# Age of Peers and Early Childhood Development

**Donald B. Bailey, Jr., Margaret R. Burchinal, and R. A. McWilliam**

*Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill*

BAILEY, DONALD B., JR.; BURCHINAL, MARGARET R.; and MCWILLIAM, R. A. *Age of Peers and Early Childhood Development.* CHILD DEVELOPMENT, 1993, **64**, 848–862. We examined the developmental trajectories of 59 children between 21 and 67 months of age randomly assigned to same-age or mixed-age child care groups. Children were assessed at 6-month intervals with the Battelle Developmental Inventory. Using a hierarchical linear model for analysis, we computed individual and group growth curves for overall development and development within 5 domains: communication, cognitive, motor, adaptive, and personal-social. Group assignment affected the developmental trajectories for communication, cognitive, motor, and overall development, with mixed-age children showing a quadratic trajectory and same-age children a linear trajectory. For adaptive and personal-social development, however, the trajectories were similar. Across all domains, the mixed-age children tended to score higher than the same-age children at the younger ages, but these average differences decreased over time and had disappeared by age 5 years.

Same-age and mixed-age groups are assumed to constitute fundamentally different ecologies and thus affect behavior and development in different ways. The nature and direction of the effects, however, are not clear from existing literature, despite almost universal agreement that *something* is different. This was the context when Hartup (1976) wrote, "Thus we know little about cross-age interaction—either the circumstances of its occurrence in naturalistic socialization or its functional significance. Intuition suggests that children should be influenced differently by peers who are younger or older than they are by peers who are roughly their own age. But this supposition has only the anecdotal reports of sharp-eyed group workers, teachers, and parents to support it" (p. 4).

Since that time, a small number of studies have sought to identify outcomes that reflect the differential impact of homogeneous and heterogeneous age groupings. Most of the studies have involved relatively short-term observations of children in dyadic contexts to determine effects on social interactions and communication. Virtually nothing is known about the longitudinal effects of age mixture on children's development, despite indications that such effects are likely. Although all children naturally experience both same-age and mixed-age groups, it is in the context of group child-care settings where the issue becomes especially relevant because of the potential effects on cognitive and social development.

Several studies suggest that mixed-age groups may enhance the development of younger children. Exposure to more competent models of behavior, the availability of older "tutors," and the likely occurrence of developmentally complex games, routines, and activities all point to the potential benefits of mixed-age settings for younger children. Brownell (1990), for example, observed dyads of 18- and 24-month-old toddlers and found that the 18-month-old children initiated more, produced more complex initiations, and imitated partners more when paired with an older child than with an age-mate. Howes and Farver (1987) found that 2-year-old toddlers engaged in more social pretend play when paired with 5-year-old partners than when paired with other 2-year-old age-mates. Lougee, Grueneich, and Hartup (1977) observed dyads of preschoolers and concluded that younger children engaged in higher levels

This study was funded in part by a grant from the U.S. Department of Education (G008730079). Requests for reprints should be addressed to Don Bailey, Frank Porter Graham Child Development Center, CB# 8180, University of North Carolina, Chapel Hill, NC 27599.

of positive social interactions when paired with a companion who was 16 months older than when paired with a same-age peer. In contrast, Graziano, French, Brownell, and Hartup (1976) found that first graders placed in triads with other first graders and given a timed performance task talked more and contributed more to task success than did first graders placed in triads containing one or two third graders.

Additional support for the benefits of mixed-age groupings is found in a study by Furman, Rahe, and Hartup (1979). Building on earlier animal research by Suomi and Harlow (1972), the authors paired normally developing children who exhibited low rates of peer social interactions with age-mates or with younger children for 10 dyadic play sessions. Both groups increased their socialization skills as a result of the play sessions. The effect was especially strong, however, for the children who were paired with younger playmates. This finding is consistent with Suomi and Harlow's (1972) finding that isolate-reared monkeys gained social skills when paired with immature playmates but not when paired with same-age, normally reared peers.

