Journal of Educational Psychology
2008, Vol. 100, No. 4, 741–753

0022-0663/08/$12.00 DOI: 10.1037/a0013099

Classroom Age Composition and Developmental Change in 70 Urban Preschool Classrooms

Arlen C. Moller
Children's Institute, Inc. and Gettysburg College

Emma Forbes-Jones and A. Dirk Hightower
Children's Institute, Inc. and University of Rochester

A multilevel modeling approach was used to investigate the influence of age composition in 70 urban preschool classrooms. A series of hierarchical linear models demonstrated that greater variance in classroom age composition was negatively related to development on the Child Observation Record (COR) Cognitive, Motor, and Social subscales. This was true when controlling for class size, general classroom quality, and socioeconomic status at the classroom level and for age, gender, and baseline ability at the child level. Additionally, to address possible concerns related to nonrandom assignment to classrooms, a series of models were run including variance in developmental age (i.e., baseline ability) at the classroom level and at the child level. The results were consistent for chronological age composition and developmental age composition at the classroom level; greater variance in classroom developmental age composition was negatively related to Time 2 scores on the COR Cognitive, Motor, and Social subscales. Furthermore, a cross-level interaction indicated that negative influence of greater variance in classroom developmental age composition was stronger for children older in developmental age. Implications for early childhood education policy are discussed.

Keywords: preschool, mixed age, single age, age composition, Child Observation Record

There is a great deal of variability in the age composition of U.S. preschool classrooms. Some preschool programs tightly restrict the range of ages in their classrooms (e.g., 3-year-olds only). Other programs permit, and in some cases even promote, a wide range of ages in the same classroom (e.g., from 2.5 years to as old as 5 years). Such variability in practice follows from a lack of consensus in both the theoretical and the empirical literature on the effects of age composition of classrooms on individual learning.

Research on the effects of classroom age composition on individual growth is not new. The theoretical writings of two pioneers in the field of early childhood development, Vygotsky and Piaget, are at odds with respect to the predicted impact of exposure of individual children to peers of variant ages on cognitive and social development (Bailey, Burchinal, & McWilliams, 1993; Hartup, 1993; Piaget, 1932; Rubin, Bukowski, & Parker 1998; Vygotsky, 1930/1978). From one perspective, that of Piaget (1932) and others, interacting with peers who are close in age is preferable. Children close in age are also likely to be more closely matched in terms of knowledge, skill, and power. As a result, children in such a setting are able to engage in and resolve cognitive and social conflicts with each other and play a variety of roles in their resolution. By contrast, Vygotsky and others argued that interaction with older, more competent children provides an optimal context for cognitive and behavioral development (Vygotsky, 1930/1978).

Interestingly, in the years since, relatively little empirical consensus has emerged. Yet despite a lack of consensus, the National Association for the Education of Young Children has given mixed-age classrooms its endorsement[1] (Bredekamp, 1987), in line with several researchers in the field (e.g., Katz, Evangelou, & Hartman, 1990; Whaley & Kantor, 1992). We now explore the extant literature relating to preschool classroom age composition more closely.

Support for Mixed-Age Preschool Classrooms

Proponents of mixed-age preschool classrooms typically focus on the age-specific benefits most relevant to the younger children in such groupings. For example, as already noted, Vygotsky (1930/1978) spoke of the (perhaps) optimal challenges inherent in classroom contexts that include children who are older and thus more mature. Similarly, social learning theorists have maintained that

Arlen C. Moller, Department of Psychology, Gettysburg College, and Children's Institute, Inc., Rochester, New York; Emma Forbes-Jones, Department of Psychiatry, University of Rochester School of Medicine and Dentistry, and Children's Institute, Inc.; A. Dirk Hightower, Department of Clinical and Social Sciences in Psychology, University of Rochester, and Children's Institute, Inc.

We would like to thank those responsible for managing the Rochester Early Childhood Assessment Partnership Project, including Lauri Brugger, Susan Greenberg, Andrew MacGowan, Beverly Miller, and Kathy Slater. In particular, we thank Walter Gramiak for his help and expertise in managing and cleaning the dataset from which these analyses were derived. Finally, we thank Guillermo Montes, Aaron Metzger, and Janis Cameron for their insightful comments and feedback on earlier versions of this article.

Correspondence concerning this article should be addressed to Arlen C. Moller, Department of Psychology, Gettysburg College, Box 0407, 300 North Washington Street, Gettysburg, PA 17325. E-mail: amoller@gettysburg.edu

[1] The National Association for the Education of Young Children's (1996) most recent position statement does not take a position with regard to the age composition of preschool classrooms, and the recommendation for mixed-age classrooms extended by Bredekamp (1987) has not been retracted.

younger children benefit from being exposed to and modeling the behavior of older children (Bandura, 1986). In support, Rothstein-Fisch and Howes (1988) found that when toddlers interacted with slightly older peers, the younger children exhibited more complex language and engaged in more complex play in a family childcare setting. Howes and Farver (1987) similarly found that toddlers whose play partners were older peers engaged in more complex social play.

Other advocates for mixed-age classrooms have posited that older children, too, may benefit from exposure to younger children (Katz et al., 1990; Lloyd, 1999). For example, in such settings older children are in a position to more readily practice prosocial behavior (Derscheid, 1997; Urberg & Kaplan, 1986) and leadership skills (French, Waas, Stright, & Baker, 1986), given their elevated status in mixed-age group interactions. Chapman (1994) found that older children in a mixed-age elementary school setting served as mentors for younger children both spontaneously and when asked. Although the specific benefits that have been linked to mixed-age peer interaction have thus far been age specific (i.e., only relevant to younger or older children), advocates of mixed-age classrooms maintain that the age-specific advantages are presumably passed on as a child progresses through such a program (i.e., younger children will one day be older children in the same classroom), and thus an egalitarian ethic is preserved.

Further research has suggested additional benefits of mixed-age classrooms for all children, regardless of age. For example, research by Goldman (1981) has suggested that both 3- and 4-year-olds in mixed-age preschool classrooms engaged in significantly less parallel (as opposed to interactive) forms of play, relative to those in same-age classrooms. An investigation by Urberg and Kaplan (1986) found that both older and younger children in mixed-age preschool classrooms exhibited greater interactive functional play and less parallel functional play, relative to those in same-age classrooms. Rubin, Maioni, and Hornung (1976) have argued that interactive play represents a more mature level of play, relative to parallel play. Thus, these authors argued that mixed-age classrooms facilitate more mature play and therefore more optimal social–cognitive development.

Whaley and Kantor (1992) suggested that mixed-age classrooms may help to lessen the likelihood of parents and caregivers to make comparisons among children in terms of development. Instead, mixed-age rooms allow parents to focus on each child's development as an individual. As such, mixed-age classrooms may decrease competition between peers, focusing children instead on more mastery-oriented goals. Many studies conducted with children, adolescents, and college students confirm that focusing on individual standards (i.e., mastery goals) is optimal in terms of maintaining both intrinsic motivation and achievement in the classroom (see Elliot & Dweck, 2005, and Nicholls, 1989, for reviews). Blasco, Bailey, and Burchinal (1993) found that children with developmental disabilities, in particular, were more likely to engage in mastery play when they were enrolled in mixed-age settings.

Whaley and Kantor (1992) further speculated that benefits extend to both children and teachers as a result of the "family climate" and continuity of care fostered by mixed-age classrooms. That is, these authors suggested that the abrupt transitions in climate encountered when moving between restricted-age classrooms at larger childcare facilities could be stressful for both the teachers and the children. However, an empirical investigation into the perceived stress levels of children and teachers in such facilities has yet to be published.

Several authors have noted that gender segregation is less prevalent in mixed-age classrooms relative to restricted-age settings (Field, 1982; Lougee, Grueneich, & Hartup, 1977; Roopnarine et al., 1992; Roopnarine & Johnson, 1984). It seems that in mixed-age contexts, age plays a more powerful role than gender in peer playmate selection. Thus, at a society level, long-term promotion of gender equality might be better served by more preschools moving toward a greater diversity in terms of classroom age composition. Such an argument is highly speculative, as no empirical explorations into a relation between preschool classroom age composition and attitudes related to gender equality have been conducted.

