Dunn, 1997), and one study (NICHD) used the Preschool Language Survey (PLS). The PPVT is a test of receptive vocabulary that relates to other measures of language, literacy, and academic achievement (Chow & McBride-Chang, 2003; Dunn & Dunn, 1997). Children are shown a set of four pictures and asked to select the picture that best represents the meaning of a word spoken by the examiner. According to the scale's authors, the α coefficient of all the items on this scale ranges from .92 to .98, with a median reliability of .94, and test–retest reliability is reported as ranging from .85 to .90.

The NICHD SECCYD administered the Reynell Developmental Language Scale (RDLS; Reynell, 1991) at 36 months and the Preschool Language Scale (PLS-3; Zimmerman, Steiner, & Pond, 1979) at 54 months. The RDLS is composed of two 67-item scales that assess verbal comprehension and expressive language; Cronbach's αs = .93 and .86. The PLS-3 measures a range of language behaviors, including vocabulary, morphology, syntax, and integrative thinking, grouped into two subscales: Auditory Comprehension and Expressive Language; Cronbach's αs = .89 and .92.

Prereading skills. Six of the seven studies (all except MAF) included the Letter–Word Identification subtest of the Woodcock–Johnson Tests of Achievement (Woodcock, McGrew, & Mather, 2001) as a

measure of pre- and emerging reading skills. Woodcock–Johnson Tests of Achievement are well-established standardized measures of academic achievement. This subtest measures word-identification skills. The child is initially asked to identify letters. The remaining items require the child to read and pronounce written words correctly. Three of the studies (GECS, NCEDL, and PCER) used the Woodcock–Johnson III, whose reliability coefficient for the 3- to 5-year-old age group ranges from .97 to .99 according to the measure's authors. Three studies (EHS, NICHD, and FACES) that began data collection earlier used the Woodcock–Johnson Revised (Woodcock & Johnson, 1990). That test has a reliability coefficient of .92 for 4-year-olds.

Early math skills. All seven studies included the Applied Problems Subtest of the Woodcock–Johnson Tests of Achievement as a measure of early math skills. The Applied Problems subtest examines the child's ability to analyze and solve math problems. Four of the studies (GECS, MAF, NCEDL, and PCER) used the Woodcock–Johnson III, whose reliability coefficient for the 3- to 5-year-old age group ranges from .92 to .94 according to the measure's authors. Three studies (EHS, NICHD, and FACES) that began data collection earlier used the Woodcock–Johnson Revised (1990). This test has a reliability coefficient of .91 for 4-year-olds.

The NICHD SECCYD used a different test at 36 months to assess school readiness. The School Readiness Composite of the Bracken Basic Concepts Scale (Bracken, 1984) consists of 51 items grouped into five categories reflecting children's knowledge of colors, letters, numbers/counting, comparisons, and shapes; the score analyzed was the percentile rank.

Control Variables

Models contain a common set of control variables; however, different instrumentation in the various studies prevented specifying the variables identically across studies. Instead, each study included whichever variable they had that was closest to the desired control. For the classroom-level analyses, the control variables were site/state, adult-to-child ratio, class size, length of school day, and teacher ethnicity (White, African American, Latino, or Other/Multi-Racial), proportion of White students in class, and proportion of poor students in class (using any measure/definition of poverty the study included). Whenever available, the models also controlled for two key interactions between: (1) the variable of interest and state/site and (2) the variable of interest and length of the school day. Table 4 presents descriptive information from each study for the classroom control variables. NA on Table 4 indicates that the study did not control that variable, because the data were not collected for the entire sample, or there was no variance (e.g., MAF and GECS did not control length of school day or state/site because all classrooms in those studies were full-day and in a single state). The last row of Table 4 lists a few additional controls that were needed for some studies because of different circumstances and sampling strategies.

Child-level analyses controlled for site/state, child gender, ethnicity of child (White, African American, Latino, or Other/Multi-Racial), years of maternal education, poverty/family income, and previous assessment score (e.g., previous PPVT score when analyzing PPVT scores). When available, the models also control for two key interactions between: (1) the variable of interest and state/site and (2) the variable of interest and poverty/family income. Table 5 presents descriptive information about the control variables for participating children. Again, NA on this table indicates that the control was not included. The final row of this table includes the few additional variables that were included to account for studies' special circumstances.

Analysis Plan

Each study was provided with detailed instructions and SAS code to spell out precisely how each model should be specified. For each of the research questions, four models were estimated, one for each of the outcomes: classroom quality, receptive vocabulary, prereading skills, and early math skills. The control variables described above were included in each model. Standard errors used in hypothesis testing were adjusted for cluster effects and other design effects based on the specific study. The analyses used hierarchical linear modeling (HLM) to adjust for dependencies in the data when multiple children from the same classroom were included. HLM was not used for EHS and NICHD because those studies included only one child per classroom. For two projects that involved stratified random sampling (GECS and FACES), the analyses weighted the data to represent the populations.

Effect sizes (*d*) were computed when statistically significant associations were obtained. The effect sizes were computed as the difference between group means divided by the standard deviation of the instrument for the sample used for development or norming. The standard deviation of the ECERS–R in

Table 4
Descriptive Information About Control Variables in Classroom-Level Models

	EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Mean (*SD*) child-to-adult ratio	NA	6.18 (2.00)	10.02 (2.19)	6.23 (2.58)	7.55 (3.35)	8.69 (3.95)	7.28 (2.59)
Mean (*SD*) class size	17.50 (5.05)	14.38 (3.12)	16.82 (4.38)	15.99 (2.35)	18.50 (5.56)	15.49 (5.73)	16.32 (4.73)
Mean (*SD*) school day hours	**	4.88 (2.04)	6.50 (0.00)	NA	4.58 (2.52)	4.86 (2.91)	NA
Mean proportion poor in class	NA	NA	0.54 (0.37)	NA	0.58 (0.32)	NA	NA
Mean proportion White in class	NA	NA	0.49 (0.37)	0.30 (0.29)	0.41 (0.37)	NA	NA
Teacher's ethnicity, *n* (%)				NA			
Latino	101 (12)	69 (22)	2 (2)		93 (14)	27 (4)	7 (9)
African American	173 (21)	101 (33)	41 (34)		74 (11)	41 (6)	25 (33)
White	518 (63)	127 (41)	75 (63)		459 (67)	556 (87)	34 (45)
Other/mixed	35 (4)	13 (4)	1 (1)		55 (8)	14 (2)	10 (13)
Missing	60	0	9		42	1	0
Other controls	Head Start versus not		GA Pre-K versus Head Start versus private	% of MAF children in class; school year	Program located in a public school versus not; full-day program or not		

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NA = not available and therefore not controlled; NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
''Proportion poor in class'' should not be compared across studies because poverty definitions varied; GECS used whether or not the family had ever received TANF since child's birth, NCEDL used 150% of federal poverty guidelines. EHS controlled full- versus half-day; 50% of classes were full day (i.e., 30 or more hours per week instead of ''school day hours''; All MAF classes meet for a full-day school (6 – 6.5 hr). MAF only collected demographic data on MAF children and not on other children in those classrooms; ''Mean Proportion White'' indicates the percent of MAF children in each classroom who are White.
**Similar variable was controlled.