The developmental benefits of mixed-age groupings for older children are less obvious; indeed, some evidence suggests the possibility of negative effects. Lougee et al. (1977), for example, found that older children in mixed-age dyads engaged in fewer positive social interactions than older children in same-age dyads. Brownell (1990) also found that 24-month-olds made fewer social overtures to 18-month-old dyadic partners than to other 24-month-old dyadic partners. Initiations to younger partners, however, were longer and more complex; the author concluded that this may have been an effort to enhance the salience of the behavior as a social overture. Goldman (1981) found that 4-year-olds placed in mixed-age groups with younger children engaged in more solitary play than 4-year-olds placed in same-age groups. Since the complexity of play was not coded in this study, the developmental effects of this finding are uncertain. Several studies have demonstrated that older children adjust their communication skills to the level of the play partner (e.g., Guralnick & Paul-Brown, 1984, 1986; Shatz & Gelman, 1973). Although this obviously benefits the younger children, it is unlikely that abbreviated communication patterns have any developmentally beneficial effects for older children.

Finally, in related literature on the development of children in a family context, Zajonc (1986) argued that the birth of a new child reduces the "intellectual climate" in a home and subsequently has a negative effect on the development of older children and family members. Although Zajonc's model and data have been criticized for failure to account for a number of key variables (e.g., Rodgers, 1988), the relative role of siblings and birth order on child development remain an important issue. McCall's (1984) longitudinal work found that the IQ scores of children who experienced the birth of a younger sibling dropped an average of 10 points compared with singleton children. These findings lend further support to the possibility that mixed-age groups may not benefit the development of older children.

The lack of longitudinal data makes any conclusion about the long-term developmental outcomes of same- versus mixed-age groupings speculative at best. Thus the purpose of the study reported in this article was to determine the effects of age of peers on early childhood development. We hypothesized that mixed-age groups would result in developmental benefits, across a number of domains, for younger children exposed to older peer models. The article is also intended to demonstrate how use of a rotating panel design and a hierarchical linear model of analysis can enhance the longitudinal study of development.

## Method

The study involved the longitudinal developmental assessment of children randomly assigned to same-age or mixed-age groups in a university-based child-care program. The period of study was from fall of 1986 through spring of 1990.

### *Subjects*

A total of 59 children participated in the study. Criteria for participation were random assignment to a same-age or mixed-age group and enrollment in the assigned group for a minimum of 1 year. The children ranged in age at time of assessment from 21 to 67 months. The group consisted of 27 girls and 32 boys. Fifty-four percent were members of a minority group, predominantly African-American. Approximately one-third of the children were from low-income families, one-third from middle-income families, and one-third from upper-income families. Minority and income status were equally distributed across the ages as well as across experimental conditions. None of the children reported on here had a diagnosed handicapping condition or developmental delay.

*Setting*

*Classroom information.*—The children were enrolled in one of eight child-care groups in a university-based day-care program; each group consisted of six children (two handicapped and four normally developing) and a day-care teacher. Each group had a self-contained space designed to meet the needs of the children in the group. This included a variety of play areas, toys and materials placed on low shelves for easy access, cots for naps, and access to child-sized sinks and toilets. Classroom areas were separated by low dividers approximately 4 feet high.

*Teachers.*—Nine teachers worked in the eight classrooms used in this study; the room for 1-year-olds had two teachers in order to provide an appropriate adult-child ratio. All of the teachers were women, with a mean age of 46.5 years (range = 33.4–59.5), a mean length of employment at the center of 15.3 years ($SD$ = 3.7), and mean education (expressed in grades completed, including higher education equivalents of one grade level for each year completed) of 13.9 ($SD$ = 1.5).

*Curriculum.*—All children at the center participated in the Learningames curriculum (Sparling & Lewis, 1979, 1984). The curriculum consists of 200 carefully developed but simple games to encourage learning and social-emotional growth. A typical schedule for the day-care center consisted of arrival and free play (7:30–9:00), breakfast (9:00–9:30), diapering and cleanup (9:30–10:00), planned group and individual activities (10:00–10:45), circle and story time (10:45–11:00), outdoor play (11:00–11:30), preparation for lunch (11:30–11:50), lunch and cleanup (11:50–12:30), nap and quiet activities (12:30–3:00), getting dressed and snack (3:00–4:00), and free play (4:00–4:30). During the free play times at the beginning and the end of the day, the children were grouped by age (all 1-year-olds together, all 2-year-olds together, etc.). Between 9:00 A.M. and 4:00 P.M., children were in their assigned experimental group.