Finally, some research has supported the notion that mixed-age peer interactions occur naturally, outside the classroom context (Ellis, Rogoff, & Cromer, 1981). An observational study by Ellis et al. (1981) of 436 children ages 1–2 found that target children were with same-age peers in only 6% of the observations recorded, compared with 55% of the observations involving mixed-age peer interactions (with children who differed in age by at least 1 year). These findings related to naturalistic social interactions, however, are inconclusive to the extent that this may be because of peer accessibility or children actively seeking out companionship from mixed-age peers. Goldman's (1981) research on mixed- versus same-age preschool classrooms demonstrated that—within mixed-age classrooms—age did not significantly influence choice of playmates. Thus, some authors have suggested that children are naturally better suited to developing in the context of mixed-age groups (Ellis et al., 1981; Whaley & Kantor, 1992).

Support for Restricted-Age Preschool Classrooms

The central argument in favor of restricting the range of age in preschool classrooms concerns a desire to provide each child with an optimal degree of challenge and the utility of specialization toward meeting this goal. As such, a criticism that has been levied against mixed-age classrooms in favor of more restricted-age preschool groupings concerns the development of age-appropriate curriculum and staff training. Specifically, when teacher training and curriculum are uniquely focused on what is developmentally appropriate for a more restricted age range of children, the quality of the classroom climate will be higher.

Although some authors have maintained that mixed-aged classrooms facilitate more mature forms of play (Goldman, 1981; Howes & Farver, 1987; Rothstein-Fisch & Howes, 1998; Urberg & Kaplan, 1986), other researchers have found evidence to the contrary or suggestive of a more complex effect than is often highlighted. For example, Roopnarine et al. (1992) observed more social and nonsocial dramatic play in preschoolers attending same-age relative to mixed-age classrooms, whereas more social–constructive and nonsocial manipulative play was seen in mixed-age classrooms. Roopnarine et al. maintained that dramatic play was the more complex form of play relative to manipulative play, suggesting that same-age classrooms facilitated more sophisticated play behaviors. Urberg and Kaplan (1986) reported finding a greater incidence of interactive, mature forms of play in mixed-age preschool classrooms relative to same-age classrooms; however, these authors also reported several observed costs associated with mixed-age classrooms. For instance, younger children in mixed-age classrooms tended to be involved in a greater number of negative interactions. This was attributed to the potential for frustration in older children resulting from interactions with less mature younger children. Furthermore, Urberg and Kaplan found that older children in mixed-age classrooms were more likely to be onlookers, detaching themselves from social interaction. Similarly, Goldman (1981) found more solitary play and less participation in teacher-directed activities among older children in mixed-age classrooms. Urberg and Kaplan speculated that teachers in the mixed-age classrooms might have had a more difficult time providing a sufficient variety of activities suitable for keeping children of varied age and ability interested and involved.

Proponents of mixed-age classrooms often focus on benefits to younger children, but the potential costs to the older children in such classrooms may sometimes be overlooked. A study of mixed-age elementary school classrooms (children ages 6 to 9 years) by Byrnes, Shuster, and Jones (1994) confirmed that both older students and their parents rated the quality of the academics in multiage classrooms less positively than that in single-age classrooms. Despite failing to report any benefits found for older, more able students in their study, the authors suggest that administrators ought to pay special attention to helping "more able students in seeing how they benefit from being in multiage classrooms" (p. 21). In another observational investigation of play behavior in different classrooms, Dunn, Kontos, and Potter (1996) monitored family childcare homes and found that frequency of interactions with older peers were associated with more complex cognitive play, whereas interactions with younger and same-age peers were associated with less complex social and cognitive play and lower receptive language scores. One lab experiment (Tudge, 1992) compared mixed- versus matched-competence peers (ages 5–9 years) working in dyads. The results of this experiment suggested that regression in thinking was as likely a consequence of mixed groupings as was improvement. These results suggest that at least with regard to some cognitive outcomes, older children in mixed-age groups may suffer. Furthermore, support for such a hypothesis was found by Winsler et al. (2002). Winsler et al. studied longitudinally one preschool's transition from two single-age classrooms (separate 3- and 4-year-olds) to two mixed-age classrooms (combined 3- and 4-year-olds). In this study, statistical interactions were found between child age and classroom age composition in predicting task focus in the classroom. Although 3-year-olds seemed to do better after the change to mixed-age classrooms, 4-year-olds were observed to be less focused and goal directed. Moreover, children of both ages were observed demonstrating significantly more overt positive affect (e.g., smiles and laughter) and less neutral affect in the single-age classrooms compared with the mixed-aged classrooms.

The 1995 meta-analysis conducted by Veenman concluded that there were no academic or social differences between elementary school children attending mixed-age and same-age classrooms. Similarly, Blasco et al. (1993) reported no differences in terms of complex mastery play in same- versus mixed-age classrooms when their observations focused on typically developing preschool-age children; the benefits attributed to mixed-age classrooms and reported above were (again) specific to children with developmental disabilities.

Ability-Based Groupings

As noted, many of the theoretical assertions for or against mixed- and restricted-age groupings are based on the assumption that age is closely related to ability. By this token, ability-based groupings may be expected to have an influence on developmental outcomes similar to that of age-based groupings. Mirroring the literature supporting the efficacy of mixed-age classrooms, many researchers have championed mixed-ability groupings in terms of benefits to be reaped by disadvantaged children who are lower in ability (Guralnick, 1981; Rogers & Ross, 1986; Tach & Fargas, 2006). For example, Tach and Fargas (2006) found that within kindergarten and first-grade classrooms in which ability-based groupings were used, placement into a higher ability group exerted a uniformly positive effect on student learning-related behavior and reading achievement. Typically, these studies have implicitly assumed that the costs for higher ability children in such mixed-ability groupings are negligible, although several exceptions (e.g., Guralnick, 1981) have reported no detrimental impact on higher ability children in mixed-ability groupings. Nevertheless, a theme that runs through these studies advocating for mixed-ability groupings is that an important and focal goal is ensuring that the needs of students who have experienced situations that place them at risk or are disadvantageous for their academic development are met. From an entirely different perspective, some researchers have advocated expressly for the interests of the most gifted and talented children. Researchers focused on those children higher in ability tend to advocate for more restricted-ability groupings, such that high-ability children have more opportunities to interact with peers of similar ability (e.g., Gallagher, 1986). To summarize, as is the case in the literature focused on age-related groupings, the literature on ability-related groupings is mixed; furthermore, the most controversial aspect of the debate concerns those children in mixed contexts who are older, higher in ability, or both.

Shortcomings of the Extant Literature Addressed by the Present Investigation

Many of the studies that have expressly focused on classroom age composition have targeted an elementary school population, which makes it difficult to extend the conclusions to preschool classrooms (e.g., Chapman, 1994; French et al., 1986; Veenman, 1995). Early childhood (ages 2 to 5) is a period of relatively rapid development; thus, the impact of age differences in classroom composition within this range is likely to be magnified, if not distinctly different, in preschool classrooms relative to that found in elementary school settings among older children.

Furthermore, as Winsler et al. (2002, p. 306) pointed out, "most of the (extant) research exploring the effects of age-composition on children's peer interactions has been conducted in laboratory or quasi-laboratory settings." In other words, with few exceptions, the extant literature lacks empirical research on preschool classroom age composition conducted in preschool classrooms. The exceptions (Bailey et al., 1993; Goldman, 1981; Winsler et al., 2002) have been limited to research conducted with a relatively small number of classrooms, typically restricted to a single organization. For example, Bailey et al.'s (1993) investigation tracked the developmental trajectories of just 59 children between the ages of 21 and 67 months. Goldman's (1981) sample consisted of just three classrooms, and Winsler et al.'s (2002) sample included only four. Furthermore, the samples included in the vast majority of studies reviewed above came from affluent or middle-class suburban communities. From the perspective of triage at the level of public policy, an argument could be made for focusing attention first on the neediest and most underserved populations. Yet, urban samples of lower socioeconomic standing have typically been ignored in this literature. An additional limitation of the extant literature is that researchers have typically focused on a limited range of outcomes, perhaps masking a richer, more ambiguous phenomenon.

We undertook an investigation into the question of preschool classroom age composition on a scale far larger than previously conducted (806 children in 70 classrooms), a sample that included every preschool classroom under the administrative umbrella of a public school district in a mid-sized city (approximate population = 212,000). We included as outcomes in our research well-validated measures of various aspects of development (cognitive, social, and motor) and controlled for baseline levels on entering the classroom. Furthermore, we assessed and controlled for general classroom quality. As a result, this research should offer fresh and rich insight into questions concerning the consequences of preschool classroom age composition that can be extended more easily to urban populations.

