

Experimental Impacts of a Teacher Professional Development Program in Chile on Preschool Classroom Quality and Child Outcomes

Hirokazu Yoshikawa
New York University

Diana Leyva
Davidson College

Catherine E. Snow
Harvard University

Ernesto Treviño
Universidad Diego Portales

M. Clara Barata
Instituto Universitário de Lisboa

Christina Weiland
University of Michigan

Celia J. Gomez
Harvard University

Lorenzo Moreno
Mathematica Policy Research, Princeton, New Jersey

Andrea Rolla, Nikhit D'Sa, and Mary Catherine Arbour
Harvard University

We assessed impacts on classroom quality and on 5 child language and behavioral outcomes of a 2-year teacher professional-development program for publicly funded prekindergarten and kindergarten in Chile. This cluster-randomized trial included 64 schools (child $N = 1,876$). The program incorporated workshops and in-classroom coaching. We found moderate to large positive impacts on observed emotional and instructional support as well as classroom organization in prekindergarten classrooms after 1 year of the program. After 2 years of the program, moderate positive impacts were observed on emotional support and classroom organization. No significant program impacts on child outcomes were detected at posttest (1 marginal effect, an increase in a composite of self-regulation and low problem behaviors, was observed). Professional development for preschool teachers in Chile can improve classroom quality. More intensive curricular approaches are needed for these improvements to translate into effects on children.

Keywords: preschool education, teacher professional development, Latin America, classroom quality, experiment

Supplemental materials: <http://dx.doi.org/10.1037/a0038785.supp>

In recent years, governments across the world have increased their investments in early childhood education programs (Engle et al., 2011; Myers, 2006; UNESCO, 2006; Yoshikawa & Kabay,

2014). This trend reflects the proven value of providing early learning opportunities in promoting later educational attainment and earnings (Cunha & Heckman, 2007). While the initial focus

Hirokazu Yoshikawa, Department of Applied Psychology, New York University; Diana Leyva, Department of Psychology, Davidson College; Catherine E. Snow, Graduate School of Education, Harvard University; Ernesto Treviño, School of Education, Universidad Diego Portales; M. Clara Barata, Instituto Universitário de Lisboa (ISCTE-IUL), Cis-IUL; Christina Weiland, School of Education, University of Michigan; Celia J. Gomez, Graduate School of Education, Harvard University; Lorenzo Moreno, Mathematica Policy Research, Princeton, New Jersey; Andrea Rolla, David Rockefeller Center for Latin American Studies, Harvard University; Nikhit D'Sa, Graduate School of Education, Harvard University; and Mary Catherine Arbour, Brigham and Women's Hospital, Harvard University.

The primary funder of this project was the Fundación Educacional Oportunidad in Santiago, Chile. Initial seed funding was also provided by

the Harvard Graduate School of Education, the Harvard Center on the Developing Child, the David Rockefeller Center for Latin American Studies at Harvard University, and UNICEF's Chile country office. Treviño acknowledges funding from Fondecyt Iniciación Project N°11100493. We express deep gratitude to the Chilean Ministry of Education, Kathleen McCartney, and all members of the Un Buen Comienzo teams in Chile and the United States for their contributions to the design, implementation, and evaluation of Un Buen Comienzo. We also express our appreciation to the teachers, families, school principals, and municipalities' personnel who served as partners in this project.

Correspondence concerning this article should be addressed to Hirokazu Yoshikawa, 726 Broadway, 5th Floor, New York University, New York, NY 10003. E-mail: hiro.yoshikawa@nyu.edu

for those interested in improving early childhood education in low- and middle-income countries was access, emphases are shifting to quality and equity, particularly in nations that have made large investments in access (Chavan, Yoshikawa, & Bahadur, 2013).

The evaluation of *Un Buen Comienzo* (A Good Start; UBC) is the first large-scale, randomized study of an effort to improve the quality of preschool education in South America. *Un Buen Comienzo* is a 2-year program that provides professional development to prekindergarten and kindergarten teachers in Chile, with the goal of enhancing children's language, literacy, health, and socioemotional outcomes. It was developed through extensive discussions with leaders of preschool education in Chile and pilot implementation and research in 2007. In 2008, a cluster-randomized trial was launched to determine the effects of the program on classroom quality and ultimately on multiple domains of child development. We focus on the following question: Does an intensive, 2-year teacher professional development program improve classroom quality and child language, literacy, and behavioral outcomes?

The Rationale for Investing in Quality in Preschool Education

The importance of the first years of life for later physical, cognitive, educational and social development and adult productivity and health has been established (Harvard Center on the Developing Child, 2010; Knudsen, Heckman, Cameron, & Shonkoff, 2006). Economic evidence highlights the particularly high returns to early childhood education programs, when compared to later remedial investments during the primary or secondary school years (Heckman & Krueger, 2004). A large literature across high-, middle-, and low-income nations shows that early childhood education in the preprimary year(s) can positively impact both early and later human development, particularly in cognitive domains (Aboud, 2006; Britto, Yoshikawa, & Boller, 2011; Currie, 2001; Engle et al., 2011; Myers, 1992; UNESCO, 2006; Woodhead & Oates, 2009; Yoshikawa et al., 2013).

Early childhood education might not have substantial positive effects on children's development unless it is implemented with sufficient structural and process quality (structural quality encompasses features such as class size, child-adult ratio, teacher qualifications, and physical environment, while process quality includes features of teacher-child interaction and instructional characteristics; National Institute of Child Health and Human Development (NICHD) Early Child Care Research Network, 2006). As Chile and other nations expand their early childhood education systems, concerns persist about quality. In Chile, researchers have raised concerns about process and instructional quality (Herrera & Bellei, 2002; Eyzaguirre & Le Foulon, 2001). Strasser and Lissi (2009) conducted systematic observations in kindergarten classrooms in nine schools with diverse sources of funding serving children from different socioeconomic backgrounds in the metropolitan region of Santiago (Strasser & Lissi, 2009). They found that noninstructional activities—snacks, behavior management, and recess time—occupied more than half of the overall time. Instructional activities typically focused on unstructured conversations and arts and crafts. On average, teachers spent only 5 minutes reading books with children and only 1 minute teaching the names and/or sounds of letters per observation. Ac-

tivities explicitly building vocabulary (e.g., discussing the meaning of words) were virtually nonexistent. Another study conducted in 38 Chilean schools (over 150 prekindergarten through second-grade classrooms) found that opportunities for children to develop vocabulary and concepts during instruction were rare, and almost 50% of the classroom activities followed no organized structure (Pontificia Universidad Católica de Chile, 2011). Taken together, these studies highlight some of the aspects of Chilean instructional quality that might make it challenging to implement a professional development program. At the same time, Chile is in a position of relatively advanced economic development that makes it feasible to implement large-scale changes in preschool education with adequate human and financial resources. Hence, Chile offers both challenge and promise as a country in which to evaluate the impact of an intensive teacher professional development program.

Multiple Foci for Quality Improvement of Preschool Education in Chile

The Government of Chile has identified early childhood development policy as a key priority in efforts to reduce economic inequality. Shortly after taking office in 2006, President Michelle Bachelet named an advisory council to come up with a proposal for establishing a comprehensive program of early childhood development for Chile. The resulting program—*Chile Crece Contigo*, or “Chile Grows with You”—incorporated comprehensive services for socially vulnerable children from birth to school entry (Ministerio de Planificación, 2007; Peralta, 2011; Vegas & Santibáñez, 2010). The expansion of preschool under Bachelet continued a progression begun in the two prior decades, with coverage for 4-year-olds at 80.4% in 2012 (Ministerio de Educación, 2012), relatively high for Latin America (Schady, 2013).

Stakeholders, preschool leaders and educators in Chile, in the process of developing UBC, cited major areas in which preschool education needed to be improved. These included support for a) language and literacy development; b) improvements in classroom behavior management; and c) the coordination of early childhood education with family health care. UBC incorporated these three areas of quality improvement. The principal mechanism through which the program was intended to have an effect on children's development was classroom quality, in domains of instructional and emotional support as well as classroom organization. In this article, we report impacts on these three dimensions of classroom quality, as well as children's language, literacy, and socioemotional skills (aggression, impulse control, and positive social behavior).

Targeting Classroom Instruction to Support Language and Literacy Development

Early childhood language development is the foundation for the oral language and reading skills that strongly predict success in primary education and later life (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Scarborough, 2001; Snow, Burns, & Griffin, 1998). Evidence suggests that teachers could be better supported to improve multiple aspects of children's language and literacy development in Chile. Teacher professional development interventions in the United States have successfully improved the quality of language instruction and thereby improved

children's language and early literacy skills (Beck & McKeown, 2007; Biemiller & Boote, 2006; Dickinson & Smith, 1994; Hargrave & Sénéchal, 2000; Juel, Biancarosa, Coker & Deffes, 2003; Justice, Kaderavek, Fan, Sofka, & Hunt, 2009; Landry, Anthony, Swank, & Monseque-Bailey, 2009; Lonigan & Whitehurst, 1998; Neuman & Cunningham, 2009; Silverman, 2007). Many are based on interactive or dialogic reading principles, including a well-known program tested in the U.S. Head Start preschool program (public preschool for children in poverty; Whitehurst et al., 1994). Two recent experimental evaluations demonstrate that the combination of detailed feedback or didactic coursework and in-classroom observation and coaching were more effective in improving preschool teachers' classroom language and literacy practices than feedback, coursework, or coaching alone (Landry et al., 2009; Neuman & Cunningham, 2009).