*Assignment to groups.*—The children were assigned to eight groups, four consisting of children of mixed ages and four consisting of children of the same age. The four same-age groups each consisted of children within 12 months of age of each other. Thus there were a class of 1-years-olds, a class of 2-year-olds, a class of 3-year-olds, and a class of 4-year-olds. The four mixed-age groups consisted of three older children and three younger children, with an average difference of 25.5 months between the older and younger children. A 2-year age difference was used to approximate family-like spacing of ages and to maximize group heterogeneity. Two mixed-age groups consisted of 1- and 3-year-olds, and two contained 2- and 4-year-olds. Two children with disabilities were enrolled in each group, but were not included in the present study.

The same-age classes remained intact, with the exception of three children who left the day-care center, across the study period. For example, the class of 2-year-olds in 1987 became the class of 3-year-olds in 1988. The mixed-age classes remained intact for a 2-year period. For example, the class of 1- and 3-year-olds in 1987 became the class of 2- and 4-year-olds in 1988. Each mixed-age class changed composition after 2 years due to movement of older children into kindergarten. For example, in the fall of 1989, the 4-year-olds (who by that time were mostly 5-year-olds) in a class of 2- and 4-year-olds moved on to kindergarten. The remaining children became the 3-year-olds in a class of 1- and 3-year-olds, with a new group of 1-year-olds added to the class.

Assignment of groups used a stratified random assignment procedure. First, children were matched on the basis of age, sex, race, and socioeconomic status (based on Hollingshead's [1975] Four-Factor Index of Social Status). Then they were randomly assigned to either a same-age or mixed-age group. The distribution of these variables across treatment groups is displayed in Table 1. The equivalence of the groups at the time they entered the study was examined by comparing groups on gender, age, race, SES, and the total and domain (cognitive, communication, adaptive, personal-social, motor) scores of the Battelle Developmental Inventory (BDI; Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984). No differences between groups were found on gender, age, race, or SES.

Comparison of the BDI scores was complicated by the fact that at recruitment many children were too young for developmental assessments to show good across-time reliability (i.e., were assessed initially at ages that were excluded from subsequent analyses). A two-way analysis of variance (AN-

TABLE 1

COMPARABILITY OF MIXED-AGE AND SAME-AGE GROUPS AT FIRST ASSESSMENT

| | TREATMENT GROUP | |
|---|---|---|
| | Mixed-Age ($n$ = 28) | Same-Age ($n$ = 31) |
| Age in months at entry into study | 26.8 (11.6) | 29.5 (14.0) |
| Percentage African-American | 57.1 | 51.6 |
| Percentage female | 42.9 | 48.4 |
| Mean BDI score: | | |
| Total* | 112.5 (15.1) | 104.9 (12.2) |
| Adaptive* | 114.0 (15.7) | 106.7 (12.0) |
| Cognitive | 106.8 (12.5) | 106.6 (13.5) |
| Communication | 107.2 (16.2) | 102.0 (12.1) |
| Motor** | 108.6 (11.1) | 100.3 (13.0) |
| Personal-Social | 112.4 (16.2) | 105.7 (13.0) |

NOTE.—Numbers in parentheses are the standard deviations.
* $p$ between .05 and .10.
** $p < .05$.

OVA) was used to determine whether age at initial testing or treatment condition were related to mean differences in initial BDI scores. Age at initial testing was included to determine whether the scores for toddlers under 21 months of age differed from those of the children who entered the study at older ages. Treatment condition was included in the analysis to determine whether random assignment resulted in developmentally equivalent groups. No differences were found as a function of age at first assessment. Although the two treatment groups were virtually identical in cognitive development, a statistically significant difference was observed on the Motor scale, $F(1, 55) = 5.79$, $p = .02$, and marginally significant differences were found on the total score, $F(1, 55) = 3.73$, $p = .059$, and the Adaptive scale, $F(1, 55) = 3.89$, $p = .054$. These comparisons suggest that although the two groups of children were both within the average range of development, the mixed-age children tended to be slightly more advanced than the same-age children at entry into the study. As described in the analysis and results sections, post hoc analyses using initial BDI scores as a covariate were run to determine whether our results were influenced by this difference.