Table 5
Descriptive Information About Control Variables in Child-Level Models

	EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Male (%)	442 (50)	503 (49)	336 (53)	380 (49)	1459 (49)	322 (50)	348 (52)
Mothers with less than high school education (%)	542 (61)	310 (31)	66 (14)	NA	539 (19)	40 (6)	116 (20)
Poor (%)	**	655 (68)	258 (53)	685 (89)	1605 (58)	144 (23)	378 (76)
Child's ethnicity *n* (%)							
Latino	218 (25)	346 (32)	14 (2)	117 (15)	764 (26)	35 (5)	113 (18)
African American	316 (36)	348 (35)	231 (40)	336 (43)	533 (18)	61 (10)	275 (44)
White	325 (37)	224 (23)	283 (49)	273 (35)	1200 (41)	514 (80)	187 (30)
Other/multiracial	28 (3)	73 (10)	54 (9)	59 (8)	401 (14)	29 (5)	55 (9)
Missing	0	50	48	0	68	0	37
Other controls	Head Start versus not		GA Pre-K versus Head Start versus private	School year	Program located in a public school versus not; full-day program or not		

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NA = not available and therefore not controlled; NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
''Poor'' values should not be compared across studies because poverty definitions varied; FACES used 100% of federal poverty guidelines, GECS used whether or not the family ever received TANF since child's birth, MAF used free/reduced price lunch eligibility, NCEDL used 150% of federal poverty guidelines, NICHD and PCER used 200% of federal poverty guidelines. EHS controlled family income (M = $1,847/month, SD = $1,202) rather than poverty.
**Similar variable was controlled.

this sample is 1.0 and 15 for the PPVT, PLS, and Woodcock–Johnson scales. Typically, effect sizes of $<.30$ are considered modest, .30–.60 are considered moderate, and $>.60$ are considered large (Cohen, 1988).

Missing data occurred in these longitudinal projects due to attrition and failure to complete all assessments. Missing data were imputed using multiple imputation (Rubin, 1987; Schafer, 1997; Schafer & Graham, 2002) under the assumption that missing data were ignorably missing. Given the many measures of demographic, child, family, and classroom measures, each study had sufficient information to estimate missing data accurately. We used Schafer's recommended procedure, an iterative E–M algorithm in which missing values for each variable are estimated iteratively using a logistic or multiple regression from all the other variables using the data for individuals with observed values on that variable. The prediction model is used to predict missing values, and random variability is added as the missing data are predicted. The process is repeated for each variable until the differences in predicted values across iterations are miniscule. For each study, five data sets were created in which all observed data are represented and missing data are estimated. Consequently, analyses were conducted five times, using each of the five imputation data sets. The results of these analyses were combined using the recommended procedures of Schafer (1997). The test statistics and regression coefficients were combined by averaging them across the five analyses, and the standard errors for the coefficients were combined by combining within- and between-model variability.

Results

Question 1: Degree

The first research question asked whether the highest degree attained by the lead teacher predicted classroom quality and children's academic gains, with special focus on the value of a Bachelor's degree. We conducted 27 analyses for this question (4 outcomes $\times$ 7 studies, with MAF unable to test Letter–Word). As can be seen in Table 6, only 8 of these 27 analyses provided any evidence of an association, and the direction was negative (i.e., more education associated with less positive outcome) for 2 of 8. For classroom quality, two studies (EHS and NICHD) found evidence that more educated teachers had higher quality classrooms and that quality was higher when teachers had a Bachelor's degree. One study, FACES, found that teachers with a Bachelor's degree had lower quality classrooms compared with teachers without a Bachelor's. The remaining studies found no evidence supporting an association between education and quality.

EHS found a linear association between quality and teacher's degree. Classrooms in which the teacher had a graduate degree had higher ECERS–R scores than classrooms where the teacher had a Bachelor's degree ($d = .65$). Classrooms where the teacher had a Bachelor's received higher ECERS–R scores than classrooms where the teacher had an Associate's ($d = .42$) or a high school degree ($d = .55$). There was no difference in ECERS–R scores between classrooms with an Associate's versus high school level teacher. In comparing all teachers with a Bachelor's degree or higher with those without a Bachelor's degree, EHS found evidence that quality was moderately higher when teachers had a Bachelor's ($d = .45$).

NICHD, using the positive caregiving composite from the ORCE, found that teachers with a graduate degree had classrooms with moderately higher quality scores than teachers with a Bachelor's ($d = .44$) or Associate's degree ($d = .47$). Teachers with a Bachelor's degree, in turn, had substantially higher quality classrooms than high school level teachers ($d = .68$). There was no difference between the classroom quality of teachers with Bachelor's versus Associate's degree or Associate's versus high school. For this reason, a smaller, but statistically significant difference was observed when the teachers with and without a Bachelor's degree were compared ($d = .23$).

FACES also found a significant association between teachers' education and quality. Classrooms where the teacher had a Bachelor's degree or more were rated lower on the ECERS–R than classrooms where teachers did not have a Bachelor's degree ($d = -.26$). The more refined comparison of the four levels of teacher degree was not statistically significant.

In addition, three of the seven studies reported a significant interaction between education and site. The significant interaction in the EHS, NCEDL, and NICHD studies indicated that a stronger positive association between education and quality was obtained in some sites, but a weaker and sometimes negative association was obtained in other sites, with no clear pattern to the findings.

For academic outcomes, there was very limited evidence of an association between teachers' highest degree and scores at the end of the 4-year-old year, controlling for previous skills and the other demo-