In Latin America, a program providing tutoring, classroom support, and parent involvement in promoting children's language and literacy skills resulted in large gains in concepts about print and letter identification among kindergarten children in Costa Rica (Rolla, Arias, Villers, & Snow, 2006). The approach of the UBC program was based in part on this program, which itself was adapted from a professional development program carried out in low-income districts in the state of Ohio (Porche, Pallante & Snow, 2012). UBC was also based on research-based principles for enriching preschool language and literacy environments (Snow et al., 1998). Importantly, although the program encouraged and supported the use of specific language-building strategies, it did not include prescribed frequency or specified sequences of these activities during the day or week, as such structured curriculum features were ruled out in the stakeholder discussions during the program development phase. We further discuss the role of curricula in early childhood education later in this section.

There is little evidence about the intensity of on-site professional development required to impact children's developmental outcomes. In North America, in-service professional development incorporating both didactic instruction and weekly or biweekly contact with mentors or coaches has resulted in gains in preschool classroom quality (instructional, organizational, and/or emotional climate dimensions, depending on the study), as well as children's cognitive and behavioral development (Bierman et al., 2008; Landry et al., 2009; Neuman & Cunningham, 2009; Raver et al., 2009; Zaslow, Tout, Halle, Whittaker, & Lavelle, 2010). Gains have been apparent in some, though not all, of these programs after only 1 year of implementation. Such programs have varied in content focus, with different classroom processes and instructional content targeted for improvement, depending on the domain of child skills. The existing evidence suggests that important elements of teacher professional development include a coaching component, a focus on language, and content-specific strategies (Landry et al., 2009; Neuman & Cunningham, 2009).

Targeting Classroom Behavior Management

Socioemotional skills are central to children's school readiness and adaptation and play a critical role in academic success. Children benefit more from instruction in the classroom when they are better at self-regulating their emotions and behaviors, less likely to be distracted, and show low levels of aggressive and disruptive behaviors and high levels of prosocial behaviors. Such children are

more likely to be positively viewed by peers and teachers and show higher levels of readiness across academic and social domains than children with lower levels of these skills (Bierman et al., 2008; Birch & Ladd, 1997; Blair, 2002; McClelland, Morrison, & Holmes, 2000; Raver, 2002).

Many teacher professional development programs in the United States have targeted classroom behavior management, positive emotional climate, and scaffolding of children's self-regulation to reduce behavior problems. Evaluations of these classroom-based programs have demonstrated positive effects on children's socioemotional and executive function development in the preschool and primary school years (Brown, Jones, LaRusso, & Aber, 2010; Diamond, Barnett, Thomas & Munro, 2007; Raver et al., 2009; Webster-Stratton, Reid, & Hammond, 2004). In the Caribbean region, the Incredible Years teacher training program for behavior management adapted for Jamaica produced small to moderate reductions in children's conduct problems, as reported by independent observers, teachers, and parents (Baker-Henningham, Scott, Jones, & Walker, 2012). In Chile, there is growing interest in training teachers in classroom management, yet to date no published studies evaluating such programs.

Coordination of Early Childhood Education With Health Services

The domains of physical health, cognitive development, and socioemotional development are intertwined especially closely in the first years of life (Shonkoff, 2010). Integration across education and health is widely recommended in early childhood programs (Chan, 2013); such integration has shown some evidence of effectiveness in improving health outcomes (e.g., Ludwig & Miller, 2007). The UBC program aimed to coordinate early education and health services across preschools and community health centers. Due to constraints on space, impacts on the key outcomes of the health component of the intervention will be published in a separate report.

Standards and Curricula in Preschool Education in Chile

In 2005, the Chilean Ministry of Education developed a set of eight learning objectives to be used by teachers as guidelines for their instruction of preschool children (Ministerio de Educación, 2005). This set of learning standards included identity (e.g., self-esteem and emotion understanding) and verbal expression (oral and written). Because in Chile there is a strong belief in the principle of curricular freedom (Peralta, 2011), these national standards are intentionally broad, and do not specify the content, frequency, or sequence of learning activities—a common characteristic of curricula in the United States. Teachers in Chile are expected to identify the individual academic and emotional needs of children and their communities, and based on this information to individualize and adapt the learning standards to their own classrooms (Peralta, 2012). Attempts by the Ministry of Education to make this national framework more structured (e.g., by including a monitoring tool to assess progress in achieving the learning standards) have faced strong criticisms from some scholars and educators in Chile. Some of the concerns center on potential homogenization of teacher practices or excessive emphasis on

academics rather than play (Peralta, 2012), concerns that are not specific to Chile. In most Latin American countries many stakeholders advocate for solely play-based preschool education, and relatively few support more targeted and specific instructional strategies, particularly for accountability purposes. Given this context, the developers of the UBC program refrained from imposing a curriculum with prescribed dosage or frequency of instructional or classroom strategies.

Development of Un Buen Comienzo

The UBC program emerged from interest during the first Bachelet Administration (2006–2010) in increased investment in preschool education in Chile. A board of Chilean policymakers, leaders and teachers in preschool education (from the Ministry of Education and the three major organizations involved in the provision of early childhood care and education: Junta Nacional de Jardines Infantiles or JUNJI, Integra, and Hogar de Cristo) was established and met monthly for 18 months. This board, or *Mesa Técnica Interinstitucional*, specified the goals of the UBC program; chose teacher professional development as the overarching approach; and developed a pilot intervention. In addition, actors from one municipality (*comuna*) of Santiago, including municipal administrators, school principals, and preschool teachers, were centrally involved. These local actors were key in providing information on the needs of the preschool education system in Santiago. This resulted in an intervention that not only reflected the latest research on professional development and early childhood development, but also had the buy-in and commitment of stakeholders in Chilean early childhood education. Three main resulting innovations of UBC for the Chilean context included: a) the intensity of the professional development, including the implementation of in-classroom coaching; b) the training of all adults in the classrooms (teachers and aides) in the same topics and strategies, and c) the collaborative nature of the development of the program.

The intervention was first piloted in 2007 in four schools with prekindergarten and kindergarten classrooms in one lower-income municipality of Santiago. A total of 22 classrooms with 325 children ages 4–6 participated. Half the schools were assigned to the intervention, and half to a control condition. We piloted all aspects of the intervention as well as measures. Based on this work, the intervention was revised, and some measures were either omitted or revised prior to the full experimental impact evaluation of UBC, which began in 2008.

We expected that improvement in quality of classroom practices—the initial and proximal target for intervention—would be apparent in prekindergarten classrooms after 1 school year of support provided by UBC, and in kindergarten classrooms after 2 years of support. We also hypothesized that the program would positively impact children’s language, literacy, and socioemotional skills at the end of the 2-year program, among children recruited at the beginning of prekindergarten. Finally, we investigated whether associations of the UBC treatment with child outcomes might be mediated by levels of classroom quality.

Method

Procedures

Research design. We estimated the causal impacts of the UBC program using a cluster-randomized design, with random

assignment of schools to either the UBC intervention (“Full UBC” condition or a much-reduced intervention (“Comparison” condition). Schools were randomly assigned within municipality. The full sample of 64 schools was recruited over 3 consecutive years (2008–2010). This number was chosen based on an initial multi-level power analysis (Moreno & Lugo-Gil, 2007; Schochet, 2008). We projected a minimum detectable effect size of .60 for classroom-level outcomes and .20 for child outcomes, based on relevant prior studies (e.g., Raver et al., 2009). The assumptions for child minimum detectable effect (MDE) were: a) intraclass correlation coefficients (ICCs) of .10; b) .05, two-tailed significance level; c) an average of two classrooms per school, 30 children per classroom; and d) projected attrition of schools of 1/6 or 17%. Clustering (ICCs) values were based on preliminary analysis in which ICCs were computed at the school level for public schools in 10 low-income municipalities in Santiago based on the 1999 national academic assessment scores (Chilean fourth-graders’ reading and math skills, the Sistema de Medición de la Calidad de la Educación, or SIMCE). Ultimately our obtained ICCs for child outcomes were below .10.

The results presented here include data from a pretest assessment (child outcomes and classroom quality); an end-of-prekindergarten assessment (classroom quality); and a posttest, end-of-kindergarten assessment (classroom quality and child outcomes).