Method

Participants

Participants included 806 preschool children from 70 preschool classrooms. The classrooms in the sample all fell under the administrative umbrella of a public school district in a mid-sized city in the northeastern United States. The mean age of children in the sample at Time 1 was 4.15 years ($SD = 0.50$; 51% male). The ethnic composition included 57% African American, 17% White, 15% Hispanic, 2% Asian, and 9% children of unreported race or ethnicity. Overall, the students enrolled in the district were low in socioeconomic status (SES). Across the district, 86% of students came from homes below the poverty line, and 84% of the students qualified for free or reduced-price lunch. The mean number of students in a classroom was 16.34 (range = 8 to 21). The mean difference in age between the oldest and youngest child in a classroom was 1.58 years (range = 0.73 to 2.73). The mean standard deviation of children's age in classrooms was 0.44 (range = 0.10 to 0.70).

Measures

Child Observation Record (COR)

This instrument (High/Scope Educational Research Foundation, 1992) measures several relevant domains of preschool children's development. The COR includes 32 items, with response options ranging from *least developed* to *most developed* on a 5-point scale, to be filled out by teachers immersed in the classroom environment from observations of naturally occurring behaviors. Before teachers used the COR they received 3 hr of training familiarizing them with this instrument. A recent investigation, focused on validating the COR, found that a three-factor solution, which included Cognitive Skills, Social Engagement, and Coordinated Movement/Motor Skills, best accounted for the correlations among the 32 COR items (Fantuzzo, Hightower, Grim, & Montes, 2002). Thus, in the present investigation, we used this three-factor solution, as well as the total COR (a combined score across all three subscales). Convergent and divergent validity studies have found significant relations between COR dimensions and other classroom competency constructs, such as peer play interactions (Penn Interactive Peer Play Scale; Fantuzzo, Coolahan, Mendez, McDermott, & Sutton-Smith, 1998; Fantuzzo et al., 2002; McWayne, Fantuzzo, & McDermott, 2004; Sekino & Fantuzzo, 2005), psychological adjustment (Teacher-Child Rating Scales [Fantuzzo et al., 2002; Perkins & Hightower, 1999, 2000] and Adjustment Scales for Preschool Intervention [Lutz, 1999; Lutz, Fantuzzo, & McDermott, 2002; Sekino & Fantuzzo, 2005]), receptive vocabulary (Peabody Picture Vocabulary Test-III; Dunn & Dunn, 1997; Fantuzzo et al., 2002), general learning skills (Preschool Learning Behaviors Scale [McDermott, Green, Francis, & Stott, 1996; McWayne et al., 2004; Sekino & Fantuzzo, 2005] and McCarthy Scale of Children's Abilities [McCarthy, 1972; Schweinhart, McNair, Barnes, & Larner, 1993]), early reading ability (Test of Early Reading Ability, 3rd ed.; Reid, Hreski, & Hammill, 2001; Sekino & Fantuzzo, 2005), early math ability (Test of Early Mathematics Ability, 2nd ed.; Ginsburg & Broody, 1990; Sekino & Fantuzzo, 2005), kindergarten literacy (Dynamic Indicators of Basic Early Literacy Skills; Good & Kaminski, 2002; Sekino & Fantuzzo, 2005), and kindergarten academic success (Early Screening Inventory—Revised Kindergarten Version; Meisels, Marsden, Wiske, & Henderson, 1997; McWayne et al., 2004).

Early Childhood Environment Rating Scale—Revised (ECERS–R)

The ECERS–R is among the most widely used observational tools allowing for objective assessment of preschool classroom quality (Henry et al., 2004; Howes & Smith, 1995; Scarr, Eisenberg, & Deater-Deckard, 1994). The seven areas of classroom quality that the ECERS–R measures are space and furnishing, personal care routines, language and reasoning, activities, interaction, program structure, and parents and staff. Each area contains 5–10 items that represent various elements of that area. For the present investigation, we averaged scores across all seven areas to create an overall, composite measure of classroom quality.

Procedure

Seventy teachers completed the COR for each child in their classroom, once in the fall and again in the spring. Teachers participated as part of an ongoing community assessment partnership that was open to all providers of services to preschool children in the area. Teachers were trained on the COR according to High/Scope standards, including procedures for observation and note taking, before completion of COR protocols. Trainees matched observations of videotaped classrooms against a standard to establish reliability levels of at least 80%. Master teachers and consultants were available to any teacher who had any questions or issues after the initial training for the entire school year. After COR ratings were completed, the first round of data sheets were returned to us for processing in October; the second round of data sheets were returned to us for processing in late April. Thus, the typical length of time between COR assessments was 6–7 months. COR ratings were returned for 95% of the original sample.[2]

ECERS–R observations of classroom quality were conducted by 24 observers midway through the academic year (in the months of February, March, and April). For classroom observers in their 1st year of training, ECERS–R observers attended a 15-hr training program and reached an interrater reliability of 85% agreement with a master observer who was trained by the ECERS–R authors. For observers in their 2nd year, an additional 4–5 hr of training were required. All observers maintained an interrater reliability of 80% agreement, with 20% of their observations being checked.

Results

COR Scoring

We calculated COR scores for each of the three subscales—Cognitive Skills (Time 1 [T1], $\alpha = .92$; Time 2 [T2], $\alpha = .91$), Social Engagement (T1, $\alpha = .93$; T2, $\alpha = .93$), and Coordinated Movement/Motor (T1, $\alpha = .87$; T2, $\alpha = .90$)—and the Combined/Total COR (T1, $\alpha = .96$; T2, $\alpha = .96$). The range for each COR subscale was from 0 to 5. Mean COR scores and standard deviations at Time 1 were, for Cognitive Skills, 2.08 and 0.69, respectively; for Social Engagement, 2.63 and 0.75, respectively; for Coordinated Movement/Motor, 2.68 and 0.68, respectively; and for Total COR, 2.43 and 0.66, respectively. Mean COR scores and standard deviations at Time 2 were, for Cognitive Skills, 3.07 and 0.77, respectively; for Social Engagement, 3.72 and 0.80, respectively; for Coordinated Movement/Motor, 3.65 and 0.80, respectively; and for Total COR, 3.48 and 0.72, respectively. Intercorrelations among the three COR factors ranged from .78 to .80 at Time 1 and from .71 to .76 at Time 2 (see Table 1). The high degree of correlation among the three COR factors indicates that analyzing each factor and the total COR score is justified.

Primary Analysis

We used hierarchical linear modeling (HLM; Bryk & Raudenbush, 1992) for our primary tests. HLM appropriately addresses the hierarchically nested design of our data set, in which lower level units, children, were nested within a higher level unit, classroom. HLM treats classroom as a random rather than a fixed effect, thereby permitting generalizations of the findings to a wider population. Unlike regression models, this analysis allows for the possibility that the within-group (i.e., classroom) slopes may differ significantly from one classroom to another. Such an approach makes it possible to see whether the relationship between child's age and change in COR score differs from classroom to classroom. For example, a positive main effect for classroom age range would imply that classrooms that are more mixed in terms of age composition are related to greater developmental growth, as assessed by COR change.

We calculated intraclass correlation coefficients (ICCs) for each of the four difference outcomes variables from the unconditional models: total COR score, ICC = .25, $\chi^2(69, N = 70) = 329.30$, $p < .001$; COR Social, ICC = .25, $\chi^2(69, N = 70) = 330.56$, $p < .001$; COR Motor, ICC = .23, $\chi^2(69, N = 70) = 301.19$, $p < .001$; and COR Cognitive, ICC = .25, $\chi^2(69, N = 70) = 333.65$, $p < .001$. These findings confirmed that a HLM approach would provide substantial benefits over a standard fixed-effects model approach for the analysis of these data.

Models 1–8: Chronological Age Effects

In the first set of models run, the primary classroom-level (Level 2) predictor was classroom chronological age composition, which was operationalized in two ways: (a) the range in chronological age between the youngest and oldest child in a given classroom and (b) the standard deviation of chronological age within a given classroom. Covariate predictors entered at the classroom level included the number of students in the classroom ($M = 16.34$). At the person or child level (Level 1), child's chronological age at Time 1, gender, and COR assessment score at Time 1 were included as predictors. HLM estimates classroom-level and child-level effects simultaneously. Thus, classroom-level effects are statistically independent of child-level effects. All predictors (both classroom and child level) were centered on the sample (grand) means. The multilevel equations used for this basic model are described in Appendix A. This basic model was run independently using two different indicators of classroom chronological age composition (chronological age range in classroom and standard deviation of chronological ages in classroom) and four different outcomes (total COR score and each subscale from the COR), resulting in Models 1–8.