The study employed a rotating panel design (i.e., a design in which children from several birth cohorts are studied, resulting in incomplete longitudinal data when age is the time metric of interest; Nesselroade & Baltes, 1979). In this design, seven birth cohorts were followed for 1 to 4 years. This design, illustrated in Figure 1, involved the random assignment in 1986 of children in the first four birth cohorts, all of whom were already enrolled in a university-based child-care program, to treatment groups (eight 4-year-olds, eight 3-year-olds, eight 2-year-olds, and eight 1-year-olds). These children remained in the study until they left the center for kindergarten in the fall of their sixth year. Each fall from 1987 to 1989 another birth cohort of eight 1-year-olds was randomly assigned to treatment groups and followed until the study ended in the spring of 1990. A total of 56 children were assigned to treatment in this manner. Three children left the center at some point during the study (at about 3.5, 4.5, and 5.5 years) and were replaced by new children with similar ages, thus accounting for the total sample of 59 children. In summary, 32–33 children were observed at each age, but only eight children participated for a full 4 years.

This design has the advantage of allowing longitudinal analyses to be performed, while enabling valid cross-sectional estimates to be made for the ages of the children included in the study. In particular, longitudinal analyses can address questions about the relation between predictors and patterns of development for the observed ages, under the assumption that there are no age × cohort interactions (e.g., that development between 1 and 5 years of age was affected by the same factors whether you were born in 1982 or 1988). By assuming that tem-

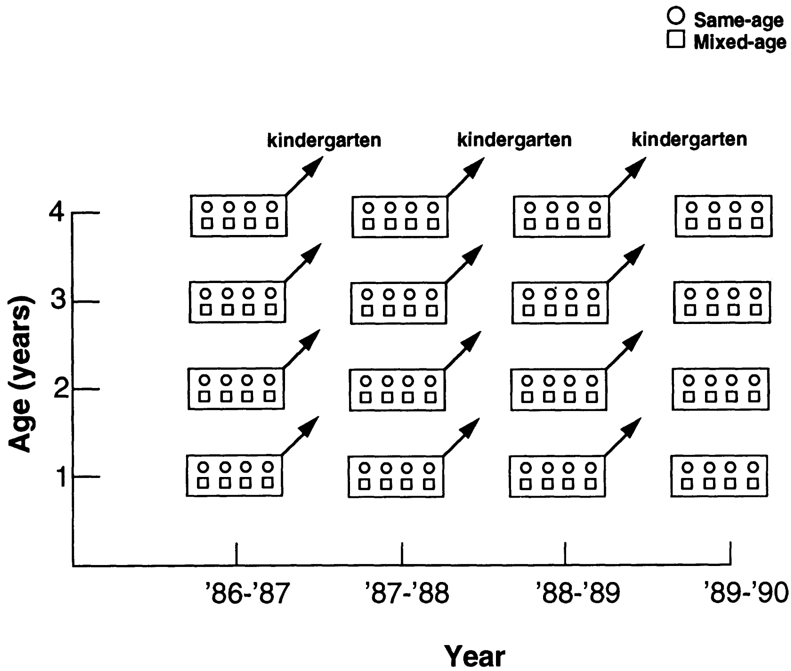



FIG. 1.—Ages and group assignments of children by year

poral factors can be ignored, the incomplete data for most children can be used to estimate developmental trajectories. While not as powerful as a study in which 59 children are followed longitudinally from 1 to 5 years of age, the design offered a practical alternative at much less expense and time.

*Instrumentation*

We had hypothesized that exposure to older peers might influence children's development in a broad array of developmental domains. Thus we needed a measure that not only assessed the basic areas of cognitive and language development, but also described motor, social, and adaptive development. For this reason, and because the model of analysis used would be enhanced if a common instrument could be used across domains, children's developmental status was assessed using the BDI. The BDI assesses development of children from birth to 8 years of age and consists of 341 items distributed across five domains and 22 subdomains: (1) Personal-Social (adult interaction, expression of feelings/affect, self-concept, peer interaction, coping, and social role); (2) Adaptive (attention, eating, dressing, personal responsibility, and toileting); (3) Motor (muscle control, body coordination, locomotion, fine muscle, and perceptual motor); (4) Communication (receptive and expressive); and (5) Cognitive (perceptual discrimination, memory, reasoning and academic skills, and conceptual development). Most of the scores are based on children's responses to items directly administered. Some items may be scored on the basis of an interview with a knowledgeable caregiver, and the personal-social items are based on observations of freely occurring behavior. Score summaries include a developmental age for domain and total scores, as well as standard scores ($z$, $T$, deviation quotient, or normal curve equivalent) for total, domain, and subdomain scores. The examiner's manual reports high levels of test-retest reliability, with correlation coefficients ranging across age groups from .90 to .99 for total scores and from .84 to .98 for domain scores.