Table 6

Question 1: Educational Degree as a Predictor of Quality and Children's Outcomes

		EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Classroom quality		ECERS – R	ECERS – R	ECERS – R	ECERS – R	ECERS – R	ORCE Teacher Sensitivity	ECERS – R
Highest degree	*F*	**$F(3, 886) = 9.42^{***}$**	$F(3, 297) = 2.22$	$F(3, 127) = 0.95$	$F(2, 222) = 0.56$	$F(3, 670) = 0.77$	**$F(3, 589) = 3.57^{*}$**	$F(3, 53) = 0.15$
High school	*M*(*SE*)	**4.44 (0.19)^{a}**	4.39 (0.16)	4.30 (0.34)	NA	3.85 (0.20)	**10.44 (0.69)^{a}**	3.80 (0.65)
Associate's	*M*(*SE*)	**4.58 (0.28)^{a}**	4.16 (0.14)	4.41 (0.29)	5.03 (0.15)	3.58 (0.22)	**11.86 (0.35)^{ab}**	3.72 (0.90)
Bachelor's	*M*(*SE*)	**5.06 (0.07)^{b}**	4.03 (0.07)	4.21 (0.08)	5.07 (0.05)	3.74 (0.08)	**11.93 (0.21)^{b}**	3.60 (0.25)
Graduates	*M*(*SE*)	**5.80 (0.20)^{c}**	3.71 (0.28)	4.21 (0.37)	4.93 (0.12)	3.84 (0.17)	**12.88 (0.39)^{c}**	3.18 (0.77)
Bachelor's	*F*	**$F(1, 886) = 24.61^{***}$**	**$F(1, 301) = 5.91^{*}$**	$F(1, 127) = 2.19$	$F(1, 232) = 0.00$	$F(1, 692) = 0.38$	**$F(1, 589) = 7.40^{**}$**	$F(1, 61) = 0.12$
No	*M*(*SE*)	**4.55 (0.14)^{a}**	**4.26 (0.10)^{a}**	4.38 (0.23)	5.04 (0.12)	3.85 (0.12)	**11.51 (0.28)^{a}**	3.72 (0.56)
Yes	*M*(*SE*)	**5.06 (0.07)^{b}**	**4.04 (0.07)^{b}**	4.19 (0.08)	5.05 (0.05)	3.81 (0.08)	**12.01 (0.20)^{b}**	3.67 (0.19)
Receptive language		PPVT	PPVT	PPVT	PPVT	PPVT	PLS	PPVT
Highest degree	*F*	$F(3, 855) = 0.83$	$F(3, 290) = 0.42$	$F(3, 490) = 0.74$	$F(2, 648) = 0.80$	$F(3, 2247) = 1.65$	$F(3, 582) = 0.06$	$F(3, 58) = 0.70$
High school	*M*(*SE*)	90.00 (2.19)	86.32 (1.24)	96.92 (1.79)	NA	94.38 (1.43)	96.82 (4.52)	91.23 (3.89)
Associate's	*M*(*SE*)	89.54 (2.64)	86.24 (1.08)	94.05 (1.58)	94.43 (4.33)	92.12 (1.50)	97.72 (2.37)	93.44 (3.58)
Bachelor's	*M*(*SE*)	91.08 (0.93)	86.26 (0.42)	95.12 (0.55)	89.55 (0.72)	94.14 (0.52)	97.97 (1.21)	93.59 (0.93)
Graduate	*M*(*SE*)	94.37 (2.52)	85.08 (2.03)	93.97 (2.03)	86.32 (2.69)	95.37 (1.19)	98.60 (2.61)	96.23 (2.66)
Bachelor's	*F*	$F(1, 855) = 1.44$	$F(1, 292) = 0.29$	$F(1, 494) = 1.61$	$F(1, 678) = 0.17$	$F(1, 2250) = 2.66$	$F(1, 589) = 0.06$	$F(1, 66) = 0.46$
No	*M*(*SE*)	89.94 (1.46)	86.09 (0.73)	94.47 (1.26)	88.12 (2.29)	93.29 (0.96)	97.50 (1.80)	92.18 (2.61)
Yes	*M*(*SE*)	90.77 (0.84)	86.40 (0.39)	95.33 (0.56)	88.86 (0.72)	94.08 (0.47)	97.84 (1.11)	93.40 (0.84)
Prereading		WJ-R	WJ-R	WJ-III	NA	WJ-III	WJ-R	WJ-III
Highest degree	*F*	$F(3, 855) = 0.40$	$F(3, 290) = 0.21$	**$F(3, 490) = 4.31^{***}$**		**$F(3, 1466) = 2.78^{*}$**	$F(3, 582) = 0.14$	**$F(3, 58) = 2.85^{*}$**
High school	*M*(*SE*)	88.41 (2.20)	99.28 (1.74)	**98.54 (2.26)^{a}**		101.62 (1.84)^{a}	102.04 (3.40)	101.27 (3.35)^{a}
Associate's	*M*(*SE*)	87.60 (3.14)	98.85 (1.50)	**101.19 (1.77)^{ab}**		100.44 (1.57)^{a}	99.76 (1.82)	105.56 (3.56)^{a}
Bachelor's	*M*(*SE*)	89.45 (0.94)	99.77 (0.58)	**103.91 (0.61)^{b}**		102.08 (0.67)^{a}	100.88 (0.92)	103.06 (0.86)^{a}
Graduate	*M*(*SE*)	91.53 (3.16)	100.84 (2.72)	**110.00 (2.45)^{c}**		102.11 (1.37)^{a}	101.15 (1.98)	106.98 (2.82)^{a}
Bachelor's	*F*	$F(1, 855) = 1.15$	$F(1, 292) = 0.57$	**$F(1, 494) = 11.39^{***}$**	NA	**$F(1, 1475) = 4.81^{*}$**	$F(1, 604) = 0.33$	$F(1, 66) = 1.33$
No	*M*(*SE*)	88.33 (1.51)	99.98 (1.02)	**100.39 (1.14)^{a}**		**100.78 (1.02)^{a}**	100.12 (1.38)	104.89 (2.56)
Yes	*M*(*SE*)	89.51 (0.85)	99.67 (0.53)	**103.82 (0.62)^{b}**		**102.20 (0.65)^{b}**	100.73 (0.83)	102.59 (0.74)
Math		WJ-R	WJ-R	WJ-III	WJ-III	WJ-III	WJ-R	WJ-III
Highest degree	*F*	$F(3, 855) = 0.65$	$F(3, 290) = 0.75$	$F(3, 490) = 0.56$	$F(2, 648) = 0.70$	**$F(3, 2247) = 3.55^{*}$**	$F(3, 582) = 1.99$	$F(3, 58) = 0.59$
High school	*M*(*SE*)	87.55 (2.49)	93.68 (1.68)	95.24 (1.99)	NA	**95.05 (1.27)^{a}**	109.23 (4.12)	97.91 (3.33)
Associate's	*M*(*SE*)	84.17 (3.72)	91.12 (1.31)	98.14 (1.54)	88.19 (4.47)	**98.80 (1.38)^{ab}**	104.65 (2.14)	97.89 (3.94)
Bachelor's	*M*(*SE*)	87.72 (1.15)	92.34 (0.54)	97.37 (0.59)	92.83 (0.73)	**97.99 (0.54)^{b}**	103.42 (1.08)	98.37 (0.85)
Graduate	*M*(*SE*)	89.61 (3.97)	91.11 (2.52)	98.10 (2.21)	94.86 (2.82)	**98.30 (1.20)^{ab}**	98.88 (2.36)	100.75 (2.73)
Bachelor's	*F*	$F(1, 855) = 0.94$	$F(1, 292) = 0.24$	$F(1, 494) = 0.17$	$F(1, 678) = 1.28$	**$F(1, 2250) = 4.27^{*}$**	**$F(1, 604) = 4.73^{*}$**	$F(1, 66) = 0.01$
No	*M*(*SE*)	86.53 (1.73)	92.03 (0.91)	97.23 (1.03)	90.47 (2.45)	**96.90 (0.82)^{a}**	**105.81 (1.62)^{a}**	98.41 (2.54)
Yes	*M*(*SE*)	87.78 (1.00)	92.41 (0.50)	97.40 (0.59)	92.69 (0.73)	**98.04 (0.49)^{b}**	**103.10 (1.02)^{b}**	98.24 (0.76)

Note. ECERS – R = Early Childhood Environment Rating Scale-Revised; EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NA = not analyzed (measure not included in study, too little variance or subgroup too small); NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research; PPVT = Peabody Picture Vocabulary Test. Significant effects appear in bold. Adjusted means that are significantly different from one another ($p < .05$) have different superscripts. For MAF, the high school group was too small to analyze separately and hence they have been dropped for the "highest degree" analyses only.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

graphic characteristics. None of the seven studies found an association between the highest degree and receptive language skills, and only a few studies reported associations with reading or math.