Selection of municipalities in Santiago de Chile. A purposive sample of municipalities was drawn, in a two-stage process. First, to be invited to participate in the intervention and our study, each municipality had to show all of the following criteria: a) a high proportion of at-risk children, as defined below, in the primary grades (a minimum of 20% of the children enrolled, though most municipalities had substantially higher percentages); b) a minimum of eight schools with prekindergarten and kindergarten classrooms; and c) location in the metropolitan area of Santiago. At-risk children were identified using a national measure collected by the Chilean government for all households, which included family income, parent education, and whether the family was a beneficiary of government social/health benefits. Other characteristics, such as the absence of other school-based interventions (particularly language development interventions) being implemented at the same time, were also taken into consideration when selecting municipalities. Fourteen municipalities met these selection criteria and were invited to apply to participate in the UBC program and its evaluation.

Second, interviews with municipal representatives from mayors’ offices and departments of education and health were conducted to make sure the goals of the program were clear, to explain the evaluation design, and to answer any questions. In this stage, six municipalities were ultimately selected. Additional criteria were also considered in the selection: the total number of schools available to contribute to our target of about 60, based on our power analysis; geographic spread across the city of Santiago; and responsiveness of the municipality leadership. Three cohorts of children were aggregated over a 4-year time period as follows: Cohort 1 from one municipality (starting in 2008), Cohort 2 from two additional municipalities (starting in 2009), and Cohort 3 from three additional municipalities (starting in 2010).

To understand the generalizability of our sample, we conducted analyses comparing schoolchildren in the six selected municipal-

ities to those in municipalities that were eligible but not selected. These analyses showed that there were no significant differences in the national academic test scores (SIMCE scores measuring Chilean fourth-graders' reading and math skills) between schoolchildren living in municipalities that participated in the study and children from other candidate municipalities. Similarly, no significant differences were found in students' household socioeconomic risk. Hence, our sample of municipalities was representative of other municipalities in Santiago with similar characteristics.

Program services provided in the Full UBC and Comparison conditions. This program was a collaboration among three organizations: (a) a Chilean private educational foundation that housed the intervention team that implemented the program in schools randomized to the Full UBC condition; (b) a U.S. research institution that supported the design and evaluation of the program; and (c) a Chilean research institution that conducted data collection and some analyses. The UBC program consisted of 12 modules overall (6-month-long modules per year) combining didactic (content-focused strategies) and coaching components, and provision of a library of roughly 100 books per classroom. Teachers began training 1 month after the beginning of each school year. Each module consisted of four weekly activities, beginning with a 4-hr didactic workshop to introduce a particular topic and the corresponding instructional strategies (e.g., supporting children's predictions in read-alouds). This workshop was followed by two coaching sessions over the next 2 weeks of each module. During the first coaching session, the coach modeled for the teacher and aide the strategies introduced at the workshop. In the second coaching session, the teacher and aide either implemented the strategies in the classroom with the coach observing, or complemented the strategy with the coach. Hence, coaches visited classrooms twice per month. Every 2 months, a group reflection session at the school took place to discuss the successes and challenges of the module's topic and strategies.

Each coaching session included three phases: (a) a brief meeting between the coach and the teacher or aide to plan and share the activity plan, (b) the implementation of the activity plan in the classroom, and (c) an immediate postobservation meeting to discuss what went well and what could be improved. Coaches went through training before visiting schools. The training involved "conceptual" and "practical" components. The conceptual component included familiarization with the goals and content of the UBC project. This training was given by expert coaches and lasted one week. The practical component involved two visits to schools during which new coaches "shadowed" expert coaches in the classrooms to learn how the UBC coaching was done. In addition, expert coaches observed the first, and, if needed, additional coaching sessions of new coaches and gave feedback. Most coaches had a master's degree in early childhood education or a related discipline and prior classroom experience.

During Cohort 1's first year of implementation of UBC, the coaches did not always model for teachers and aides in the first session; teachers and aides implemented the strategy while coaches observed, and coaches returned for a second session in which they provided feedback to the teacher and aide on the previous session. End-of-year focus groups indicated that teachers and aides needed more scaffolding. Thus, for Cohort 1's second year of implementation of UBC and for both years of implemen-

tation of Cohorts 2 and 3, the first coaching session within each module involved modeling and immediate feedback.

The UBC modules addressed the domains of preschool quality identified as central by Chilean stakeholders, policymakers and educators:

1. Oral language and early literacy development. Teachers were trained on strategies to develop children's vocabulary, oral production and comprehension, and emergent writing skills. Vocabulary and oral skills were developed in the context of book-reading sessions. Teachers were trained in selecting teachable words (i.e., high-utility, multiple-meaning words) from a book, reading the book and discussing the meaning of these new words with children, and then providing opportunities for children to use these words in other classroom activities. The use of open-ended questions and language modeling techniques (repeating and extending what children said; summarizing and predicting content of stories) were encouraged in interactions with children. "Word walls" were used by teachers to help children become aware of print and phonemic features of words. Children were provided with frequent opportunities in the classroom to use emergent writing to communicate information to peers, teachers, and parents. Six of the 12 modules included topics related to language and literacy.
2. Socioemotional development. Teachers were trained in behavior management strategies (e.g., creating a system to keep track of student participation in class), establishing a positive classroom climate, and individual case management for children with challenging behaviors (e.g., identifying children who needed individualized attention, making sure that the classroom is arranged to maximize their opportunities to learn and focus). Five of the 12 modules included coverage of these topics.
3. Coordination of early childhood education with health services. Teachers were equipped with specific skills and materials to address health problems affecting preschool-aged children in Chile, which include respiratory illnesses and lack of sufficient well-child visits (Ministerio de Salud, 2006; Centers for Disease Control, 2004). Ten of the 12 modules included coverage of these health topics.

The Comparison condition was a much-reduced version of the Full UBC program. In schools assigned to this condition, 10 books (rather than 100) were distributed per classroom, and one workshop on self-care and stress reduction was provided for teachers and aides.

Recruitment of schools, randomization of schools, and study sample. Officials in the departments of education of all participating municipalities engaged the directors of schools that met our selection criteria (in Chile, each municipality has its own department of education). The recruitment rate for schools was 98.5% across all cohorts. Teachers participated in the study at the request of their school principals. The recruitment rate for teachers was 99.1%. Recruitment rates reflect the number of eligible partici-

pants (parents/teachers) affiliated with a school participating in the study who agreed to be part of the project.

Within each of the six municipalities, a public lottery took place to randomly assign schools within each municipality to the Full UBC condition (32 schools) or Comparison condition (32 schools). Because the intervention was at the school level, classrooms in the same school were always assigned to the same condition, to avoid contamination across conditions within schools (e.g., adoption of aspects of UBC in the comparison condition; Murnane & Willett, 2010). Successful random assignment should result in groups that are not statistically significantly different in their pretest characteristics measured before the intervention. To confirm that the randomization process was successful, we conducted *t* tests by experimental condition at the school level on the following teacher, child, and classroom pretest characteristics: teacher years of experience and degree qualifications; children's age and gender; and outcomes at the classroom- and child-level at pretest. These analyses yielded no statistically significant differences at the .05 level across 11 variables tested. This pattern suggests that the random assignment to conditions was successful in producing comparable groups.

A total of 64 schools, 90 classrooms, 119 teachers, 94 aides and 1,876 4-year-old children participated in the study (32 schools, 51 classrooms, 66 teachers, 54 aides, and 1,033 children in the Full UBC condition and 32 schools, 39 classrooms, 53 teachers, 40 aides and 843 children in the Comparison condition). Cohort 1 included one municipality and six schools; Cohort 2 included two municipalities and 29 schools; and Cohort 3 included three municipalities and 29 schools. Children were 53.5 months of age at baseline, on average ($SD = 3.7$ months), and 48.9% were male. Teachers were 45.5 years of age on average, with a range of 27 to 60 years; 13% had taught less than 5 years, while a striking 53% had taught more than 15 years. Forty-four percent had had postgraduate education. Postgraduate qualification in Chile is not equivalent to a master's degree. It is a 1-year professional development course taken at a university and leading to a diploma.

Overall, 1.9% (36 of 1,876) of children changed condition, of whom six children changed conditions between the beginning and the end of prekindergarten and 30 between the end of prekindergarten and the end of kindergarten. Specifically, 2.3% (24 of 1,033) of children assigned to the Full UBC condition moved to the Comparison condition, while 1.4% (12 of 843) of children assigned to the Comparison condition moved to the Full UBC condition. Because the status changes were rare, no adjustments for crossovers were implemented in the analysis.

Receipt of Full-UBC-condition services. We calculated teachers' and aides' participation in the intervention across three sessions per module (one training session and two coaching sessions; a total of 36 sessions). Overall, teachers and aides attended an average of 79% of the UBC sessions ($SD = 0.20$). Specifically, prekindergarten teachers attended 89% of UBC sessions in the first year ($SD = 0.14$, range = 0.39–1.0). Kindergarten teachers attended 80% in the second year ($SD = 0.11$, range = 0.5–1.0; note that all teachers experienced 2 years of intervention, but our end-of-first-year assessments are in prekindergarten classrooms, while our posttest assessments are in kindergarten classrooms). Prekindergarten aides attended 79% of sessions in the first year ($SD = 0.19$, range = 0.22–1.0), and kindergarten aides attended 69% of session in the second year ($SD = 0.16$, range = 0.39–1.0).