*Procedure*

The BDI was administered to each child at 6-month intervals throughout the period of participation in the study. Assessments occurred during a 6-week period in early fall of each year and 6 months later in the spring. The BDI was administered by a number of specially trained doctoral students in special education or school psychology. The testers each had extensive prior experience in assessing young children with a variety of standardized measures. Each participated in a 1-day training session on administering the

BDI and practiced on several nonstudy children until familiar with the measure and materials. Scoring of 100% of the protocols was checked by a second rater to detect any errors in calculating raw scores or converting to age equivalents. A number of different testers were involved over the course of the study; none tested the children on more than two occasions. All testers were blind to the purpose of the study and to group assignment.

Each child was assessed on at least two occasions. Of the 59 children, 20 received two assessments, 7 received three assessments, 14 received four assessments, 4 received five assessments, and 14 received six assessments. The mean age at first testing was 31.75 months (SD = 10.15), and the mean age at last testing was 49.47 months (SD = 14.24). The children ranged between 22 and 67 months of age at assessment. The mean length of time in the study was 17.73 months (SD = 11.7).

While the BDI was administered to all study children, only the BDI scores of tests when the children were over 20 months of age were included in the analyses. The test scores from children between 12 to 20 months were excluded because it is probable that the traits measured by developmental tests of children less than 21 months of age are different from those measured after verbal skills are assessable (Harris, 1983). Also, the BDI scores for the children with disabilities were excluded because of the small sample size and the heterogeneous nature of the group.

*Data Analysis*

The hierarchical linear model (HLM) approach to analyzing unbalanced, incomplete longitudinal data was employed (Bryk & Raudenbush, 1987; Hocking, 1985; Laird & Ware, 1982). Children were observed in the spring and fall between the ages of 21 months and entry to kindergarten while they participated in the study. The number of observations per child ranged from two to six, depending on the year and the child's age at entry into the study.

The HLM (sometimes called a random effects or general linear mixed model) was selected for several reasons. First, traditional repeated-measures methods could not be used for this study because of the design. While data were collected on 59 children of ages 1 to 5 years, they were collected at all ages on only eight children. Thus, only the data from these eight children could be included in longitudinal analyses using traditional methods such as the multivariate repeated-measures ANOVA, the univariate-corrected ANOVA, or structural modeling. In contrast, the HLM approach estimates individual curves based on the data collected on each child, allowing children with missing data to be included under the assumption that there are no time or cohort × age interactions (e.g., that classroom type had a similar impact on the development of children of a particular age regardless of whether they were that age in any year between 1986 and 1990).

Furthermore, the HLM is designed to allow examination of individual patterns of development and interindividual differences in these patterns. Traditional methods estimate population regression curves that represent the expected or average patterns of development over time exhibited within a group of interest. These methods adjust or take into account the lack of independence among repeated observations from a given individual, but are not designed to estimate that individual's pattern of development. One purpose of this study was to examine individual developmental trajectories, so the HLM was the more appropriate analytic approach.

Using the HLM, we can estimate individual and group growth curves under the assumption that the parameters of the individual curves are normally distributed about the parameter values of the population curves and that within-subject variation about the individual growth curve is independent and identically distributed for all individuals (see Hocking, 1985, and Rogosa, Brandt, & Zimowski, 1982, for more detail). The HLM estimates individual and population growth curves from fixed and random variables specified in the analysis model. The fixed effect variables are used to estimate the population growth curves, while the random effect variables are used to estimate the variation between individual and population growth curves. In our study, the dependent measures observed over time are the child's BDI scores, and the predictor variables are classroom type and the ages at which the child was observed. We assumed that a cubic polynomial curve should be sufficient to describe the relation between chronological age and developmental age; thus, Age, $Age^2$, and $Age^3$ were the specific age variables entered into the model.