Of the six studies that included a prereading measure (Woodcock–Johnson Letter–Word Identification), two studies reported significant differences associated with whether the teacher had a Bachelor's degree. Prereading scores were significantly, albeit modestly, higher when teachers had a Bachelor's degree in GECS ($d = .05$) and NCEDL ($d = .09$). One of those studies, GECS, also found an association between the particular degree and prereading. Two (NCEDL and PCER) found an overall association between the four levels of education and reading scores, but pairwise comparisons revealed no significance between group differences due to small effect sizes and small sample sizes in educational-level cells aside from the Bachelor's group. Finally, three studies (EHS, NICHD, and FACES) found no association between prereading skills and teacher's highest degree. A significant interaction between Bachelor's versus no Bachelor's and site in the PCER study indicated that the association between Bachelor's degree and prereading scores was stronger in some sites than others.

Five of the seven studies found no association between early math skills (Woodcock–Johnson Applied Problems) and either teacher's highest degree or whether the teacher had a Bachelor's degree. NCEDL found that children whose teacher had a Bachelor's degree or higher had slightly higher math scores ($d = .07$) than children whose teacher did not have a Bachelor's degree. This association was the basis of the significant association between the more refined levels of education and math scores in the NCEDL sample. Children's math scores were higher when teachers had a Bachelor's degree than when teachers had only a high school degree ($d = .19$). NICHD, on the other hand, found the opposite effect: Children whose teacher did not have a Bachelor's degree scored higher on the Applied Problems test than children whose teacher had a Bachelor's degree ($d = -.18$).

Question 2: Highest Education Level Among Teachers With ECE Major

The second question asked whether highest level of education predicts classroom quality and/or children's academic gains among teachers whose highest degree was in early childhood education or child development (see Table 7). We conducted 19 analyses for this question (4 outcomes × 5 studies, with MAF unable to test Letter–Word). As can be seen in Table 7, only 2 of 19 analyses provided any evidence of an association.

Of the five studies that were able to address this question, three found no association with observed quality or the child's language, prereading, or math skills. EHS found that among those who majored in early childhood education or child development, teachers with a graduate degree had higher quality classrooms than those with a Bachelor's degree ($d = .46$), who in turn had higher quality classrooms than those with an Associate's degree ($d = .57$). Furthermore, an interaction between site and teachers' education among ECE majors in EHS indicated that stronger associations were obtained in some sites and weaker associations in other sites. EHS did not find a significant association between teacher degree and any of the child outcomes in classrooms in which the teacher majored in early childhood education or child development. In addition, teacher degree among teachers with an early childhood education or child development major was significantly associated with prereading scores in NCEDL, but none of the pairwise comparisons was sufficiently large to achieve statistical significance. Furthermore, teacher degree among ECE/CD majors was not associated with quality or the child's language or math scores in the NCEDL study.

Question 3: ECE Major Among Teachers With a Bachelor's

The third question asked whether the teacher's major made a difference, among teachers with a Bachelor's degree (see Table 8). For this question, each study split the teachers whose highest degree was Bachelor's into three groups: (1) teachers who majored in early childhood education or child development, (2) teachers who majored in any type of education other than early childhood or child development, and (3) teachers who majored in anything else. We conducted 23 analyses for this question (4 outcomes × 6 studies, with MAF unable to test Letter–Word ID). As can be seen in Table 8, only 1 of 23 analyses found a significant effect for teacher major for any of the outcomes. In the FACES study, there was a significant association between teacher's major and children's spring PPVT scores; however, pairwise comparisons revealed no significant between-group differences. No other significant associations were found between teacher major and any of the outcomes of interest in any study. A significant interaction between teacher's major and

Table 7
Question 2: Highest Degree as a Predictor of Quality and Children's Outcomes Among Teachers Who Majored in Early Childhood Education or Child Development

		EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Classroom quality		**$F(2, 822) = 5.80^{**}$**	$F(2, 299) = 1.48$	NA	$F(2, 125) = 2.37$	$F(2, 314) = 1.18$	$F(2, 338) = 2.64$	NA
Associate's	$M(SE)$	**4.49 (0.24)[a]**	4.33 (0.17)		5.11 (0.16)	3.90 (0.22)	11.10 (0.62)	
Bachelor's	$M(SE)$	**5.15 (0.07)[b]**	4.04 (0.09)		5.03 (0.07)	3.92 (0.11)	12.94 (0.57)	
Graduate	$M(SE)$	**5.67 (0.19)[c]**	3.78 (0.29)		4.63 (0.18)	4.11 (0.20)	15.67 (2.55)	
Receptive language		$F(2, 793) = 1.82$	$F(2, 284) = 0.28$	NA	$F(2, 361) = 1.23$	$F(2, 951) = 2.24$	$F(2, 331) = 0.02$	NA
Associate's	$M(SE)$	88.42 (2.41)	86.92 (1.20)		95.87 (5.65)	92.47 (1.28)	97.28 (4.39)	
Bachelor's	$M(SE)$	91.18 (0.90)	86.32 (0.42)		89.24 (1.00)	94.61 (0.67)	97.64 (3.57)	
Graduate	$M(SE)$	94.88 (2.51)	86.50 (2.10)		83.68 (3.52)	96.33 (1.55)	99.03 (16.95)	
Prereading		$F(2, 793) = 0.48$	$F(2, 284) = 0.36$	NA	NA	**$F(2, 551) = 3.08^{*}$**	$F(2, 331) = 0.01$	NA
Associate's	$M(SE)$	87.62 (2.86)	99.15 (1.65)			100.85 (1.85)[a]	102.24 (3.29)	
Bachelor's	$M(SE)$	89.96 (0.93)	99.55 (0.57)			101.61 (0.98) [a]	103.38 (2.69)	
Graduate	$M(SE)$	91.60 (2.76)	101.41 (2.84)			99.92 (2.15)[a]	104.60 (12.83)	
Math		$F(2, 793) = 1.02$	$F(2, 284) = 0.49$	NA	$F(2, 361) = 0.49$	$F(2, 951) = 0.26$	$F(2, 331) = 0.09$	NA
Associate's	$M(SE)$	84.53 (3.31)	91.50 (1.44)		88.80 (5.85)	98.11 (1.48)	102.75(4.05)	
Bachelor's	$M(SE)$	88.48 (1.11)	92.42 (0.54)		93.48 (0.99)	97.76 (0.74)	104.08 (3.28)	
Graduate	$M(SE)$	91.44 (3.45)	91.58 (2.57)		95.23 (3.50)	97.56 (1.69)	107.92 (15.65)	

Note. EHS = Early Head Start; GECS = Georgia Early Childhood Study; FACES = Family and Child Experiences Survey; MAF = More at Four; NA = not analyzed (measure not included in study, too little variance, or subgroups that are too small); NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
Significant effects appear in bold. Adjusted means that are significantly different from one another ($p < .05$) have different superscripts.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

family poverty in the NCEDL study indicated that after adjusting for covariates, prereading scores were similar among children who were poor and those who were not poor if the teacher majored in early childhood education or child development; however, when the teacher did not major in any type of education (i.e., group 3 as outlined above), the children who were not poor scored lower than the poor children.