Coaches delivered the training according to the UBC guidelines (Mendive, Weiland, Yoshikawa, & Snow, 2014).

Data collection procedures. Parental permission to participate in the study and parent questionnaires on children's socio-emotional outcomes were collected in schools, at small-group meetings, or at home by the evaluation team, who were mostly master's students or graduates in education and the social sciences. Trained assessors administered the questionnaires and read through the questions upon request from parents who had low literacy levels. If parents did not know how to write, the assessors wrote the answers for them. The recruitment rate for families was 98%. Introduction and consent were conducted by the evaluation team during a parent-teacher meeting at the beginning of the prekindergarten year. The number of children whose parents did not consent to participate was 37. Of these, 19 were in the Full UBC group and 18 in the Comparison group. *T* tests showed no statistically significant differences between consenters and non-consenters on available child demographic variables (child age and gender).

Background information on teachers was collected directly from them. Data on child age and gender were collected from school records, double-checked with parental questionnaires and medical records, and resolved with further inquiry as needed. Teachers were asked to fill out questionnaires about children's socioemotional skills and were given approximately 1 week to complete them. Assessors also filled out surveys about these skills after assessment sessions.

Classrooms were videotaped at pretest, end-of-first year and posttest data collection periods (with the interim and posttest tapings occurring toward the end of each school year). In Chile, the school year starts in March and ends in December. Thus, pretest assessments took place between March and May, and posttest assessments took place between October and December of the following year. Trained assessors videotaped an entire randomly selected school day (approximately 4 hours) for each classroom. Assessors were trained to be unobtrusive to the extent possible in the classroom while videotaping. Classroom-level assessments at pretest and end-of-first year involved the same prekindergarten teachers, whereas assessments at the posttest involved a combination of the same (39% who looped with their children) and new (61%) kindergarten teachers. We followed the children, not the teachers, over a 2-year period. Prekindergarten and kindergarten teachers received training over 2 years.

Child outcome data used in the current report were collected by the evaluation team, supervised by doctoral and postdoctoral researchers, at two time points, pre- and posttest. To give children time to adapt to the classroom setting, pretest assessments began 2 weeks after school was in session. Pretest data were collected before randomization in Cohorts 2 and 3 and immediately after randomization in Cohort 1 (8 out of the full sample of 64 schools were in Cohort 1) because of delays in establishing the data collection protocols for that first cohort.

Measures

Classroom outcomes. To measure classroom quality, we used a Spanish version of the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008), which focuses on interactions between students and teachers. The CLASS measure is

predictive of a host of children's academic and nonacademic outcomes in prekindergarten and early elementary school in the United States (Burchinal et al., 2008; Burchinal, Vandergrift, Pianta, & Mashburn, 2010; Howes et al., 2008; Mashburn et al., 2008). The CLASS measure consists of a 7-point scale with higher scores indicating higher quality. A score of 7 is considered "excellent," 5 is "good," 3 is "minimal," and 1 is "inadequate." Criteria to select four 20-min classroom video segments (cycles) for coding were based on the general live observation procedure guidelines included in the CLASS manual (Pianta et al., 2008). No meaningful differences have been found between classroom video coding and live coding for this measure (Pianta et al., 2008). Five Chilean coders were trained in the CLASS in the United States and then developed reliability using videos of Chilean students. Coders were blind to the intervention condition of the videos they coded. Reliability calculation immediately after training showed an average pairwise kappa of .89. During the coding phase, several reliability checks (mean of 4) were conducted for each coder against a master coder (average kappa .94). Reliability coefficients were calculated using Cohen's kappa with adjacent reliability criteria, which means that cases where the difference between coders was zero or one point were considered reliable (Pianta et al., 2008).

As this study represented the first administration of the CLASS assessment in Latin America, we conducted Confirmatory Factor Analysis (CFA) to examine whether the psychometric structure of the CLASS as collected in Chile duplicated the structure found in North American studies using this measure. We found support for a factor structure consisting of three subscales that is similar to that of the CLASS preschool version in the United States (Leyva et al., *in press*). The subscales were: Emotional Climate (consisting of positive climate, negative climate, regard for student perspectives, and teacher sensitivity), Instructional Support (consisting of concept development, quality of feedback and language modeling), and Classroom Organization (consisting of behavior management, productivity, and instructional learning formats).

Child outcomes. Children's language and early literacy skills were assessed using the Woodcock-Muñoz Language Survey, Revised Spanish Form (WMLS-R; Woodcock, Muñoz-Sandoval, Ruef, & Alvarado, 2005). The Picture Vocabulary subtest was used to examine receptive language skills, and the Letter-Word Identification and Dictation subtests were used to examine early literacy skills. These subtests have high levels of internal reliability and validity (Schrack et al., 2005) and in our sample showed Cronbach's alpha estimates of .76 to .97. Raw scores were used in all analyses because of the relatively restricted range of age of the sample.

Children's socioemotional skills were assessed using parents', teachers', and child assessors' reports across several domains: prosocial and positive behaviors, impulse control, attention, and externalizing and internalizing behaviors. Parents and teachers rated children's prosocial and externalizing behaviors, while teachers also rated children's internalizing behaviors, impulse control, and attention. Items were drawn from the Early Development Instrument (EDI; Janus & Offord, 2007), the Teacher Observation of Child Adaptation (TOCA-R; Werthamer-Larsson, Kellam, & Wheeler, 1991), and the Social Competence Scale, Teacher Version (Conduct Problems Prevention Research Group [CPPRG], 1990) and adapted for Chile. The child assessors rated positive

behaviors, impulse control, and attention immediately after administering the child assessment battery using the Task Orientation Questionnaire (TOQ; adapted from Smith-Donald, Raver, Hayes & Richardson, 2007). An online appendix contains details about the items drawn from each scale and the informants (teachers, parents, and assessors).

In creating our socioemotional outcome variables, we randomly split the sample of children in two halves, and conducted exploratory factor analysis in one half and then confirmatory factor analyses in the other. We found adequate fit for a two-factor model (fit statistics: chi-square = 125.17, $p < .001$; RMSEA = 0.056; CFI = 0.98; TLI = 0.95). A first factor included Prosocial and Positive Behaviors (9 items), and a second factor, which we refer to as Self-Regulation and Low Problem Behavior (27 items), included impulse control, attention, and externalizing and internalizing behaviors (with externalizing and internalizing reverse coded, so the construct is positively valenced). However, despite the adequate fit statistics, the reliability of the Prosocial and Positive Behaviors scale was unacceptably low ($\alpha = .27$ at pretest and $.35$ at posttest). Therefore we use only the Self-Regulation and Low Problem Behavior scale in the impact analyses ($\alpha = .73$ at pretest and posttest). Sample items on the Prosocial and Positive Behaviors scale were: "Gets along with other children" and "Shows pleasure in accomplishment and active task mastery." Sample items on the Self-Regulation and Low Problem Behavior scale were: "Breaks things on purpose" and "Feelings are easily hurt." See online appendix for details on items included in these two scales.

Covariates. We included in our impact models covariates that controlled for child and teacher characteristics. This is recommended practice in the analysis of cluster-randomized trials in schools to increase precision of the impact estimates (Bloom, Richburg-Hayes, & Black, 2007). In analyses of classroom-level outcomes, we controlled for child age (school average) and gender (school proportion). Child age was measured in months as the average for children in the classroom on the day of pretest assessment, and proportion male constituted the gender variable. We also controlled for teacher demographic characteristics that are considered important in the literature on teacher-child interactions (see Bierman et al., 2008), and also of relevance in the Chilean context: teacher age (in years), private school teaching experience (a binary indicator), teaching experience (in years), and teacher postgraduate education (a binary indicator). All teacher variables were obtained from teacher surveys. In analyses of classroom-level outcomes, we also included the pretest score of the school quality measure. This was entered at the school level (that is, at the level of randomization), as is recommended in the analysis of cluster-randomized trials (Raudenbush, Martinez, & Spybrook, 2007). In analyses of child-level outcomes, we controlled for child age and gender (at the individual level) and teacher demographics, and included the child's pretest score on the corresponding outcome measure (language, literacy, or socioemotional) at the school level, again as recommended in the literature on cluster-randomized trials (Hedges & Hedberg, 2007).