Chronological age and the composite total COR. The first version run of the basic model described above included chronological age range as the index of classroom chronological age composition and total COR score at Time 2 as the outcome measure (Model 1). The results of this hierarchical linear model are summarized in Table 2. Overall, there was a significantly negative main effect at the classroom level for classroom chronological age composition ($\gamma_{01} = -.19$, $p < .05$), and a nonsignificant effect for class size ($\gamma_{02} = .02$, $p = .37$). As expected, there were significant main effects observed at the child level for child's chronological age ($\gamma_{10} = .23$, $p < .001$) and Time 1 total COR score ($\gamma_{20} = .83$, $p < .001$). There was also a significant main effect at the child level for gender ($\gamma_{30} = .06$, $p < .05$), such that girls scored higher than boys on the COR at Time 2, controlling for

Table 1
Correlations Between the Three COR Factors at Time 1 and Time 2 (N = 806)

Factor	1	2	3	4	5	6
1. T1 Social	—					
2. T1 Cognitive	.78**	—				
3. T1 Motor	.79**	.80**	—			
4. T2 Social	.60**	.50**	.52**	—		
5. T2 Cognitive	.57**	.63**	.55**	.71**	—	
6. T2 Motor	.52**	.52**	.58**	.75**	.76**	—

** $p < .01$.

[2] Forty-two children were dropped from the study (roughly 5%) because Time 2 COR scores were not available. These 42 children did not differ significantly from the retained sample ($n = 806$) in terms of mean age, racial, or gender distribution.

Table 2
Fixed Effects for Models 1–4 (Classroom Age Range as Focal Level 2 Predictor)

Parameter	Model 1: Total COR Coeff	Model 1: Total COR *SE*	Model 2: COR social Coeff	Model 2: COR social *SE*	Model 3: COR motor Coeff	Model 3: COR motor *SE*	Model 4: COR cognitive Coeff	Model 4: COR cognitive *SE*
Average classroom intercept (γ_{00})	3.53***	0.04	3.75***	0.05	3.67***	0.04	3.12***	0.04
Age range (γ_{01})	−0.19*	0.10	−0.22†	0.11	−0.16†	0.09	−0.19*	0.09
Class size (γ_{02})	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02
Intercept for child's age (γ_{10})	0.23***	0.05	0.20***	0.05	0.31***	0.05	0.33***	0.06
Age range × Child's age (γ_{11})	−0.10	0.08	−0.04	0.09	−0.08	0.10	−0.12	0.11
Intercept for T1 total COR (γ_{20})	0.83***	0.04	0.72***	0.04	0.73***	0.05	0.76***	0.05
Intercept for child's gender (γ_{30})	0.06*	0.03	0.09**	0.03	0.09*	0.04*	0.03	0.03

Note. Model 1 = Total COR as dependent outcome; Model 2 = COR Social as dependent outcome; Model 3 = COR Motor as dependent outcome; Model 4 = COR Cognitive as dependent outcome; COR = Child Observation Record; Coeff = coefficient; T1 = Time 1.
* $p < .05$. ** $p < .01$. *** $p < .001$. † $p < .10$.

Time 1 score. Finally, the cross-level interaction between classroom chronological age composition and child's chronological age was nonsignificant ($\gamma_{11} = -.10$, $p = .23$). This nonsignificant cross-level interaction suggested that the negative classroom-level main effect for chronological age composition was not moderated by the child-level chronological age of children in these classrooms.

The total variance at Level 1 (total COR at Time 2) was 1,648.63. This was estimated by specifying an unconditional model that included only a random intercept for each classroom. The error variance unaccounted for by the conditional Model 1 was reduced to 984.81. Comparison of these results indicates that 40.26% of the variance in total COR at Time 2 was accounted for by Model 1.

Chronological age and the three COR subscales: Social, Motor, and Cognitive. Next, we ran three hierarchical linear models using age range as the index of classroom age composition, independently predicting the three subscales from the COR as outcomes: Social, Motor, and Cognitive (Models 2, 3, and 4, respectively). The estimates from each model are summarized in Table 2, and the observed pattern of results was remarkably consistent across all three COR subscales. The negative main effect for classroom chronological age range was marginally significant when predicting COR Social and COR Motor and significant when predicting COR Cognitive. Child's chronological age was a significant predictor of all three COR subscales, as were Time 1 scores (for each subscale, respectively). Child's gender was a significant predictor of COR Social and COR Motor (in each case, girls scored higher), but did not significantly predict COR Cognitive. The cross-level interaction between child's chronological age and classroom chronological age range was nonsignificant for all three COR subscales.

The total variance at Level 1 was estimated by specifying an unconditional model that included only a random intercept for each classroom for each of the three COR subscales: Social, 1,813.15; Motor, 1,825.91; and Cognitive, 1,755.37. The error variance unaccounted for by Model 2 was reduced to 1,291.08; thus, the model accounted for 28.79% of the variance. The error variance unaccounted for by Model 3 was reduced to 1,367.06; thus, the model accounted for 25.13% of the variance. The error variance unaccounted for by Model 4 was reduced to 1,179.26; thus, the model accounted for 32.82% of the variance.

Next, we reran the four models described above, substituting the standard deviation of chronological ages for chronological age range at the classroom level (Models 5–8). Standard deviation of chronological ages in classrooms represents an alternative index of classroom chronological age composition that is more resistant to extreme values (i.e., if a single child were much younger or older than all the other children in a classroom). The pattern of results replicated those found using chronological age range as the index of classroom chronological age composition. For a summary of these results, see Table 3.

The error variance unaccounted for by Model 5 was reduced from 1,648.63 to 978.51; thus, the model accounted for 40.65% of the variance. The error variance unaccounted for by Model 6 was reduced from 1,813.15 to 1,284.49; thus, the model accounted for 29.16% of the variance. The error variance unaccounted for by Model 7 was reduced from 1,825.91 to 1,361.85; thus, the model accounted for 25.42% of the variance. The error variance unaccounted for by Model 8 was reduced from 1,755.37 to 1,172.09; thus, the model accounted for 33.23% of the variance.

Models 9–16: Developmental Age Effects

In the second set of models run, the primary classroom-level (Level 2) predictor was classroom developmental age composition. Developmental age was operationalized as COR score at Time 1. As such, in the interest of conceptual clarity, in the following models Time 1 COR score is referred to as *developmental age.* Developmental age range was operationalized in two ways: the range in score between the lowest and highest Time 1 COR score in a given classroom and the standard deviation of COR scores at Time 1 within a given classroom. Covariate predictors entered at the classroom level included the number of students in the classroom ($M = 16.34$).

At the person or child level (Level 1), child's developmental age (i.e., Time 1 COR score), gender, and age at Time 1 were included as predictors. Finally, we included a term representing the cross-level interaction between child's developmental age (Time 1 COR) and classroom developmental age composition (range in COR scores or standard deviation of COR scores). HLM estimates classroom-level and child-level effects simultaneously. Thus, classroom-level effects are statistically independent of child-level

Table 3
Fixed Effects for Models 5–8 (Classroom Age Standard Deviation as Focal Level 2 Predictor)

Parameter	Model 5: Total COR		Model 6: COR social		Model 7: COR motor		Model 8: COR cognitive	
	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*
Average classroom intercept (γ_{00})	3.53***	0.04	3.75***	0.05	3.67***	0.04	3.12***	0.04
Age *SD* (γ_{01})	−0.74*	0.31	−0.79*	0.34	−0.60*	0.30	−0.71*	0.30
Class size (γ_{02})	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02
Intercept for child's age (γ_{10})	0.20***	0.05	0.17***	0.05	0.20***	0.05	0.30***	0.06
Age *SD* × Child's age (γ_{11})	0.17	0.32	0.44	0.34	−0.23	0.45	0.18	0.43
Intercept for T1 Total COR (γ_{20})	0.82***	0.04	0.72***	0.04	0.73***	0.05	0.76***	0.04
Intercept for child's gender (γ_{30})	0.06*	0.03	0.09**	0.03	0.09*	0.04	0.03	0.03

Note. Model 1 = Total COR as dependent outcome; Model 2 = COR Social as dependent outcome; Model 3 = COR Motor as dependent outcome; Model 4 = COR Cognitive as dependent outcome; COR = Child Observation Record; Coeff = coefficient; T1 = Time 1.
* $p < .05$. ** $p < .01$. *** $p < .001$.

effects. All predictors (both classroom and child level) were centered on the sample (grand) means. The multilevel equations used for the second basic model are described in Appendix B. (As noted earlier, the multilevel equations used for the first basic model are described in Appendix A.) Paralleling Models 1–8, the second basic model was run independently, using two different indicators of classroom developmental age composition (developmental age range in classroom and standard deviation of developmental ages in classroom) and four different outcomes (total COR score and each subscale from the COR), resulting in Models 9–16.