Discussion

Using seven recent, major studies of classroom-based educational programs for 4-year-olds, these analyses, taken together, do not provide convincing evidence of an association between teachers' education or major and either classroom quality or children's academic gains. Most of the analyses yielded null findings. Although there were some statistically significant associations, no clear pattern emerged. For instance, two of the studies indicated that quality was higher when the teacher had a Bachelor's degree or more, one study indicated that quality was lower when the teacher had a Bachelor's degree or more, and four studies found no association. Although the quick—and incorrect—conclusion might be that anybody can effectively teach preschool, we believe that the results from this study must be interpreted cautiously and that they raise more questions than answers. We start the Discussion by considering a few possible reasons why we did not find evidence of associations.

Why Did We Not Find Associations?

We expected to find associations between teachers' education and both classroom quality and children's outcomes. Generally, we (like the rest of the public) still think education should matter for teachers as well as for a host of other professions. Why did we not find evidence to confirm our belief? It is possible that the associations between teachers' education or major and either classroom quality or children's academic gains do exist but that we did not find evidence of associations because of limitations of the individual studies or this replicated secondary data analysis project. Although we recognize the limitations of this work as described in a later section, we do not believe that the limitations prevented us from finding associations. More likely, the lack of significant findings reflects the current reality of the field. Three different reasons for the

Table 8
Question 3: Major as a Predictor of Quality and Children's Outcomes Among Teachers Whose Highest Degree is Bachelor's

		EHS	FACES	GECS	MAF	NCEDL	NICHD	PCER
Classroom quality		NA	$F(2,70) = 0.07$	$F(1,24) = 0.16$	$F(2,168) = 0.23$	$F(2,304) = 1.52$	$F(2,158) = 0.49$	$F(2,30) = 0.12$
ECE/CD	*M*(*SE*)		4.10 (0.38)	4.21 (0.30)	5.04 (0.07)	3.92 (0.13)	3.55 (0.23)	3.5 (0.2)
Other Ed.	*M*(*SE*)		4.06 (0.43)	NA	5.08 (0.10)	3.88 (0.16)	3.61 (0.60)	3.6 (0.6)
Other	*M*(*SE*)		4.06 (0.48)	4.34 (0.50)	5.12 (0.12)	4.18 (0.26)	3.30 (0.66)	3.3 (0.7)
Receptive language		NA	$\mathbf{F(2,70) = 3.45^{*}}$	$F(1,31) = 0.19$	$F(2,510) = 0.12$	$F(2,1070) = 0.14$	$F(2,151) = 1.38$	$F(2,35) = 0.33$
ECE/CD	*M*(*SE*)		87.74 (0.72)[a]	97.19 (1.46)	88.76 (0.98)	94.57 (0.78)	93.91 (0.92)	93.9 (0.9)
Other Ed.	*M*(*SE*)		89.31 (1.86)[a]	NA	92.14 (7.49)	94.62 (1.16)	92.89 (2.19)	92.9 (2.2)
Other	*M*(*SE*)		90.57 (1.95)[a]	96.51 (2.50)	86.39 (4.67)	93.86 (3.22)	95.55 (3.03)	95.5 (3.0)
Prereading		NA	$F(2,70) = 0.77$	$F(1,31) = 0.04$	NA	$F(2,815) = 1.67$	$F(2,151) = 0.12$	$F(2,35) = 0.28$
ECE/CD	*M*(*SE*)		100.60 (1.08)	105.81 (1.81)		102.27 (0.95)	101.73 (0.85)	101.7 (0.9)
Other Ed.	*M*(*SE*)		99.60 (2.78)	NA		103.36 (1.34)	101.10 (2.36)	101.1(2.4)
Other	*M*(*SE*)		103.79 (3.28)	106.06 (3.13)		100.26 (1.49)	101.25 (2.70)	101.3 (2.7)
Math		NA	$F(2,70) = 0.29$	$F(1,31) = 1.64$	$F(2,510) = 0.13$	$F(2,1070) = 1.24$	$F(2,151) = 1.65$	$F(2,35) = 1.06$
ECE/CD	*M*(*SE*)		92.84 (1.07)	96.34 (1.43)	92.63 (0.97)	99.22 (0.92)	97.17 (0.83)	97.2 (0.8)
Other Ed.	*M*(*SE*)		93.87 (2.79)	NA	92.23 (7.67)	97.44 (1.40)	97.62 (2.25)	97.6 (2.3)
Other	*M*(*SE*)		93.40 (3.20)	94.18 (2.53)	91.84 (4.73)	103.22 (3.84)	94.18 (2.80)	94.2 (2.8)

Note. EHS = Early Head Start; FACES = Family and Child Experiences Survey; GECS = Georgia Early Childhood Study; MAF = More at Four; NA = not analyzed (measure not included in study, too little variance, or subgroups too small); NCEDL = National Center for Early Development and Learning; NICHD = National Institute of Child Health and Human Development; PCER = Preschool Curriculum Evaluation Research.
Significant effects appear in bold. Adjusted means that are significantly different from one another ($p < .05$) have different superscripts. For GECS, "other field of education" ($n = 5$) has been combined with "other"($n = 5$) due to small group sizes.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

lack of associations are considered below: the teacher preparation system, support within programs, and market forces.

First, the teacher preparation system may not have prepared teachers adequately to teach preschoolers. A recent review of K–12 teachers' education research suggests that the research base for coursework, methods, and pedagogy in teacher preparation programs is lacking (Cochran-Smith & Zeichner, 2005). This may also be true in the field of early childhood. Furthermore, the U.S. Department of Education has recently increased the standards for research on effective educational practices (Coalition for Evidence-Based Policy, 2003; Shavelson & Towne, 2003). Findings from new research using these higher standards were probably not incorporated into the teacher preparation programs when the majority of these teachers were in school over 10 years ago. Likewise, today's more educated teachers may have completed their university training at a time when math and prereading skills were deemphasized for young children.

Along these same lines, some researchers in early childhood education believe that trusting, respectful relationships between children and adults form the basis for much learning in early childhood, including academic skills (see Espinosa, 2002; Hamre & Pianta, 2005; National Research Council and Institute of Medicine, 2000; Pianta, 1999, for reviews). If this is true, it may be that this aspect of early childhood development has been underemphasized in early childhood teacher preparation programs, possibly leaving teachers with content knowledge around academic instruction, but lacking the needed skills for forming individual relationships that can serve as the base for academic learning.

We could not directly test questions about the quality or content of the teacher preparation programs because the studies in this project did not include the needed data. A careful, systematic program of research is needed to understand the impact of early childhood teacher preparation programs on teacher quality, classroom quality, and child outcomes, a conclusion reflected in recent federal funding initiatives in early childhood (U.S. Department of Education, IES, 2005). Likewise, more fine-grained research is needed to address which aspects of teachers' attitudes, knowledge, and behavior are affected by participation in higher education and in-service training. The field currently lacks information about what is taught in teacher preparation and in-service programs, how that instruction

is delivered, and the mechanisms by which it translates into classroom practice.

A second possible reason for the lack of associations is that teachers may not receive sufficient support to implement effectively what they have learned. Recent research suggests that entry-level teachers often feel overwhelmed and would appreciate monitoring or coaching during the transition from teacher preparation to actual teaching (Hart, Stroot, Yinger, & Smith, 2005). Although many early childhood educators become teachers before finishing their education, the need for support as teachers attempt to implement what they learn in school undoubtedly still applies. Furthermore, perhaps teachers feel pressure to abandon what they were taught in their teacher education programs and adhere to the school's standards and teaching strategies once they complete their education.