Data Analysis Strategy

To estimate program impacts on classroom quality and child outcomes, we used multilevel linear regression models that ac-

counted for the nesting of students within classrooms (for the child-level analysis) and classrooms within schools (for both the individual- and the classroom-level analysis; Bloom et al., 2007; Hedges & Hedberg, 2007; Raudenbush et al., 2007). For all analyses, we pooled data across our cohorts. Because cohort was collinear with municipality (a total of 6 municipalities nested within 3 cohorts), we adjusted for municipality rather than cohort. We used two-tailed alpha of .05 to assess statistical significance.

For child-level impact analyses, we used the following equation:

$$\begin{aligned} Outcome_{jklm} = & \beta_0 + \beta_1(Treat)_{lm} + \beta_2(Prestest)_{lm} + \beta_3(X)_{klm} \\ & + \beta_4(M)_m + \beta_5(Teacher)_{klm} + (\delta_{lm} + \gamma_{klm} + \varepsilon_{jklm}) \end{aligned} \quad (1)$$

where the subscripts j , k , l , and m refer to students, classrooms, schools, and municipalities respectively; *Outcome* is the child-level outcome at the end of kindergarten; *Treat* is a school-level, dichotomous variable set equal to 1 if the classroom was in a Full UBC prekindergarten and to 0 if the classroom was in a Comparison prekindergarten; *Prestest* is the school-level beginning pretest score of the particular outcome; *X* is a vector of child-level characteristics (age and gender, aggregated to the classroom level); *M* is a vector of five dichotomous variables indicating which of six municipalities the school was located; *Teacher* is a vector of five classroom-level teacher covariates (teacher age, private school teaching experience, less than 5 years of teaching experience, between 5 and 14 years of teaching experience, and postgraduate education, all dichotomous except for teacher age); δ is a school-level random intercept; γ is a classroom-level random intercept; and ε is a student-level random error term. β_1 is the estimate of the impacts of UBC on the examined outcome. Note that not all child-level analyses required a classroom-level random intercept (e.g., when the intracluster correlation at the classroom was zero) and accordingly, a classroom-level random intercept was left out of models for those outcomes (specifically, the language and literacy outcomes).

Model specification for classroom-level analyses was similar to equation (1), except that the outcome was at the classroom level and there was no student-level error term. We calculated effect sizes in both child- and classroom-level analyses by dividing the β_1 parameter by the standard deviation of the control group on the relevant outcome measure at posttest.

As a robustness check, we also used an alternative model specification—ordinary least squares regression with correction of the standard errors (Huber-White) for clustering. This approach is more commonly used in economics in cluster-randomized trials (Murnane & Willett, 2010). For the child-level impacts, the robust standard errors accounted for the clustering of children in both classrooms and schools. For the classroom quality impacts, the robust standard errors accounted for the clustering of classrooms within schools. We utilized this approach to ensure that our conclusions were robust to decisions about clustering adjustments in our analysis. For all analyses, we used the STATA statistical software package.

Missing data and completion rates. A total of 6.7% of classrooms ($N = 6$) were lost to the study between pretest and end of prekindergarten. In the Full UBC group, the rate was 5.9% (3/51); in the Comparison group, the rate was 7.7% (3/39). A total

of 20% of classrooms ($N = 18$) were lost to the study between pretest and posttest. In the Full UBC group the rate was 23.5% (12/51) and in the Comparison group the rate was 15.3% (6/39). Missing classroom data were mainly due to difficulties scheduling videotaping sessions in these classrooms.

A total of 17.5% of children ($N = 328$) were lost to the study between pretest and posttest. In the Full UBC group, the rate was 19.4% (200/1,033); in the Comparison group, the rate was 15.2% (128/843). Missing child data were mainly due to the following: family issues, illness, or change of residence out of the municipality or the Santiago metropolitan region.

We had complete classroom data and teacher/child demographic data on 85 classrooms at the pretest (94.4% completion rate): 76 classrooms at the end of prekindergarten (84.4% completion rate), and 72 classrooms at the posttest (80% completion rate). On average, we had complete child outcome data and teacher/child demographic data on 1,584 (language and literacy; 84% completion rate), and 1,204 (self-regulation and low problem behavior; 64.1% completion rate) children at the pretest, and 1,354 (language and literacy; 72.1% completion rate), and 1,078 (self-regulation and low problem behavior; 57.5% completion rate) children at the posttest.

We conducted logistic regression analyses predicting a dichotomous variable representing missingness for any posttest variable, with our 14 pretest characteristics as predictors: three representing the CLASS subscales at pretest; our three pretest child language outcomes; our one pretest child behavior outcome; child age; child gender; and five teacher characteristic variables representing teacher age, experience (less than 5 years; between 5 and 14 years with 15 or more as reference category); teacher private school experience; and teacher postgraduate degree qualification. Missing cases at posttest were significantly more likely to be male; more likely to have teachers with between 5 and 14 years of experience; and less likely to have teachers with private-school experience. We also entered interactions between each of these characteristics and the treatment condition assignment variable. Of these, two reached statistical significance (just above chance levels): having a teacher with less than 5 years teaching experience and having a teacher with postgraduate qualification were more strongly associated with missingness in the full UBC condition than in the comparison condition.

The analysis sample uses only cases with complete data on all outcomes. However, we checked the sensitivity of our results to missing data problems by employing a dummy strategy for missing covariates. We entered a zero if a case was missing a given covariate and created a dummy variable set equal to one if data were missing for a covariate and zero otherwise. This approach allowed us to include all cases with nonmissing outcome data. This strategy was one of the two that performed best in a recent simulation study of different approaches for addressing missing data problems in randomized controlled trials (Puma, Olsen, Bell, & Price, 2009).

Results

We present results in two sections. First, we present the end-of-prekindergarten and end-of kindergarten impact estimates on classroom quality. Second, we present the end-of-kindergarten impacts on children's outcomes.

Classroom Quality Impacts

Instructional Support scores were consistently in the low range (1.74–1.80), whereas Emotional Support and Classroom Organization were in the middle range (4.63–4.72 and 4.14–4.29, respectively). This pattern is also common in the U.S. CLASS studies, though Instructional Support was lower in the Chile data than in the United States (Pianta et al., 2008). Table 1 summarizes the classroom-quality impacts at the end of prekindergarten and at posttest. The third column in the table corresponds to the adjusted mean (adjusted for covariates) for classrooms in the Full UBC condition, while the fourth column corresponds to the adjusted mean in the Comparison condition. The fifth column in the table represents the difference between the adjusted means for Full UBC versus Comparison group classrooms, while the sixth column displays level of statistical significance of the group difference. The last column is the Full UBC and Comparison group difference expressed as an effect size. Consistent with the consensus in the education and social sciences, we interpret effects in the .10 to .30 range as small; effects in the .30 to .60 range as moderate; and effects in the .60 and greater range as large (Hill, Bloom, Black, & Lipsey, 2007).

As shown in the first set of columns in Table 1, we found positive impacts on all three variables measuring prekindergarten classroom quality at the end of the first year. Intervention effects were positive and statistically significant for Emotional Support and Classroom Organization ($p < .05$) and positive and marginally significant for Instructional Support ($p < .10$). One effect size was in the large range; full UBC classrooms scored 0.81 standard deviation higher than Comparison classrooms on Emotional Support. Effect sizes for Classroom Organization and Instructional Support were in the moderate range (0.45 and 0.43, respectively).

As shown in the second set of columns in Table 1, we also found positive intervention effects for two of the three variables measuring kindergarten classroom quality at the end of the second year of intervention. Intervention effects were positive and statistically significant for both Emotional Support and Classroom Organization ($p < .05$), with effect sizes in the moderate range (0.37 and 0.43, respectively). The impact on Instructional Support did not reach statistical significance. Note that pretest and end-of-first year classroom assessments involved the same teachers, whereas posttest classroom assessments involved many new kindergarten teachers.

The effect sizes using our alternative estimation approach—OLS regressions with cluster corrections—were nearly identical to

those using multilevel models for first-year and second-year impact analyses. Findings were also nearly identical when we adjusted for missing data using the missing data dummy strategy as described above (Puma et al., 2009).

Child-Level Impacts

Table 2 summarizes the results of the impact analyses for child outcomes at the posttest. There were no significant effects of the UBC program on children's language, literacy, or socioemotional skills at the end of kindergarten, with the exception of a small positive impact of marginal significance ($p < .10$; $d = .16$) on self-regulation and low problem behavior (that is, an increase in self-regulation and reduction in problem behavior). As with the classroom quality impacts, results were nearly identical using OLS regression with cluster corrections, and with missing data adjustments (the missing data dummy strategy).

Indirect Associations of UBC With Child Outcomes Through Classroom Quality

Significant indirect (mediational) associations can be present even in the absence of a direct treatment effect on an outcome (Preacher, Zyphur, & Zhang, 2010). Hence, we also tested whether there were any indirect associations of experimental condition, through each of the three scales of the CLASS, on each of the four child outcomes, using the Preacher et al. (2010) approach. We did not find evidence to support indirect associations above chance.