Developmental age and the composite COR total. The basic model described above was first run using developmental age range as the index of classroom developmental age composition and total COR score at Time 2 as the outcome measure (Model 9). The results of this hierarchical linear model are summarized in Table 4. Overall, there was a significantly negative main effect at the classroom level for classroom developmental age range ($\gamma_{01} = -.23$, $p < .01$). The classroom-level effect for class size was nonsignificant ($\gamma_{02} = .02$, $p = .31$). As expected, there were significant main effects observed at the child level for child's developmental age ($\gamma_{10} = .86$, $p < .001$) and child's chronological age ($\gamma_{20} = .22$, $p < .001$). There was also a significant child-level main effect for gender ($\gamma_{30} = .06$, $p < .05$). Finally, a significant cross-level interaction between classroom developmental age range and child's developmental age was observed ($\gamma_{11} = -.17$, $p < .01$). This cross-level interaction suggested that there was little difference in terms of Time 2 COR for children lower in developmental age (i.e., 1 standard deviation below the grand mean Time 1 Total COR score) in classrooms with mixed versus restricted developmental age ranges. However, for children higher in developmental age (i.e., 1 standard deviation above the grand mean Time 1 Total COR score), there was a significant advantage to being in a classroom with a restricted developmental age range versus a classroom with a wide range in developmental age. This interaction is illustrated in Figure 1.

The total variance at Level 1 (total COR at Time 2) was 1,648.63. This was estimated by specifying an unconditional model that included only a random intercept for each classroom. The error variance unaccounted for by the conditional Model 1 was reduced from 1,648.63 to 989.45; thus, the model accounted for 40.00% of the variance.

Developmental age and the three COR subscales: Social, Motor, and Cognitive. Next, three hierarchical linear models using developmental age range as the index of classroom developmental

Table 4
Fixed Effects for Models 9–12 (Classroom Developmental Age Range as Focal Level 2 Predictor)

Parameter	Model 9: Total COR		Model 10: COR social		Model 11: COR motor		Model 12: COR cognitive	
	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*
Average classroom intercept (γ_{00})	3.54***	0.04	3.75***	0.05	3.68***	0.04	3.13***	0.04
Developmental age range (γ_{01})	−0.23**	0.07	−0.21*	0.09	−0.15*	0.07	−0.23**	0.07
Class size (γ_{02})	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Intercept for child's age (γ_{10})	0.86***	0.05	0.75***	0.04	0.74***	0.05	0.78***	0.05
Developmental age range × Child's developmental age (γ_{11})	−0.17**	0.05	−0.10*	0.05	−0.08	0.07	−0.08	0.06
Intercept for child's age (γ_{20})	0.22***	0.04	0.20***	0.05	0.30***	0.05	0.32***	0.06
Intercept for child's gender (γ_{30})	0.06*	0.03	0.09**	0.03	0.09*	0.04	0.03	0.03

Note. Model 1 = Total COR as dependent outcome; Model 2 = COR Social as dependent outcome; Model 3 = COR Motor as dependent outcome; Model 4 = COR Cognitive as dependent outcome; COR = Child Observation Record; Coeff = coefficient.
* $p < .05$. ** $p < .01$. *** $p < .001$.

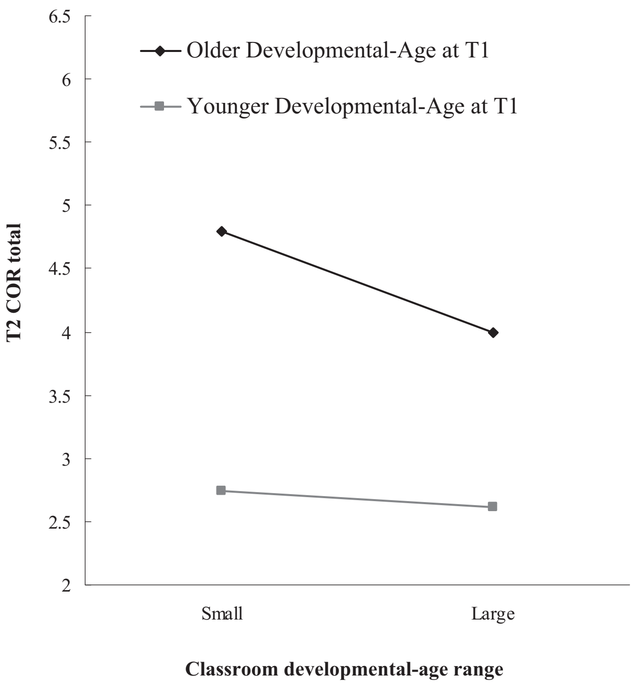

Figure 1. Classroom developmental age range by child's developmental age predicting total Child Observation Record (COR) score at Time 2, controlling for total COR score at Time 1.

age composition were run independently, predicting the three COR subscales as outcomes: Social, Motor, and Cognitive (Models 10, 11, and 12, respectively). The estimates from each model are summarized in Table 4, and the observed pattern of results was remarkably consistent across all three COR subscales. Classroom-level developmental age range was a significantly negative predictor of all three subscales from the COR. Class size was not a significant predictor of any of the three COR subscales. Child's chronological age and developmental age significantly predicted all three COR subscales. Child's gender significantly predicted COR Social and COR Motor (girls scored higher), but not COR Cognitive. Finally, the cross-level interaction between classroom-level developmental age range and child's developmental age was significant only with regard to COR Social, but did not significantly predict COR Motor or COR Cognitive. These results are summarized in Table 4.

Next, we reran the four models described above, substituting the standard deviation of developmental ages as the indicator of developmental age composition at the classroom level (Models 13–16). Standard deviation of developmental ages in classrooms represents an alternative index of classroom developmental age composition that is more resistant to extreme values (i.e., if a single child scored much higher or lower on the Time 1 COR relative to all the other children in a given classroom). The pattern of results replicated those found using developmental age range as the index of classroom developmental age composition. These results are summarized in Table 5.

Classroom quality. We next included classroom quality (ECERS–R) in the models described above exploring the influence of both chronological (Models 1–8) and developmental (Models 9–16) age composition on residual COR change. The significant effects reported for classroom age composition all remained significant when controlling for classroom quality. That is, classroom-level variance in age, chronological or developmental, was negatively related to residual COR change. Furthermore, there were no significant main effects found for classroom quality and no significant classroom-level interactions between classroom quality and classroom age composition.

It is, however, worth noting that the ECERS–R scores for the classrooms in this sample were uniformly approaching ceiling levels for excellence. The maximum possible score for overall classroom quality on this scale is 7.0. In the school district sampled, the mean score for overall classroom quality was 6.42 (Mdn = 6.75), with a standard deviation of only 0.77. Thus, the level of ECERS–R-rated classroom quality in this sample was greater than 2 standard deviations above the national average (approximately 5.0), as previously reported by Montes, Hightower, Brugger, and Moustafa (2005).

SES. Next, an indicator of classroom-level SES was included in the models (1–16) described above. In the absence of direct indicators of family income, we used mothers' level of education as an indirect indicator of child-level SES. Studies that have explored the relation between level of education and income typically report strong correlations (e.g., r = .51; Robert, 1998); researchers often use these variables as interchangeable indicators of SES in multivariate models owing to the high degree of shared variance (e.g., Winkleby, Kraemer, Ahn, & Varady, 1998). Thus, mean and median mothers' highest completed level of education were used as indicators of SES at the classroom level.[3] Across all 70 classrooms in the sample, the mean level of mothers' education was 3.25 (SD = 1.09), corresponding to a level of education between high school diploma or GED and some college. The significant effects reported for classroom age composition all remained significant when controlling for classroom-level SES. That is, greater variance in age, chronological or developmental, at the classroom level was negatively related to residual COR change. Furthermore, no significant main effects were found for classroom SES, and there were no significant classroom-level interactions between classroom SES and classroom age composition.