Our specific model represents the developmental level of the $i$th individual in the $k$th group as a separate cubic regression curve for each individual. The parameters of each individual's curve were assumed to be randomly distributed about his or her group cubic regression curve. Thus, the statistical model represents the observed score of the $i$th individual in the $k$th group at $\text{Age}_j$ as a linear combination of the $k$th group's regression parameters, the $i$th individual's deviation from those parameters, and random error:

$$\begin{aligned}\text{Battelle score}_{ijk} = {} & \mu + \alpha_k + \beta_{1k}(\text{Age}_{ijk}) \\ & + \beta_{2k}(\text{Age}_{ijk})^2 \\ & + \beta_{3k}(\text{Age}_{ijk})^3 \\ & + \delta_{ik} + \lambda_{1ik}(\text{Age}_{ijk}) \\ & + \lambda_{2ik}(\text{Age}_{ijk})^2 \\ & + \lambda_{3ik}(\text{Age}_{ijk})^3 + \epsilon_{ijk}.\end{aligned}$$

The $k$th group population growth curve (see Fig. 2 for estimated group growth curves) is estimated from fixed effect variables as

$$\mu + \alpha_k + \beta_{1k}(\text{Age}_{ijk}) + \beta_{2k}(\text{Age}_{ijk})^2 + \beta_{3k}(\text{Age}_{ijk})^3.$$

In this model, the $k$th group's regression curve is defined by the following parameters: $\mu$ is the general intercept; $\mu + \alpha_k$ is the intercept for the $k$th group's curve; and $\beta_{1k}$, $\beta_{2k}$, and $\beta_{3k}$ are the slopes with respect to Age, $\text{Age}^2$, and $\text{Age}^3$ for the $k$th group's curve. Similarly, the growth curve for the $i$th individual in the $k$th group (see Fig. 3 for estimated individual growth curves) is defined from random and fixed effect variables as

$$(\mu + \alpha_k + \delta_{ik}) + (\beta_{1k} + \gamma_{1ik})(\text{Age}_{ijk}) + (\beta_{2k} + \gamma_{2ik})(\text{Age}_{ijk})^2 + (\beta_{3k} + \gamma_{3ik})(\text{Age}_{ijk})^3,$$

where $\delta_i$ (the random effect intercept) is the difference between the individual and

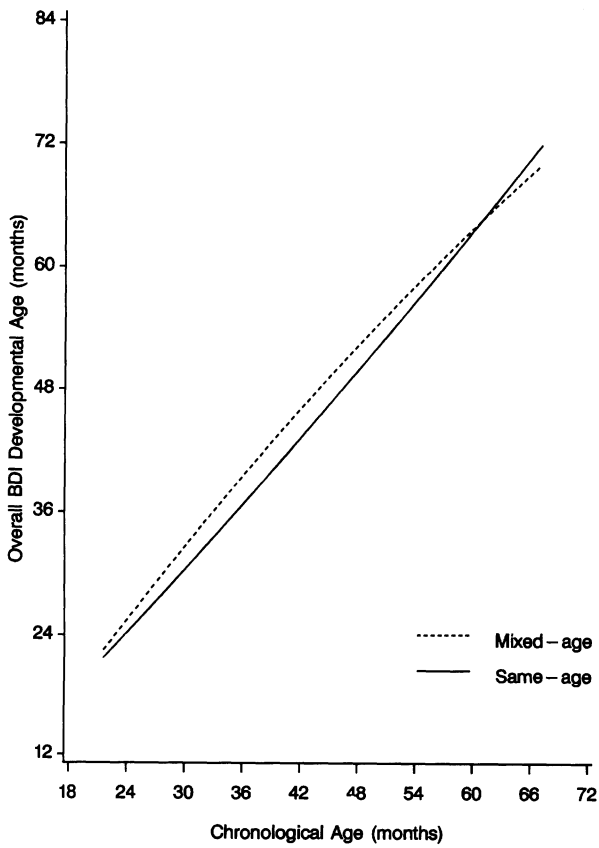
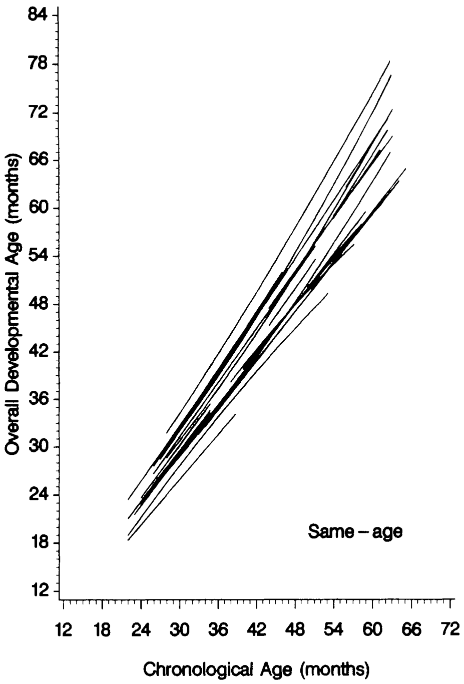



FIG. 2.—Estimated group growth curves for overall development of children in same-age and mixed-age groups.