Discussion

We report results from the first large-scale randomized evaluation of a program to improve the quality of preschool education in South America. We assessed impacts of Un Buen Comienzo on classroom quality at the end of the first and second year of the 2-year program, and child outcomes at the end of the second year. Rather than addressing the question of what impact preschool education has relative to no preschool education (what some have called a “first-generation” question in global early childhood development), this study addresses the second-generation question of how preschool education can be improved (Engle et al., 2011). Improving quality is an urgent global challenge in early childhood education, now that many low- and middle-income countries are investing more in this crucial developmental period (Britto et al.,

Table 1
Impact of the UBC Intervention on Classroom Quality Outcomes at the End of Prekindergarten and Posttest

Classroom quality outcome	N	End of prekindergarten impact				Posttest impact						
		Adj. mean full UBC group	Adj. mean comparison group	Diff.	Sig.	Effect size	N	Adj. mean full UBC group	Adj. mean comparison group	Diff.	Sig.	Effect size
Emotional support	76	4.810	4.420	0.390	***	0.814	72	4.831	4.566	0.265	*	0.376
Instructional support	76	1.804	1.605	0.199	**	0.438	72	1.842	1.813	0.029		0.056
Classroom organization	76	4.435	4.131	0.304	*	0.459	72	4.340	3.990	0.341	*	0.432

Note. Adj = adjusted mean for covariates; Diff = difference between adjusted mean in full UBC group and adjusted mean in comparison group; this is equivalent to parameter estimate (beta).

* = $p < .05$. ** = $p < .10$. *** = $p < .001$.

Table 2
Impact of the UBC Intervention on Children's Language, Literacy, and Socioemotional Skills at Posttest

Child outcome	N	Adj. Mean full UBC group	Adj. Mean comparison group	Diff	Sig.	Effect size
Language and early literacy skills						
Vocabulary	1,409	26.909	26.553	-0.356		-0.07
Letter-word identification	1,344	12.681	13.241	0.560		0.09
Early writing	1,311	9.987	9.928	-0.059		-0.03
Socioemotional skills						
Self-regulation and low problem behavior scale	1,078	4.002	3.909	0.093	*	0.161

* $p < .10$.

2011; Chavan et al., 2013; Engle et al., 2011; Myers, 2006; UNESCO, 2006).

During this launch phase of the Un Buen Comienzo program, it consisted of a combination of didactic workshops, in-classroom coaching and mentoring, and group reflection, and focused on language, literacy, socioemotional skills, and coordination of health services. We expected that improvement in classroom practices—the initial and proximal target for intervention—would be apparent after 1 school year of support provided by UBC. We found evidence, in fact, that observed classroom practices, measured through the CLASS assessment, improved as a result of the intervention. We found a large and statistically significant positive effect of the UBC program after 1 year of intervention on emotional support provided in the classroom. We also found a moderate and statistically significant increase in classroom organization. In addition, the UBC program resulted in an increase in instructional support provided by teachers at a marginal level of statistical significance (note that our study was powered to detect small effects on child outcomes and large effects on classroom quality). At the end of the second year, we found positive impacts in the moderate range for emotional support and classroom organization (effect size .43), but no significant impact on instructional quality. Importantly, analysts who coded the videotapes were blind to intervention condition and reliable based on interrater reliability calculations.

The UBC program targeted classroom instruction through its didactic and coaching components. This combination, with a content focus on language, has proven effective in improving early literacy practices in U.S. preschool classrooms. Two recent experimental trials with multiple conditions designed to tease apart these elements showed that the combination was more effective in improving practices than the didactic or coaching component alone (Landry et al., 2009; Neuman & Cunningham, 2009). The UBC program had in common with these programs such a combination of support, with each of the monthly sequences focusing both on didactic content knowledge (one session per month) and in-classroom observation/coaching with feedback (two sessions per month). This package of supports may have explained the moderate impacts on Classroom Organization (including instructional time on task, clear instructions, and lesson preparation), as measured by the CLASS assessment. However, the impacts on Instructional Support, while moderate at the end of the first year in prekindergarten classrooms, were small and nonsignificant at the end of the second year in kindergarten classrooms. This may be because of some differences between the instructional features

emphasized in the CLASS compared to the specific literacy strategies that were at the core of the UBC theory of change. Dosage analyses, in a fidelity of implementation study coded from the videotapes in both conditions, show that the number of minutes of UBC-specific literacy strategies increased significantly by the end of the second year (effect size of .69), with decreases in minutes of non-UBC literacy strategies such as rote oral routines, which do not show evidence of effectiveness in the research literature (Mendive et al., 2014).

Teachers' emotional support of children may also have improved for several reasons. The intervention included a focus on behavior management, in particular strategies for effectively addressing child disruptive behavior by providing individualized attention as necessary. In addition, teachers' perceptions that they were receiving support in their work may have increased due to the regular contact with coaches in the program. Both of these factors may have led to the increases in shared activities, warm and respectful interactions, and positive emotions and expectations, all features of this CLASS subscale (Pianta et al., 2008).

In addition to the experimental impacts, this study provides useful data on classroom quality under usual conditions in Chilean public prekindergarten and kindergarten. The pretest means showed that instructional support, as is often the case in U.S. public preschool studies, was particularly low (averaging between 1 and 2, between "inadequate" and "minimal," a bit lower than in recent U.S. data; Mashburn et al., 2008). Emotional climate and classroom organization were in the medium-quality range, similar to commonly observed U.S. means (which are typically in the 4 to 5 range). These results are in accord with prior studies of Chilean preschool quality using measures like the Early Childhood Environment Rating Scale or ECERS (Strasser & Lissi, 2009; Villalón, Suzuki, Herrera, & Mathiesen, 2002), and provide new evidence supporting the urgency of efforts to improve classroom quality in Chile.

UBC did not have a significant effect on children's language or literacy skills at the end of the 2-year program. This is surprising given that the classroom practices for which we find evidence of impact are features of observed classroom quality that have been linked to improvements in children's cognitive outcomes, over time (Burchinal et al., 2008; Howes et al., 2008; Mashburn et al., 2008). Recent studies show, however, that the associations between classroom quality and child outcomes are generally small in magnitude (Burchinal et al., 2010), and thus even moderate to large impacts on classroom quality may not translate to statistically significant impacts on children. There was one positive impact of

marginal significance on our measure of socioemotional skills (a composite of impulse control, attention, and externalizing behaviors with externalizing reverse coded; effect size .16).

Why did a combination of content-focused strategies in professional development and coaching twice a month not produce more robust child impacts? In the United States, the combination of developmentally specific curricula and coaching has produced positive effects on children's early skills (Zaslow et al., 2010). The UBC program did not provide some aspects of content-focused, developmentally specific curricula that have proven successful in North American studies, such as prescribed frequency and dosage of strategies (Beck & McKeown, 2007; Biemiller & Boote, 2006; Dickinson & Smith, 1994; Hargrave & Sénéchal, 2000; Juel et al., 2003; Justice et al., 2009; Lonigan & Whitehurst, 1998; Silverman, 2007). In the initial stakeholder meetings to develop the program, such prescribed dosage was ruled out. As we noted earlier, challenges exist in Chile to curricula that prescribe dosage (Peralta, 2012).

To further understand the role of dosage in our program impacts we conducted a study of the fidelity of implementation of UBC (Mendive et al., 2014). In this study, we examined dosage and adherence in both the full UBC and comparison groups using classroom videotaped data. We found that the effects of the UBC program were positive and statistically significant on both dosage and adherence. However, the absolute increase in time spent on targeted language and literacy instruction were not large (an increase of less than 15 minutes per observed day), which might explain the lack of significant effects on children's outcomes.

Another possible explanation for the absence of impacts on children's skills is related to children's school attendance. Absenteeism in Chilean preschools is high; more than half of children miss more than 10% of school during the year, with the average of nearly a quarter of days missed (Arbour et al., 2014). Thus, the positive effect of UBC on classrooms may have been attenuated by high absenteeism producing lower-than-intended child-level dosage.

Several limitations of this study warrant mention. First, the study's sample is limited in its generalizability to six low-income municipalities in Santiago. Although our analyses comparing general school quality (as proxied by fourth-grade standardized achievement) in these municipalities compared to other low-income municipalities in the city showed no significant differences, our municipalities may have differed in other ways from these other contexts. Second, although we conducted missing-data sensitivity analyses using a preferred method (Puma et al., 2009) and found no important differences when comparing to our main results, this does not completely rule out threats to internal validity brought about by missing data patterns. Finally, it is possible that the instruments selected were not sensitive enough to capture the specific changes that the program produced on classroom quality and children's skills. For example, the CLASS instrument does not capture specific literacy strategies targeted by UBC (e.g., requesting summaries of stories and predictions about what will happen next in the story). Others have mentioned that this lack of attention to content-specific instructional strategies used in the classroom might account for the generally small associations of the CLASS with child measures (Burchinal et al., 2010). Another example is the Woodcock-Muñoz Picture Vocabulary outcome, which sam-

ples children's word knowledge but was not targeted at the specific words children learned in the classrooms through this intervention.