A third possible explanation for these largely null findings is that current market forces have stimulated at least a short-term ripple in the labor market for early education teachers. These studies contained large numbers of publicly funded classrooms (i.e., state funded prekindergarten, federally sponsored Head Start). The higher wages and benefits provided in publicly funded programs, as compared with community-based child care, may attract and retain the best and the brightest teachers *without* a Bachelor's degree (i.e., an Associate's degree or less) to the public system. On the other hand, the lower wages provided in prekindergartens as compared with elementary schools may lead the highest quality teachers *with* a Bachelor's degree to teach older children and may lead to increased turnover among those Bachelor's-level teachers who select to work with prekindergartners. These contradictory market forces may be diminishing the link between education and quality in early care and education settings.

Likewise, within school systems where most teachers, even those at the prekindergarten level, have a Bachelor's degree there may be factors that encourage the best teachers to teach older children, leaving teachers with fewer skills in the prekindergarten settings. School administrators may place the best teachers in grades in which high-stakes accountability testing occurs. The best teachers may also seek and obtain teaching positions in higher grades because those grades do not rely on discretionary funding as prekindergarten often does. Therefore, there may be negative selection in teaching prekindergarten for the highly skilled Bachelor's or higher level teachers.

Possible Misinterpretations of the Findings

These findings do not indicate that teacher quality is unimportant. Teachers' education and teacher quality are two separate albeit related constructs. Although teacher quality is sometimes defined in policy and the press solely as teachers' level of education, the concept of teacher quality is much larger, representing more than just education. Teacher quality encompasses a broad array of knowledge, skills, and behaviors. This project looked only at teachers' formal education.

By definition, teachers who provide instruction that leads to positive child outcomes are high-quality teachers. Thus, identifying and supporting high-quality teachers is important. Appropriate delivery of high-quality, stimulating preschool education is challenging and requires a great deal of skill. These data indicate that a policy of using teachers' educational attainment and/or major will not substitute for selecting teachers with the skills needed to teach at this level.

Likewise, these findings do not mean that postsecondary education *cannot* produce high-quality teachers or that *none* of the current teacher preparation programs are adequately preparing teachers. These analyses average across many different higher education programs that have produced today's early childhood educators. Many of the teachers in these studies received their degrees 10, 20, or even 30 years ago and may have received little or no coursework or supervised practica in teaching 4-year-olds as part of their teacher preparation programs. Furthermore, the studies included in this project were not designed to evaluate postsecondary education and, therefore, did not include information about these teachers' educational experiences like course content, rigor, or field placements. We can assume, however, that the quality of higher education programs varied. Nevertheless, the results from this replicated secondary data analysis project suggest that a policy of requiring a certain degree or major will not substitute for the hard work of selecting high-quality teachers and supporting them in a way that will maximize their capacity.

Finally, these findings do not indicate that teachers' education is unimportant. It is likely that the relationships among teachers' education, classroom quality, and child outcomes in preschool programs are complex. A new set of detailed studies on professional development is needed to better understand the role of teachers' education in the multifaceted early care and education system. In-depth information about course content, grades,

supervised practica experience, in-service training experiences (quantity, quality, and content), and on-the-job support, supervision and monitoring would begin to answer some of the outstanding questions. Furthermore, additional research on curricula and instructional practices is needed to ensure that professional development focuses on aspects of the early childhood program that are most likely to lead to maximum gains for children.

From a practical perspective, teachers' education in prekindergarten programs remains important because the Bachelor's degree is the established entry point into the teaching profession within a public school setting. As prekindergarten programs are increasingly administered by public schools, one could expect prekindergarten teachers to meet the same or similar teachers' education standards as required in the K–12 grades.

Study Limitations

This replicated secondary data analysis project had several strengths, including use of many highly regarded, rigorously collected data sets; common variable definitions and data analysis procedures; and inclusion of extensive control variables. Analyzed together, they allow us to address a timely issue in the policy and research arenas without worrying that the findings are artifacts of differences in definition or analysis.

Some limitations should be noted, however. We selected these quality and child outcome measures because they were contained in most of the data sets. In turn, the investigators responsible for collecting the data selected these measures because: (1) they have been linked to other important outcomes (e.g., PPVT predicts later school success), (2) they are widely used in educational research and hence norms and standards are readily available, (3) they can be administered in a reasonably short period of time, and (4) data collectors without an extensive background in assessment can learn to administer them reliably. The ECERS–R was a particularly logical choice of outcomes because it had been linked to teachers' education in previous research (see Tout et al., 2005 for a review) and five of the studies in this project have linked ECERS–R scores to children's outcomes (Howes et al., 2005; Love et al., 2003; Mashburn, 2006; Peisner-Feinberg & Maris, 2005b; Zill et al., 1997, 2003).

Nonetheless, different yet equally important outcomes could be linked to teachers' education or major. It is possible, for instance, that the benefits children experience from a highly educated teacher can be seen only in children's higher order thinking skills or when application of knowledge is measured. Likewise, while the ECERS–R is an excellent measure of global quality it has its own limitations. Many of the items are out of the control of the teacher and many of the types of instruction that today's early childhood educators value are not measured or measured in only a single item (e.g., child engagement, literacy-rich environment).

Furthermore, these data sets did not include all of the information that would be of interest in answering the primary research questions. For instance, among Associate's-level teachers, there was insufficient variance to consider the role of major. Additionally, there are some potentially important aspects of teacher education and preparation that could not be included in these analyses because they were measured inconsistently across the studies. Namely, teaching certification/licensure and quality and quantity of ongoing in-service training may contribute to classroom quality and children's academic gains, but could not be included in the present analyses. Likewise, information about coursework, grades, practica experiences, ongoing training, and support would allow for more detailed and nuanced analyses. These data sets and analyses also lack detailed measures of instructional practices and teacher behaviors that might be helpful in understanding the links between teacher education, teacher practices, and quality. Lastly, these findings apply only to classroom-based programs for 4-year-olds. We did not include data for classrooms for younger children and we did not include data from family child-care homes.

Implications

These findings have important implications for researchers, policymakers, and practitioners.

Teacher Quality is Complex

The results from this replicated secondary data analysis project demonstrate the limitations of a conceptualization of teacher quality that relies heavily on teacher degree or major. This finding is in line with recent debates in the K–12 literature regarding whether the educational attainment and degree status are the correct targets for effectively improving student achievement (Whitehurst, 2002). A new era of research is needed to address the complexity of teacher quality. Researchers should go beyond the easy-to-measure constructs of degree and major to more fully understand teacher quality

and its relationship with classroom quality and child outcomes. This recommendation is in keeping with the recent American Educational Research Association (AERA) report summarizing research on teachers' education in the K–12 system (Cochran-Smith & Zeichner, 2005). New measures are needed to capture adequately the quantity and quality of not only teacher education but also specific teacher behaviors and instructional practices in the classroom that are tied to positive child outcomes.

Likewise, we encourage policymakers and practitioners to consider more than teachers' education in ensuring high-quality teachers. Factors such as individual teacher's skills, classroom practice, and beliefs could also be taken into account in the hiring process. A comprehensive professional development system for preservice and in-service teachers could provide the knowledge, skills, and supports for teachers to provide a high-quality early education experience that can positively impact children's development.