General Discussion

This investigation represents a unique and important contribution to the literature on preschool classroom age composition in a number of respects. First, the study included a sample far larger than that in any previously conducted research, 806 children from 70 different preschool classrooms. Second, this research is among the first to use a well-validated assessment of early childhood development (i.e., the COR) in a variety of domains (social, motor, and cognitive) and to include assessments at two time points (spaced approximately 6 months apart). Thus, we were able to explore the influence of classroom age composition on residualized change in various aspects of development. Additionally, important covariates, such as general classroom quality (i.e., the

[3] Mother's education was available for 88% of the children in the sample. To calculate mean and median mothers' education at the classroom level, missing values were replaced with the classroom (or group) mean.

Table 5
Fixed Effects for Models 13–16 (Classroom Developmental Age Standard Deviation as Focal Level 2 Predictor)

Parameter	Model 13: Total COR		Model 14: COR social		Model 15: COR motor		Model 16: COR cognitive	
	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*	Coeff	*SE*
Average classroom intercept (γ_{00})	3.53***	0.04	3.75***	0.05	3.67***	0.04	3.12***	0.04
Developmental age *SD* (γ_{01})	−0.71**	0.23	−0.59*	0.28	−0.46*	0.21	−0.72**	0.21
Class size (γ_{02})	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.02
Intercept for child's age (γ_{10})	0.86***	0.05	0.75***	0.04	0.73***	0.05	0.78***	0.05
Developmental age *SD* × Child's developmental age (γ_{11})	−0.48*	0.19	−.29†	0.15	−0.07	0.19	−0.22	0.22
Intercept for child's age (γ_{20})	0.22***	0.04	0.20***	0.05	0.30***	0.05	0.32***	0.06
Intercept for child's gender (γ_{30})	0.06	0.03	0.09**	0.03	0.09*	0.04	0.03	0.03

Note. Model 1 = Total COR as dependent outcome; Model 2 = COR Social as dependent outcome; Model 3 = COR Motor as dependent outcome; Model 4 = COR Cognitive as dependent outcome; COR = Child Observation Record; Coeff = coefficient.
* $p < .05$. ** $p < .01$. *** $p < .001$. † $p < .10$.

ECERS–R) and SES, were included. Finally, this research focuses on a population (i.e., urban preschool children) that has been understudied, yet is extremely important from the perspective of policymakers and others interested in social welfare.

The results were surprising, insofar as they contradicted assertions made by advocates of mixed-age classrooms (Bredekamp, 1987; Katz et al., 1990), yet remarkably consistent across indicators of classroom age composition and domains of development. We consistently observed a significant main effect at the classroom level for classroom age composition, which suggested that a wide range in children's ages within a classroom (and high standard deviations in terms of age) was negatively related to development. This was true in terms of overall development (i.e., total COR score) and, independently, in terms of social, motor, and cognitive domains of development (COR subscales).

Furthermore, the negative relation observed between classroom (chronological) age composition and developmental outcomes was replicated using an indicator of classroom developmental age composition (range and standard deviation of Time 1 COR scores). This alternative approach to exploring the impact of classroom age composition was important for a number of reasons.

First, the parallel set of findings observed for classroom developmental age range helps to mitigate potential confounds that might have resulted from nonrandom assignment to classrooms. For example, one might imagine that the oldest children in mixed-age classrooms may be more likely to have been held back and that the brightest children are culled from these mixed-age classrooms when they become older. Those classrooms with wider distributions in age might also have a higher frequency of older children who are slow learners or are developmentally delayed. However, the models run using Time 1 COR scores as an indicator of developmental age help to address and dismiss the notion that older children in mixed-age classrooms in our sample were developmentally delayed or otherwise noncomparable to older children in single-age classrooms, as chronological age was separated from developmental age, and the pattern of results did not change.

Second, the fact that a wider distribution of both chronological and developmental age in preschool classrooms was negatively related to development, coupled with a large and diverse sample of classrooms, makes a strong case for generalizing this effect to other urban populations.

Third, the use of Time 1 COR scores as an indicator of developmental age allowed us to identify a cross-level interaction between classroom developmental age composition and child's developmental age, which further specifies the nature of the significant main effect found for classroom developmental age composition. This cross-level interaction suggested that children high in developmental age were most negatively influenced by assignment to classrooms with wider distributions of developmental age, whereas younger children did not seem to be significantly affected by classroom assignment.

As noted, these findings stand in contrast to enthusiasm that has been expressed for movement toward more mixed-age classrooms in preschool education (e.g., Bredekamp, 1987; Katz et al., 1990). As we have argued, the empirical evidence for this enthusiasm has been decidedly mixed and inconclusive to this point and, more important, has relied on relatively small samples, typically in suburban settings. In this context, the present research strongly suggests that reconsideration of the issue of classroom age composition in early childhood education is warranted.

Most relevant, this research speaks to the consequences of variance in classroom age composition as it naturalistically occurs in urban preschool classrooms today. Although it could well be that mixed-age contexts have the potential to benefit children, in this large and diverse urban sample this was not the case. The opposite proved true. For reasons yet to be fully elaborated, and which merit further research, older children in mixed-age classrooms developed at a slower rate than older children in single-age classrooms.

Theoretical Implications

The findings from the present investigation strongly support the theory-based predictions offered by Piaget (1932) and others, who argued that interacting with peers who are close in age and ability will result in optimal learning. At the same time, these findings are not entirely inconsistent with predictions offered by Vygotsky (1930/1978) and others, who argued for mixed-age interaction principally on the basis of the merits implicit for younger children in these contexts. These authors posited that younger children could benefit from interacting with older, more competent children. In this regard, their prediction was unsupported. However, implicit in the Vygotskian model are the potential negative consequences for those who are older and more competent in mixed-age interaction. Thus, in terms of the relative advantage or disadvantage for different children in mixed-age interaction, the data from this study provide some support for the assertions made by Vygotsky and his colleagues. That is, one interpretation of the Vygotskian model is that variance in age composition or ability among interaction partners should be relatively less beneficial for those older children who are higher in ability. Our investigation supported this relativistic aspect of the Vygotskian model, such that although the influence of developmental age variance was negligible for developmentally younger children, there were significant negative consequences for developmentally older children.

The theoretical accounts for the influence of mixed- versus restricted-age groupings offered by Piaget (1932) and Vygotsky (1930/1978) focused primarily on the influence of peer interaction. However, it is likely that those supervising children (teachers, in the present investigation) may behave differently in mixed- versus restricted-age groupings as well. That is, as a result of being placed in mixed-age versus restricted-age classrooms, teachers may change their modes of instruction and interaction. Although exploring the dynamics through which teachers may be influenced and, in turn, exert influence on the children in mixed- versus restricted-age classrooms was beyond the scope of this study, it remains an important issue to consider.

Limitations and Conclusions

Differences in these findings relative to other published studies may, in part, be attributed to the time horizon used in assessing the influence of classroom age composition. The vast majority of studies on classroom age composition have been cross-sectional (i.e., single time point as opposed to change) in design (e.g., Dunn et al., 1996; Field, 1982; Goldman, 1981; Urberg & Kaplan, 1986; for exceptions, see Bailey et al., 1993, and Winsler et al., 2002). A strength of the present investigation was that development was assessed longitudinally. However, the full influence of mixed-age classrooms may take several years to manifest. One future direction for this line of research would be to explore the influence of classroom age composition on developmental change across multiple time points, extending from 6 months (as in the present research) to several years after initial exposure to these classrooms.

A strength of this research was the use of assessments of constructs at multiple time points; however, data were only collected at two time points. A logical extension of the basic approach we have outlined in this research would involve assessments at three or more time points. This would permit exploration into the possibility of both linear and nonlinear trajectories.

Another limitation of this research involves the correlational nature of these data. Empirical investigations that manipulate the age composition of preschool classrooms, with random assignment to condition, are warranted. As an alternative, when large-scale institutional policy changes are made affecting the age composition of preschool classrooms, such cases represent pseudoexperiments that may be used to more strongly establish a causal relation between classroom age composition and developmental outcomes (e.g., Winsler et al., 2002).

Finally, the uniformly high ratings of classroom quality (ECERS–R) in this sample were not conducive to appropriately testing the potential for classroom quality to moderate the negative effect of age variance observed in mixed-age classrooms. Further investigations are called for, targeting samples that include classrooms with greater variance in quality and the inclusion of additional indicators of classroom quality.