Conclusion

Our findings suggest that classroom quality in Chile can be improved by incorporating didactic and coaching supports on a weekly basis into in-service training. In the UBC program, teachers and aides not only learned and practiced strategies to enhance children's learning experiences, but more importantly were continuously supported by coaches in improving their practices in the classroom. In-service professional development with intensive coaching is rare in preprimary education in Latin America; paired with an emphasis on content expertise, such models have shown promising results in North America (Zaslow et al., 2010).

Some of the lessons learned from our findings from this first school-randomized evaluation of a teacher professional development program in a Latin American country correspond to decisions made in implementation during, before, and after the trial. First, we learned the value of working collaboratively with the stakeholders in the country to design an intervention that is meaningful and aligned with the needs of the community: buy-in at different levels, from teachers, to school principals, to municipality mayors, was vital for the success of a program. Second, the degree of flexibility to adapt and evolve the program may differ across experimental and postexperimental phases of quality improvement in early childhood development: during the randomized trial, we did not make substantial changes to the program when implemented in different communities; however, in its current version, the UBC program has integrated continuous quality improvement with networked learning collaboratives (Berwick, 2003; Bryk, Gomez, & Grunow, 2011) to improve the intervention and promote sustainability in scale-up. Quasiexperimental analyses show encouraging patterns of child language (Treviño, Godoy Ossa, Yoshikawa, Snow, & Marzolo, 2014) relative to the control group of the original experiment. Third, an observational fidelity of implementation study (based on recoding of our videotaped classroom data) can aid in understanding why classroom quality effects might not have translated to child impacts (Mendive et al., 2014). Fourth, taking into account absenteeism and attendance when implementing an intervention program may be critical; evidence suggests significant positive impacts of the intervention on early literacy for children with the highest levels of attendance (Arbour et al., 2014). In sum, the pathway to improvement of a sector of education with demonstrated changes in child learning may involve multiyear efforts, of which a randomized trial may occupy only a segment.

References

- Aboud, F. E. (2006). Evaluation of an early childhood preschool program in rural Bangladesh. *Early Childhood Research Quarterly, 21*, 46–60. <http://dx.doi.org/10.1016/j.ecresq.2006.01.008>
- Arbour, M. C., Yoshikawa, H., Murnane, R. J., Snow, C. E., Weiland, C., Barata, M. C., & Treviño, E. (2014). *Experimental impacts of a preschool intervention in Chile on children's language outcomes: Moderation by student absenteeism*. Manuscript under review.
- Baker-Henningham, H., Scott, S., Jones, K., & Walker, S. (2012). Reducing child conduct problems and promoting social skills in a middle-income country: Cluster randomised controlled trial. *The British Journal*

- of *Psychiatry*, 201, 101–108. <http://dx.doi.org/10.1192/bjp.bp.111.096834>
- Beck, I., & McKeown, M. (2007). Increasing young low-income children's oral vocabulary through rich and focused instruction. *The Elementary School Journal*, 107, 251–271. <http://dx.doi.org/10.1086/511706>
- Berwick, D. M. (2003). Disseminating innovations in health care. *The Journal of the American Medical Association*, 289, 1969–1975. <http://dx.doi.org/10.1001/jama.289.15.1969>
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, 98, 44–62. <http://dx.doi.org/10.1037/0022-0663.98.1.44>
- Bierman, K. L., Domitrovich, C. E., Nix, R. L., Gest, S. D., Welsh, J. A., Greenberg, M. T., . . . Gill, S. (2008). Promoting academic and social-emotional school readiness: The Head Start REDI program. *Child Development*, 79, 1802–1817. <http://dx.doi.org/10.1111/j.1467-8624.2008.01227.x>
- Birch, S. H., & Ladd, G. W. (1997). The teacher–child relationship and children's early school adjustment. *Journal of School Psychology*, 35, 61–79. [http://dx.doi.org/10.1016/S0022-4405\(96\)00029-5](http://dx.doi.org/10.1016/S0022-4405(96)00029-5)
- Blair, C. (2002). School readiness. Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, 57, 111–127. <http://dx.doi.org/10.1037/0003-066X.57.2.111>
- Bloom, H., Richburg-Hayes, L., & Black, A. R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis*, 29, 30–59. <http://dx.doi.org/10.3102/0162373707299550>
- Britto, P., Yoshikawa, H., & Boller, K. (2011). Quality of early childhood development programs and policies in global contexts. *Social Policy Report*, 25, 1–31.
- Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving classroom quality: Teacher influences and experimental impacts of the 4Rs Program. *Journal of Educational Psychology*, 102, 153–167. <http://dx.doi.org/10.1037/a0018160>
- Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. T. Hallinan (Ed.), *Frontiers in sociology of education* (pp. 127–162). Amsterdam, Netherlands: Springer Netherlands. http://dx.doi.org/10.1007/978-94-007-1576-9_7
- Burchinal, M., Howes, C., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. *Applied Developmental Science*, 12, 140–153. <http://dx.doi.org/10.1080/10888690802199418>
- Burchinal, M., Vandergrift, N., Pianta, R., & Mashburn, A. (2010). Threshold analysis of association between child care quality and child outcomes for low-income children in pre-kindergarten programs. *Early Childhood Research Quarterly*, 25, 166–176. <http://dx.doi.org/10.1016/j.ecresq.2009.10.004>
- Centers for Disease Control. (2004). *Global school-based health survey Chile*. Retrieved from <http://www.cdc.gov/gshs/countries/americas/chile.htm>
- Chan, M. (2013). Linking child survival and child development for health, equity, and sustainable development. *The Lancet*, 381, 1514–1515. [http://dx.doi.org/10.1016/S0140-6736\(13\)60944-7](http://dx.doi.org/10.1016/S0140-6736(13)60944-7)
- Chavan, M., Yoshikawa, H., & Bahadur, C. (2013). *The future of our children: Lifelong, multi-generational learning for sustainable development* (Report of the Workgroup on Early Childhood Development, Education and the Transition to Work). New York, NY: U.N. Sustainable Development Solutions Network.
- Conduct Problems Prevention Research Group (CPPRG). (1990). *Social Competence Scale (Teacher version)*. University Park, PA: Pennsylvania State University.
- Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, 97, 31–47. <http://dx.doi.org/10.1257/aer.97.2.31>
- Currie, J. (2001). Early childhood education programs. *The Journal of Economic Perspectives*, 15, 213–238. <http://dx.doi.org/10.1257/jep.15.2.213>
- Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science*, 318, 1387–1388. <http://dx.doi.org/10.1126/science.1151148>
- Dickinson, D. K., McCabe, A., Anastopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology*, 95, 465–481. <http://dx.doi.org/10.1037/0022-0663.95.3.465>
- Dickinson, D., & Smith, M. K. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*, 29, 104–122. <http://dx.doi.org/10.2307/747807>
- Engle, P. L., Fernald, L. C., Alderman, H., Behrman, J., O'Gara, C., Yousafzai, A., . . . Global Child Development Steering Group. (2011). Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *The Lancet*, 378, 1339–1353. [http://dx.doi.org/10.1016/S0140-6736\(11\)60889-1](http://dx.doi.org/10.1016/S0140-6736(11)60889-1)
- Eyzaguirre, B., & Le Foulon, C. (2001). La calidad de la educación Chilena en cifras. *Estudios Públicos*, 84 (primavera), 85–204.
- Hargrave, A., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15, 75–90. [http://dx.doi.org/10.1016/S0885-2006\(99\)00038-1](http://dx.doi.org/10.1016/S0885-2006(99)00038-1)
- Harvard Center on the Developing Child. (2010). *Early experiences can alter gene expression and affect long-term development* (Working Paper No. 10). Retrieved from http://developingchild.harvard.edu/resources/reports_and_working_papers/working_papers/wp10/
- Heckman, J., & Krueger, A. (2004). *Inequality in America*. Cambridge: MIT Press.
- Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials. *Educational Evaluation and Policy Analysis*, 29, 60–87. <http://dx.doi.org/10.3102/0162373707299706>
- Herrera, M. O., & Bellei, Y. C. (2002). *¿Qué se sabe sobre la calidad de la educación parvularia chilena?* Santiago, Chile: Universidad de Concepción and UNICEF.
- Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2007). *Empirical benchmarks for interpreting effect sizes in research*. New York, NY: MDRC.
- Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. *Early Childhood Research Quarterly*, 23, 27–50. <http://dx.doi.org/10.1016/j.ecresq.2007.05.002>
- Janus, M., & Offord, D. (2007). Development and psychometric properties of the Early Development Instrument (EDI): A measure of children's school readiness. *Canadian Journal of Behavioural Science*, 29, 1–22. <http://dx.doi.org/10.1037/cjbs2007001>
- Juel, C., Biancarosa, G., Coker, D., & Deffes, R. (2003). Walking with Rosie: A cautionary tale of early reading instruction. *Educational Leadership*, 60, 12–18.
- Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher-child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools*, 40, 67–85. [http://dx.doi.org/10.1044/0161-1461\(2008/07-0098\)](http://dx.doi.org/10.1044/0161-1461(2008/07-0098))
- Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building