Teacher Quality is Only One Component of a Larger Educational System

Teachers do not work in a vacuum but instead are part of a larger educational system. Classroom quality and positive child outcomes are influenced by a host of other system components. Even the most highly skilled teachers need, for example, adequate materials, curricular support, skilled teaching assistants, and a physical setting that is appropriate to meeting the needs of young children. Furthermore, they need administrators and supervisors who support the use of instructional practices recognized as leading to positive development in young children and ongoing professional development through mentoring, monitoring, and supervision (Howes, James, & Ritchie, 2003; Ramey & Ramey, 2005). Most likely, the impact of teacher quality on classroom quality and child outcomes is influenced by many other components of the early care and education system. We recommend that policymakers not emphasize one component (e.g., teachers' education) as more important than others but instead support policies that address multiple components of teacher quality and the factors that might influence quality, such as mentoring, monitoring and supervision, and accreditation of teacher preparation programs.

Teacher Quality Needs to be More Precisely Defined

We do not have the research base to identify the specific teacher attributes associated with high classroom quality or positive child outcomes. Although the (National Association for the Education of Young Children (2001; Hyson, 2003) has described teacher competencies as part of their standards for professional preparation, additional work is needed to further describe these competencies in more detail and provide an evidence base for the relationship between these competencies, quality programs, and good outcomes for children (Hyson & Biggar, 2005). With a more precise description of what a ''highly qualified'' preschool teacher knows and does, teacher preparation programs could more easily align their coursework and practica to these standards and researchers could better study the effects of teacher quality on classroom quality and child outcomes.

Conclusions

This project has addressed an important question in the field of early childhood education: Are policies that increase the educational attainment of preschool teachers likely to lead to increased classroom quality or children's academic gains? These data indicate that such policies alone are unlikely to have such effects. Instead, teachers' education must be considered as part of a system of factors that contribute to teacher quality, which in turn is related to classroom quality and children's gains. The findings should not be interpreted as an indictment of the role of education in high-quality programs for 4-year-olds. Rather, these findings can serve as a springboard that moves research and policy regarding the role of teachers' education and, more broadly, teacher quality to a new level that is increasingly multifaceted and nuanced.

Cochran-Smith and Zeichner (2005) in their AERA research panel report, make similar calls for researchers and policymakers to view teachers' education more complexly in the K–12 education system. The results from this project and the AERA research panel highlight the struggles shared across the education spectrum, from early care and education of preschoolers to higher education. We need more precise studies of specific types of preservice teacher preparation—and how they interact with other teacher and program characteristics—to produce a high-quality educational experience for children.

References

Administration for Children and Families. U.S. Department of Health & Human Services. (2002, June). *Making a difference in the lives of infants and toddlers and their*

families: The impacts of Early Head Start, Volume I: Final technical report. Retrieved December 16, 2005, from http://www.acf.hhs.gov/programs/core/ongoing_research/ehs/impacts_vol1/impacts_vol1.pdf

Barnett, W. S. (2003). *Better teachers, better preschools: Student achievement linked with teacher qualifications. Preschool Policy Matters (2)*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W. S., Hustedt, J. T., Robin, K. B., & Schulman, K. L. (2005). *The state of preschool. 2005 state preschool yearbook*. New Brunswick, NJ: The National Institute for Early Education Research. Retrieved June 21,2006, from http://nieer.org/yearbook/

Blau, D. M. (2000). The production of quality in child-care centers: Another look. *Applied Developmental Science*, *4*, 136–147.

Bracken, B. A. (1984). *The Bracken basic concept scale*. San Antonio, TX: The Psychological Corporation.

Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, *6*, 2–11.

Burchinal, M. R., Roberts, J. E., Riggins, R., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating quality of center-based child care to early cognitive and language development longitudinally. *Child Development*, *71*, 339–357.

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, *6*, 42–57.

Chow, B. W., & McBride-Chang, C. (2003). Promoting language and literacy development through parent–child reading in Hong Kong preschoolers. *Early Education and Development*, *14*, 233–248.

Coalition for Evidence-Based Policy. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Retrieved February 22, 2006, from http://www.ed.gov/rschstat/research/pubs/rigorousevid/index.html

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis of the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cost, Quality, & Child Outcomes Study Team. (1995). *Cost, quality, and child outcomes in child care centers, executive summary* (2nd ed.). Denver: Economics Department, University of Colorado at Denver.

de Kruif, R. E. L., McWilliam, R. A., Ridley, S. M., & Wakely, M. B. (2000). Classification of teachers' interaction behaviors in early childhood classrooms. *Early Childhood Research Quarterly, 15*, 247–268.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test* (3rd ed.). Circle Pines, MN: American Guidance Service.

Early, D., Barbarin, O., Bryant, B., Burchinal, M., Chang, F., Clifford, R., et al. (2005). *Pre-Kindergarten in eleven states: NCEDL's Multi-State Study of Pre-Kindergarten and State-Wide Early Education Programs (SWEEP) study*. Retrieved December 1, 2005, from http://www.fpg.unc.edu/~ncedl/pdfs/SWEEP_MS_summary_final.pdf

Early, D. M., Bryant, D., Pianta, R., Clifford, R., Burchinal, M., Ritchie, S., et al. (2006). Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly, 21*, 174–195.

Espinosa, L. M. (2002). The connections between social-emotional development and early literacy. In *Set for success: Building a strong foundation for school readiness based on the social-emotional development of young children* (pp. 30–44). Kansas City, MO: The Ewing Marion Kauffman Foundation.

Hamre, B., & Pianta, R. C. (2005). Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? *Child Development*, *76*, 949–967.

Hamre, B. K., & Bridges, M. (2004). *Early care and education staff preparation, quality, an child development: A review of the literature*. Unpublished manuscript. University of California at Berkeley.

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale* (Rev. ed.). New York: Teachers College Press.

Hart, P., Stroot, S., Yinger, R., & Smith, S. (2005). *Meeting the teacher education accountability challenge: A focus on novice and experienced teacher studies*. Mount Vernon, OH: Teacher Quality Partnership.

Henry, G. T., Henderson, L. W., Ponder, B. D., Gordon, C. S., Mashburn, A., & Rickman, D. K. (2003). *Report of the findings from the Early Childhood Study: 2001–02*. Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies.

Henry, G. T., Ponder, B. D., Rickman, D. K., Mashburn, A., Henderson, L. W., & Gordon, C. S. (2004). *Evaluation of the implementation of the Georgia Pre-K Program: Report of the findings from the Early Childhood Study (2002–03)*. Atlanta, GA: Georgia State University, Andrew Young School of Policy Studies.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D. M., Clifford, R., et al. (2007). *Ready to learn? Children's pre-academic achievement in pre-kindergarten programs*. Manuscript submitted for publication. University of California at Los Angeles.

Howes, C., James, J., & Ritchie, S. (2003). Pathways to effective teaching. *Early Childhood Research Quarterly, 18*, 104–120.

Howes, C., Whitebook, M., & Phillips, D. (1992). Teacher characteristics and effective teaching in child care: Findings from the National Child Care Staffing Study. *Child & Youth Care Forum. Special Issue: Meeting the child care needs of the 1990s: Perspectives on day care: II*, *21*, 399–414.

Hyson, M. (Ed.). (2003). *Preparing early childhood professionals: NAEYC's standards for programs*. Washington, DC: NAEYC.