Despite these limitations, this investigation should be considered a significant contribution, as it represents an empirical investigation of preschool classroom age composition on a far larger scale than previously completed. It introduces a novel, and potentially generative, methodological approach for investigating the influence of classroom age composition (i.e., treating age composition as a continuous variable). Furthermore, this question was explored in terms of both chronological and developmental age composition in these classrooms, and the results were remarkably consistent across indicators of classroom composition and domains of development. The results support initiatives to maintain and facilitate movement toward preschool classrooms with more restricted age composition, at least insofar as age composition is a factor considered independently from curriculum considerations. As noted, the influence of age composition may be dynamically related to teachers' modes of instruction, to the extent that teachers are influenced by classroom age composition. The present investigation leaves open the possibility that mixed-age preschool classrooms coupled with curricula tailored to such climates could indeed be as good as or even superior to the typical restricted-age classroom. Specifically, our data suggest that such mixed-age-classroom–tailored curricula should address the needs of more mature, developmentally advanced children, who seem to have fared the worst under these conditions in our sample.

Ultimately, a thorough analysis of this issue must also take into consideration the economics of mixed- versus restricted-age classrooms. In areas of the country in which preschool classrooms are easily filled to capacity, it seems most prudent at this juncture to advocate for more restricted age composition. However, in parts of the country in which preschool classrooms are less easily filled, it may become significantly more cost effective to permit greater diversity in classroom age composition. As such, we cannot unilaterally condemn mixed-age classrooms but would instead implore educators compelled to work in these settings to make special efforts to provide individualized, optimal challenges for children who may otherwise be disadvantaged.

References

Bailey, D. B., Burchinal, M. R., & McWilliams, R. A. (1993). Age of peers and early childhood development. *Child Development, 64,* 848–862.

Bandura, A. (1986). *Social foundations of thought and action.* Englewood Cliffs, NJ: Prentice Hall.

Blasco, P. M., Bailey, D. B., & Burchinal, M. A. (1993). Dimensions of mastery in same-age and mixed-age integrated classrooms. *Early Childhood Research Quarterly, 8,* 193–206.

Bredekamp, S. (Ed.). (1987). *Developmentally appropriate practice in early childhood programs serving children from birth through age 8: Expanded edition.* Washington, DC: National Association for the Education of Young Children.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Byrnes, D. A., Shuster, T., & Jones, M. (1994). Parent and student views on multiage classrooms. *Journal of Research in Childhood Education, 9,* 15–23.

Chapman, M. L. (1994). Literacy learning in a primary multiage classroom: Some findings from an investigation of peer writing events. In C. K. Kinzer & D. J. Leu (Eds.), Multidimensional aspects of literacy research, theory, and practice. *Forty-third Yearbook of the National Reading Conference* (pp. 550–559). Chicago: The National Reading Conference.

Derscheid, L. E. (1997). Mixed-age grouped preschoolers' moral behavior and understanding. *Journal of Research in Childhood Education, 11,* 147–151.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Services.

Dunn, L., Kontos, S., & Potter, L. (1996). Mixed-age interactions in family child care. *Early Education and Development, 7,* 349–366.

Elliot, A. E., & Dweck, C. (2005). *Handbook of competence motivation.* New York: Plenum Press.

Ellis, S., Rogoff, B., & Cromer, C. C. (1981). Age segregation in children's social interactions. *Developmental Psychology, 17,* 399–407.

Fantuzzo, J., Coolahan, K. C., Mendez, J., McDermott, P. A., & Sutton-Smith, B. (1998). Contextually-relevant validation of constructs of peer play with African American Head Start children: Penn Interactive Peer Play Scale. *Early Childhood Research Quarterly, 13,* 411–431.

Fantuzzo, J., Hightower, D., Grim, S., & Montes, G. (2002). Generalization of the Child Observation Record: A validation study for diverse samples of urban, low-income preschool children. *Early Childhood Research Quarterly, 17,* 106–125.

Field, T. (1982). Same-sex preferences of preschool children: An artifact of same-age grouping? *Child Study Journal, 12,* 151–159.

French, D. C., Waas, G. A., Stright, A. L., & Baker, J. A. (1986). Leadership asymmetries in mixed-age children's groups. *Child Development, 57,* 1277–1283.

Gallagher, J. J. (1986). The need for programs for young gifted children. *Topics in Early Childhood Special Education, 6,* 1–8.

Ginsburg, H. P., & Broody, A. J. (1990). *Test of Early Mathematics Ability* (2nd ed.). Austin, TX: PRO-ED.

Goldman, J. A. (1981). Social participation of preschool children in same- versus mixed-age groups. *Child Development, 52,* 644–650.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic indicators of basic early literacy skills* (6th ed.) Retrieved from University of Oregon, Institute for the Development of Education Achievement: http://dibels.uoregon.edu

Guralnick, M. J. (1981). The social behavior of preschool children at different developmental levels: Effects of group composition. *Journal of Experimental Child Psychology, 31,* 115–130.

Hartup, W. W. (1993). Peer relationships. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (4th ed., pp. 103–196). New York: Wiley.

Henry, G. T., Ponder, B. D., Rickman, D. K., Mashburn, A. J., Henderson, L. W., & Gordon, C. S. (2004). *An evaluation of the implementation of Georgia's pre-k program: Report of the findings from the Georgia early childhood study (2002–2003).* Retrieved November 24, 2006 from http://aysps.gsu.edu/publications/2005/GAPrek2004.pdf

High/Scope Educational Research Foundation. (1992). *Manual: High/Scope Child Observation Record for ages 2.5–6.* Ypsilanti, MI: High/Scope Press.

Howes, C., & Farver, J. (1987). Social pretend play in 2-year olds: Effects of age of partner. *Early Childhood Research Quarterly, 2,* 305–314.

Howes, C., & Smith, E. W. (1995). Relations among child care quality, teacher behavior, children's play activities, emotional security, and cognitive activity in child care. *Early Childhood Research Quarterly, 10,* 381–404.

Katz, L. G., Evangelou, D., & Hartman, J. A. (1990). *The case for mixed-age groupings in early childhood education.* Washington, DC: National Association for the Education of Young Children.

Lloyd, L. (1999). Multi-age classes and high ability students. *Review of Education Research, 69,* 187–212.

Lougee, M., Grueneich, R., & Hartup, W. (1977). Social interaction in same- and mixed-age dyads of preschool children. *Child Development, 48,* 1353–1361.

Lutz, M. N. (1999). *Contextually relevant assessment of the emotional and behavioral adjustment of Head Start children.* Unpublished doctoral dissertation, University of Pennsylvania.

Lutz, M. N., Fantuzzo, J., & McDermott, P. (2002). Contextually relevant assessment of the emotional and behavioral adjustment of low-income preschool children. *Early Childhood Research Quarterly, 17,* 338–355.

McCarthy, D. A. (1972). *Manual for the McCarthy Scales of Children's Abilities.* San Antonio, TX: Psychological Corporation.

McDermott, P. A., Green, L. F., Francis, J. M., & Stott, D. H. (1996). *Preschool Learning Behaviors Scales.* Philadelphia: Edumetric & Clinical Science.

McWayne, C., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool competency in context: An investigation of the unique contribution of child competencies to early academic success. *Developmental Psychology, 40,* 633–645.

Meisels, S. J., Marsden, D. B., Wiske, M. S., & Henderson, L. W. (1997). *The Early Screening Inventory* (Rev. ed.). Ann Arbor, MI: Rebus.

Montes, G., Hightower, A. D., Brugger, L., & Moustafa, E. (2005). Quality child care and socio-emotional risk factors: No evidence of diminishing returns. *Early Childhood Research Quarterly, 20,* 361–372.

National Association for the Education of Young Children. (1996). *Developmentally appropriate practice in early childhood programs serving children from birth through age* 8 (NAEYC position statement) [Online]. Retrieved September 21, 2006 from http://www.naeyc.org/about/positions/daptoc.asp

Nicholls, J. G. (1989). *The competitive ethos and democratic education.* Cambridge, MA: Harvard University Press.

Perkins, P. E., & Hightower, A. D. (1999, April). *Construct validity of the Teacher-Child Rating Scale—Revised.* Poster presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Perkins, P. E., & Hightower, A. D. (2000, August). *Improving the assessment of children's problem behaviors and competencies: Refinement of the Teacher-Child Rating Scale.* Presented at the 108th Annual Convention of the American Psychological Association, Anaheim, California.