- America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 10155–10162. <http://dx.doi.org/10.1073/pnas.0600888103>
- Landry, S., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology*, 101, 448–465. <http://dx.doi.org/10.1037/a0013842>
- Leyva, D., Weiland, C., Barata, M. C., Yoshikawa, H., Snow, C. E., Treviño, E., & Rolla, A. (in press). Teacher-child interactions in Chile and their associations with kindergarten outcomes. *Child Development*. <http://dx.doi.org/10.1111/cdev.12342>
- Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*, 13, 263–290. [http://dx.doi.org/10.1016/S0885-2006\(99\)80038-6](http://dx.doi.org/10.1016/S0885-2006(99)80038-6)
- Ludwig, J., & Miller, D. L. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *The Quarterly Journal of Economics*, 122, 159–208. <http://dx.doi.org/10.1162/qjec.122.1.159>
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79, 732–749. <http://dx.doi.org/10.1111/j.1467-8624.2008.01154.x>
- McClelland, M. M., Morrison, F. J., & Holmes, D. L. (2000). Children at risk for early academic problems: The role of learning-related social skills. *Early Childhood Research Quarterly*, 15, 307–329. [http://dx.doi.org/10.1016/S0885-2006\(00\)00069-7](http://dx.doi.org/10.1016/S0885-2006(00)00069-7)
- Mendive, S., Weiland, C., Yoshikawa, H., & Snow, C. E. (2014). *Opening the black box: Intervention fidelity in a randomized trial of a preschool teacher professional development program in Chile*. Manuscript under review.
- Ministerio de Educación. (2005). *Bases curriculares de la educación parvularia*. Santiago, Chile: Author.
- Ministerio de Educación. (2012). *Datos educación parvularia*. Santiago, Chile: Author.
- Ministerio de Planificación. (2007). *Chile Crece Contigo—Sistema de protección integral a la infancia*. Santiago, Chile: Author.
- Ministerio de Salud. (2006). *Encuesta: Calidad de vida*. Santiago, Chile: Author.
- Moreno, L., & Lugo-Gil, J. (2007). *Design consultation and data collection for a study to build better preschool education in Chile*. Princeton, NJ: Mathematica Policy Research, Inc.
- Murnane, R., & Willett, J. (2010). *Methods matter: Improving causal inference in educational research*. New York, NY: Oxford University Press.
- Myers, R. G. (1992). *The twelve who survive: Strengthening programs of early childhood development in the third world*. London, UK: Routledge.
- Myers, R. G. (2006). *Quality in programs of early childhood care education*. Paris, France: UNESCO.
- National Institute of Child Health and Human Development Early Child Care Research Network (NICHD ECCRN). (2006). Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist*, 61, 99–116. <http://dx.doi.org/10.1037/0003-066X.61.2.99>
- Neuman, S. B., & Cunningham, L. (2009). The impact of professional development and coaching on early language and literacy instructional practices. *American Educational Research Journal*, 46, 532–566. <http://dx.doi.org/10.3102/0002831208328088>
- Peralta, M. V. (2011). Early childhood education and public care policies in Chile. *International Journal of Child Care and Education Policy*, 5, 17–27.
- Peralta, M. V. (2012). Un análisis del desarrollo curricular de la educación parvularia Chilena: ¿Cuánto se ha avanzado? *Docencia*, 48, 59–71.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *The classroom assessment manual, Pre-K*. Baltimore, MD: Brookes.
- Pontificia Universidad Católica de Chile. (2011). *Informe final. Alfabetización en establecimientos Chilenos subvencionados*. Facultad de Educación. Retrieved from <http://www.educacion2.udp.cl/seminarios/201109/ppt/Alfabetizacion-Informe-final.pdf>
- Porche, M., Pallante, D., & Snow, C. E. (2012). Professional development for reading achievement: Results from the Collaborative Language and Literacy Instruction Project (CLLIP). *The Elementary School Journal*, 112, 649–671. <http://dx.doi.org/10.1086/665008>
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15, 209–233. <http://dx.doi.org/10.1037/a0020141>
- Puma, M. J., Olsen, R. B., Bell, S. H., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials* (NCEE 2009–0049). Washington, DC: Institute for Education Sciences, U.S. Department of Education.
- Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, 29, 5–29. <http://dx.doi.org/10.3102/0162373707299460>
- Raver, C. C. (2002). Emotions matter: Making the case for the role of young children's emotional development for early school readiness. *Social Policy Report*, 16, 1–20.
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M. W., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 77, 302–316. <http://dx.doi.org/10.1037/a0015302>
- Rolla, A., Arias, M., Villers, R., & Snow, C. E. (2006). Evaluating the impact of different early literacy interventions on low-income Costa Rican kindergarteners. *International Journal of Educational Research*, 45, 188–201. <http://dx.doi.org/10.1016/j.ijer.2006.11.002>
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for research in early literacy* (pp. 97–110). New York, NY: Guilford Press.
- Schady, N. (2013, October). *Early childhood development in the Latin America/Caribbean region*. Presentation at the Inter-American Development Bank, Washington, DC.
- Schochet, P. Z. (2008). Statistical power for random assignment evaluations of education programs. *Journal of Educational and Behavioral Statistics*, 33, 62–87. <http://dx.doi.org/10.3102/1076998607302714>
- Schrank, F. A., McGrew, K. S., Ruef, M. L., Alvarado, C. G., Muñoz-Sandoval, A. F., & Woodcock, R. W. (2005). *Overview and technical Suppl.* (Bateria III Woodcock-Muñoz Assessment Service Bulletin No. 1). Itasca, IL: Riverside.
- Shonkoff, J. P. (2010). Building a new biodevelopmental framework to guide the future of early childhood policy. *Child Development*, 81, 357–367. <http://dx.doi.org/10.1111/j.1467-8624.2009.01399.x>
- Silverman, R. (2007). A comparison of three methods of vocabulary instruction during read-alouds in kindergarten. *The Elementary School Journal*, 108, 97–113. <http://dx.doi.org/10.1086/525549>
- Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the preschool self-regulation assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22, 173–187. <http://dx.doi.org/10.1016/j.ecresq.2007.01.002>
- Snow, C. E., Burns, M. S., & Griffin, P. (Eds.). (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

- Strasser, K., & Lissi, M. R. (2009). Home and instruction effects on emergent literacy in a sample of Chilean kindergarten children. *Scientific Studies of Reading, 13*, 175–204. <http://dx.doi.org/10.1080/10888430902769525>
- Treviño, E., Godoy Ossa, F., Yoshikawa, H., Snow, C. E., & Marzolo, M. (2014). *Quasi-experimental impacts of continuous quality improvement in a preschool teacher professional development in Chile: A propensity score matching approach*. Manuscript under review.
- UNESCO. (2006). *Education for all: Strong foundations*. Paris, France: Author.
- Vegas, E., & Santibáñez, L. (2010). *The promise of early childhood development in Latin America and the Caribbean*. Washington, DC: The World Bank.
- Villalón, M., Suzuki, E., Herrera, M. O., & Mathiesen, M. E. (2002). Quality of Chilean early education. *International Journal of Early Years Education, 10*, 49–59.
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33*, 105–124. http://dx.doi.org/10.1207/S15374424JCCP3301_11
- Werthamer-Larsson, L., Kellam, S., & Wheeler, L. (1991). Effect of first-grade classroom environment on shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19*, 585–602. <http://dx.doi.org/10.1007/BF00937993>
- Whitehurst, G. J., Epstein, J. N., Angell, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86*, 542–555. <http://dx.doi.org/10.1037/0022-0663.86.4.542>
- Woodcock, R. W., Muñoz-Sandoval, A. F., Rief, M., & Alvarado, C. G. (2005). *Woodcock Muñoz Language Survey—Revised*. Itasca, IL: Riverside Publishing.
- Woodhead, M., & Oates, J. (2009). *Early childhood in focus*. Milton Keynes, UK: Open University.
- Yoshikawa, H., & Kabay, S. B. (2014). *The evidence base on early childhood care and education in global contexts* (Background paper, 2015 Global Monitoring Report on Education for All). Paris, France: UNESCO.
- Yoshikawa, H., Weiland, C., Brooks-Gunn, J., Burchinal, M., Espinosa, L., Gormley, W., . . . Zaslow, M. J. (2013). *Investing in our future: The evidence base on preschool education*. New York, NY: Foundation for Child Development and Ann Arbor, MI: Society for Research in Child Development.
- Zaslow, M. J., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). *Toward the identification of features of effective professional development for early childhood educators*. Washington, DC: U. S. Department of Education.

Received January 2, 2014

Revision received October 6, 2014

Accepted October 22, 2014 ■