Hyson, M., & Biggar, H. (2005). NAEYC's standards for early childhood professional preparation. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 283–308). Baltimore: Paul H. Brookes Publishing Co.

Lee, V., & Burkam, D. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Washington, DC: Economic Policy Institute.

Love, J. M., Harrison, L., Sagi-Schwartz, A., van IJzendoorn, M. H., Ross, C., Ungerer, J. A., et al. (2003). Child care quality matters: How conclusions may vary with context. *Child Development*, *74*, 1021–1033.

Mashburn, A. (2006). *Quality of the social and physical environments in preschool programs and changes in children's academic, language and literacy skills*. Manuscript submitted for publication. University of Virginia.

Maxwell, K. L., Feild, C. C., & Clifford, R. M. (2005). Defining and measuring professional development in early childhood research. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 21–48). Baltimore: Paul H. Brookes Publishing Co.

McLanahan, S. (Ed.). (2005). School readiness: Closing racial and ethnic gaps. *The Future of Children*, 15.

National Association for the Education of Young Children. (2001). *Standards for early childhood professional preparation: Initial licensure programs*. Retrieved November 19, 2005, from http://www.naeyc.org/faculty/pdf/2001.pdf

National Research Council (2001). *Eager to learn*: *Educating or preschoolers*. Committe on Early Childhood Pedagogy. B. T. Bowman, M. S. Donovan, & M. S. Burns, (Eds.), *Commission on Behavioral and Social Sciences and Education*. Washington DC: National Academy Press.

National Research Council and Institute of Medicine. (2000). *From neurons to neighborhoods*: *The science of early childhood development*. Committee on Integrating the Science of Early Childhood Development. J. P. Shonkoff, & D. A. Phillips (Eds.), Board on children, youth, and families, commission on behavioral and social sciences and education. Washington, DC: National Academy Press.

NICHD Early Child Care Research Network. (2002). Child-care structure → Process → Outcome: Direct and indirect effects of child-care quality on young children's development. *Psychological Science*, *13*, 199–206.

NICHD Early Child Care Research Network. (2005). Early child care and children's development in the primary grades: Follow-up results from the NICHD Study of Early Child Care. *American Educational Research Journal*, *42*, 537–570.

NICHD SECC. (n.d.). *The NICHD Study of Early Child Care and Youth Development*. Retrieved November 23, 2005, from http://secc.rti.org/

PCER Project. (n.d.) *Preschool Evaluation Curriculum Research Project*. Retrieved January 6, 2006, from http://pcer.rti.org/

Peisner-Feinberg, E., Burchinal, M., Clifford, R., Culkin, M., Howes, C., Kagan, S., et al. (2001). The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade. *Child Development*, *72*, 1534–1553.

Peisner-Feinberg, E. S., & Maris, C. L. (2005a). *Evaluation of the North Carolina More at Four Pre-kindergarten Program: Year 2 Report* (July 1, 2002–June 30, 2003). Chapel Hill, NC: FPG Child Development Institute. Retrieved December 19, 2005, from http://www.fpg.unc.edu/~mafeval/pdfs/MAF_Yr2_Rpt.pdf

Peisner-Feinberg, E. S., & Maris, C. L. (2005b). Evaluation of the North Carolina More at Four Pre-kindergarten Program: Year 3 Report (July 1, 2003–June 30, 2004). Chapel Hill, NC: FPG Child Development Institute. Retrieved December 19, 2005, from http://www.fpg.unc.edu/~mafeval/pdfs/MAF_Yr3_Rpt.pdf

Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott-Shim, M. (2001). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, *15*, 475–496.

Phillips, M., Brooks-Gunn, J., Duncan, G. J., Klebanov, P., & Crane, J. (1998). Family background, parenting practices, and the black/white test score gap. In C. Jencks & M. Phillips (Eds.), *The black/white test score gap* (pp. 103–145). Washington, DC: Brookings.

Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, *12*, 281–303.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.

Ramey, S. L., & Ramey, C. T. (2005). Creating and sustaining a high-quality workforce in child care, early intervention, and school readiness programs. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 355–368). Baltimore: Paul H. Brookes Publishing Co.

Reynell, J. (1991). *Reynell development language scales* (U.S. ed.). Los Angeles: Western Psychological Association.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2002). *Age 21 cost-benefit analysis of the Title I Chicago Child–Parent Centers*. Madison, WI: Institute for Research on Poverty.

Rubin, D. B. (1987). *Multiple imputation for non-response in surveys*. New York: Wiley.

Scarr, S., Eisenberg, M., & Deater-Deckard, K. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly*, *9*, 131–151.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Schweinhart, L. J. (2004). *The High/Scope Perry Preschool Project through age 40: Summary, conclusions and frequently asked questions*. Retrieved February 14, 2006, from http://www.highscope.org/Research/PerryProject/PerryAge40SumWeb.pdf

Shavelson, R., & Towne, L. (2003). *Scientific research in education*. National Research Council: National Academy Press.

Stipek, D. J., & Ryan, R. R. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology, 33*, 711–723.

Tout, K., Zaslow, M., & Berry, D. (2005). Quality and qualifications: Links between professional development and quality in early care and education settings. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 77–110). Baltimore: Paul H. Brookes Publishing Co.

Trust for Early Education. (2004, Fall). *A policy primer: Quality pre-kindergarten*. Retrieved February 15, 2006, from http://www.trustforearlyed.org/docs/TEE-Primer4.pdf

U.S. Department of Education, Institute of Education Sciences. (2004). *Digest for Education Statistics [Electronic version]* Table 161. Retrieved February 13, 2006, from http://nces.ed.gov/programs/digest/d04/tables/dt04_161.asp

U.S. Department of Education, Institute of Education Sciences. (2005). Notice inviting applications for grants to support education research for fiscal year (FY) 2006, Federal Register, 70(87), May 6, 2005.

U.S. Department of Health and Human Services , Administration for Children and Families, Office of Head Start. (2006). *Head Start Program Fact Sheet, Fiscal Year 2006*. Retrieved July 31, 2006, from http://www.acf.hhs.gov/programs/hsb/research/2006.htm.

Whitehurst, G. J. (2002). *Scientifically based research on teacher quality: Research on teacher preparation and professional development*. Presented to White House Conference on Preparing Tomorrow's Teachers, March 5, 2002. Retrieved March 3, 2006, from http://www.ed.gov/admins/tchrqual/learn/preparingteachersconference/whitehurst.html

Woodcock, R. W., & Johnson, M. B. (1990). *Woodcock–Johnson Psycho-Educational Battery—Revised*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson-III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Zill, N., & Resnick, G. (2005). Emergent literacy of low-income children in Head Start: Relationships with child and family characteristics, program factors and classroom quality. In D. Dickinson & S. Neumann (Eds.), *Handbook of early literacy research* (*Vol. II*). New York: Guilford Publications.

Zill, N., Resnick, G., Kim, K., McKey, R. H., Clark, C., Pai-Samant, S., et al. (1997). *Head Start FACES: Longitudinal findings on program performance. Third progress report*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., McKey, R. H., et al. (2003). *Head Start FACES 2000: A whole-child perspective on program performance. Fourth progress report*. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1979). *Preschool Language Scale (PLS)* (3rd ed.). San Antonio, TX: The Psychological Corporation.