Piaget, J. (1932). *The moral judgment of the child.* London: Kegan Paul.

Reid, D. K., Hreski, W. P., & Hammill, D. D. (2001). *Test of Early Reading Ability* (3rd ed.). Austin, TX: PRO-ED.

Robert, S. A. (1998). Community-level socioeconomic status effects on adult health. *Journal of Health and Social Behavior, 39,* 18–37.

Rogers, D. L., & Ross, D. D. (1986). Encouraging positive social interaction among young children. *Young Children, 41,* 12–17.

Roopnarine, J. L., Ahmeduzzaman, M., Donnely, S., Gill, P., Mennis, A., Arky, L., et al. (1992). Social-cognitive play behaviors and playmate preferences in same-age and mixed-age classrooms over a 6-month period. *American Educational Research Journal, 29,* 757–776.

Roopnarine, J. L., & Johnson, J. E. (1984). Socialization in a mixed-age experimental program. *Developmental Psychology, 20,* 828–832.

Rothstein-Fisch, C., & Howes, C. (1988). Toddler peer interaction in mixed-age groups. *Journal of Applied Developmental Psychology, 9,* 211–218.

Rubin, K. H., Bukowski, W., & Parker, J. G. (1998). Peer interactions, relationships, and groups. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 619–700). New York: Wiley.

Rubin, K. H., Maioni, T. L., & Hornung, M. (1976). Free play behaviors in middle- and lower-class preschoolers: Parten and Piaget revisted. *Child Development, 47,* 414–419.

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

Schweinhart, L. J., McNair, S., Barnes, H., & Larner, M. (1993). Observing young children in action to assess their development: The High/Scope Child Observation Record study. *Educational and Psychological Measurement, 53,* 445–455.

Sekino, Y., & Fantuzzo, J. (2005). Validity of the Child Observation Record: An investigation of the relationship between COR dimensions and social-emotional and cognitive outcomes for head start children. *Journal of Psychoeducational Assessment, 23,* 242–261.

Tach, L. M., & Fargas, G. (2006). Learning-related behaviors, cognitive skills, and ability grouping when school begins. *Social Science Research, 35,* 1048–1079.

Tudge, J. R. H. (1992). Processes and consequences of peer collaboration: A Vygotskian analysis. *Child Development, 63,* 1364–1379.

Urberg, K. A., & Kaplan, M. G. (1986). Effects of classroom age-composition on the play and social behaviors of preschool children. *Journal of Applied Developmental Psychology, 7,* 403–415.

Veenman, S. (1995). Cognitive and noncognitive effects of multigrade and multi-age classes: A best-evidence synthesis. *Review of Educational Research, 65,* 319–381.

Vygotsky, L. S. (1978). *Mind in society: The development of higher mental processes.* In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds. & Trans.), Cambridge, MA: Harvard University Press. (Original work published in 1930).

Whaley, K. L., & Kantor, R. (1992). Mixed-age groupings in infant/toddler child care: Enhancing developmental processes. *Child & Youth Care Forum, 21,* 369–384.

Winkleby, M. A., Kraemer, H. C., Ahn, D. K., & Varady, A. N. (1998, July 22). Ethnic and socioeconomic differences in cardiovascular disease risk factors: Findings for woman from the third national health and nutrition examination survey, 1988–1994. *Journal of the American Medical Association, 280,* 356–362.

Winsler, A., Caverly, S. L., Willson-Quayle, A., Carlton, M. P., Howell, C., & Long, G. N. (2002). The social and behavioral ecology of mixed-age and same-age preschool classrooms: A natural experiment. *Applied Developmental Psychology, 23,* 305–330.

Appendix A

Multilevel Equations for Models 1–8

We estimated child-level (Level 1) Time 2 (T2) Child Observation Record (COR) scores by means of the following equation:

$$\text{T2 COR} = \beta_0 + \beta_1(\text{Child's Age}) + \beta_2(\text{T1 COR}) + \beta_3(\text{Gender}) + e,$$

where β_0 refers to the intercept (i.e., the average COR score at Time 2); β_1, β_2, and β_3 represent the maximum likelihood estimate of the population slopes estimating Time 2 COR score from the child's age, child's T1 COR score, and gender, respectively; and e is error.

We estimated classroom-level (Level 2) effects as follows:

$$\beta_0 = \gamma_{00} + \gamma_{01}(\text{Age Composition}) + \gamma_{02}(\text{Class Size}) + u_0,$$

$$\beta_1 = \gamma_{10} + \gamma_{11}(\text{Age Composition}) + u_1,$$

$$\beta_2 = \gamma_{20} + u_2, \text{ and}$$

$$\beta_3 = \gamma_{30} + u_3.$$

We estimated the combined model as follows:

$$\text{T2 COR} = \gamma_{00} + \gamma_{01}(\text{Age Composition}) + \gamma_{02}(\text{Class Size}) + \gamma_{10}(\text{Child's Age}) + \gamma_{11}(\text{Age Composition})(\text{Child's Age}) + \gamma_{20}(\text{T1 COR}) + \gamma_{30}(\text{Gender}) + u_0 + u_1(\text{Child's Age}) + u_2(\text{T1 COR}) + u_3(\text{Gender}) + e,$$

where γ_{00} refers to the child-level intercept for an average classroom; γ_{01} represents the maximum likelihood estimate of the population slope estimating average levels of Time 2 COR scores across all classrooms from classroom-level age composition; γ_{02} represents the maximum likelihood estimate of the population slope estimating average levels of Time 2 COR scores across all classrooms from classroom-level size (i.e., the number of students in the classroom) and u_0 represents error in estimating this intercept; γ_{10} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from child's age and u_1 represents error related to that estimate; γ_{20} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from Time 1 COR score and u_2 represents error related to that estimate; γ_{30} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from gender and u_3 represents error related to that estimate; and γ_{11} represents the maximum likelihood estimate of the cross-level interaction between child-level age and classroom-level age composition.

Appendix B

Multilevel Equations for Models 9–16

We estimated child-level (Level 1) Time 2 COR scores by the following equation:

$$\text{T2 COR} = \beta_0 + \beta_1(\text{Child's Developmental Age}) + \beta_2(\text{Child's Age}) + \beta_3(\text{Gender}) + e,$$

where β_0 refers to the intercept (i.e., the average COR score at Time 2); β_1, β_2, and β_3 represent the maximum likelihood estimate of the population slopes estimating Time 2 COR score from the child's developmental age, child's age, and gender, respectively; and e is error.

We estimated classroom-level (Level 2) effects as follows:

$$\beta_0 = \gamma_{00} + \gamma_{01}(\text{Developmental Age Composition}) + \gamma_{02}(\text{Class Size}) + u_0,$$

$$\beta_1 = \gamma_{10} + \gamma_{11}(\text{Developmental Age Composition}) + u_1,$$

$$\beta_2 = \gamma_{20} + u_2, \text{ and}$$

$$\beta_3 = \gamma_{30} + u_3.$$

We estimated the combined model as follows:

$$\text{T2 COR} = \gamma_{00} + \gamma_{01}(\text{Developmental Age Composition}) + \gamma_{02}(\text{Class Size}) + \gamma_{10}(\text{Child's Developmental Age}) + \gamma_{11}(\text{Developmental Age Composition})(\text{Developmental Age}) + \gamma_{20}(\text{Child's Age}) + \gamma_{30}(\text{Gender}) + u_0 + u_1(\text{Child's Developmental Age}) + u_2(\text{Child's Age}) + u_3(\text{Gender}) + e,$$

where γ_{00} refers to the child-level intercept for an average classroom; γ_{01} represents the maximum likelihood estimate of the population slope estimating average levels of Time 2 COR scores across all classrooms from classroom-level developmental age composition; γ_{02} represents the maximum likelihood estimate of the population slope estimating average levels of Time 2 COR scores across all classrooms from classroom-level size (i.e., the number of students in the classroom) and u_0 represents error in estimating this intercept; γ_{10} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from child's developmental age (i.e., Time 1 COR) and u_1 represents error related to that estimate; γ_{20} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from child's age and u_2 represents error related to that estimate; γ_{30} represents the maximum likelihood estimate of the population slopes estimating Time 2 COR score from gender and u_3 represents error related to that estimate; and γ_{11} represents the maximum likelihood estimate of the cross-level interaction between child-level developmental age and classroom-level developmental age composition.

Received December 19, 2006
Revision received January 3, 2008
Accepted January 7, 2008 ■