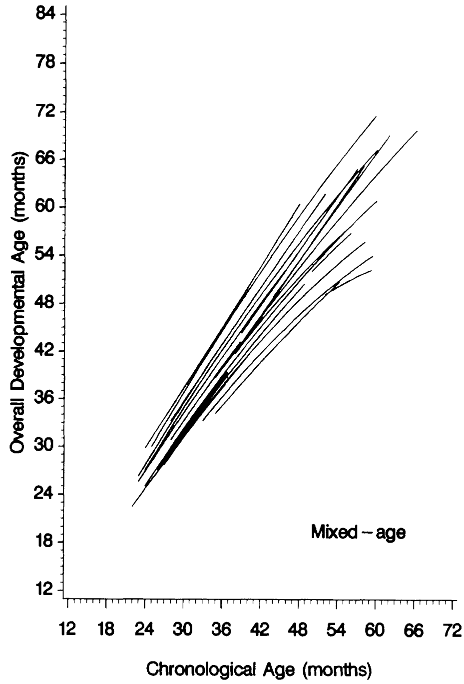




FIG. 3.—Estimated individual growth curves for overall development of children in same-age and mixed-age groups.

group intercept, and $\gamma_{1ik}$, $\gamma_{2ik}$, and $\gamma_{3ik}$ are the differences between the estimated random effect slopes for the $i$th individual's growth curve and his or her group growth curve with respect to Age, Age$^2$, and Age$^3$. Finally, $\epsilon_{ijk}$ is the random error for the $j$th time that the $i$th individual in the $k$th group was observed. In summary, group and individual growth curves are estimated under the assumption that the observed scores have a fixed component (i.e., the group parameter estimates) and two independent random components (i.e., the individual's parameter estimates and the within-individual variation).

This model was first used to analyze the overall developmental age scores. When significant group differences emerged, the five domain scores were also analyzed with this model. The overall analysis plan involved fitting separate regression curves for each of the two groups in a single analysis that included Age, Age$^2$, and Age$^3$ as predictors, assuming common error. The model was reduced, estimating a common parameter for the age variables when group differences were nonsignificant (i.e., if the test for different cubic terms was nonsignificant, then a common cubic term replaced the separate cubic terms in the model). This process continued until the most parsimonious model for each set of scores was identified, that is, until the model consisted of treatment group, age, and any statistically significant higher order age effects or age $\times$ group interaction terms. When age $\times$ group interactions or higher-order age effects were present, then all lower-order terms contained in that interaction were also retained in the model.

Regression curves were computed for the two groups and compared in two ways. First, we compared the coefficients that define the curves to identify how the average growth curves differed for the two groups. Using this approach, we determined what regression curve best described development for each group and then compared the estimated intercepts and slopes of the two regression curves to determine how the developmental growth curves differed between the two groups. Second, we computed confidence bands around the differences between these two curves to determine at which ages the groups differed. The confidence bands were inferred from results of five a priori contrasts (Stewart, 1987). These comparisons employed a Bonferroni correction to account for the number of multiple comparisons. The expected values of the two groups were compared at 24, 33, 42, 51, and 60 months of age. These ages were selected because more children were tested at these ages than at younger or older ages. Then, a confidence band was inferred by connecting the 95% confidence intervals about these comparisons.

Two post hoc analyses were conducted. First, we tested whether our findings were influenced by the initial differences between treatment groups. The final set of growth curves models were adjusted to include the first BDI score as a covariate in the analysis, predicting the children's subsequent scores from treatment, age, treatment × age, and initial BDI score.

A second post hoc analysis was performed to determine whether the amount of time that each child had been in the study influenced their developmental trajectories. This analysis was performed because the children participating in the study had varying amounts of treatment exposure. The amount of time that each child spent in the assigned same- or mixed-age classroom and an interaction between classroom type and time-in-treatment were entered as time-varying covariates into the final regression models for the BDI total score and for each domain score.

## Results

Significant differences were detected between the regression curves describing overall development for children in the same-age groups and for children in the mixed-age groups (see Figs. 2 and 3 for group and individual growth curves). The cubic term was nonsignificant and was dropped from the model. A linear regression curve described the relation between chronological age and developmental age for the children in the same-age groups, but a quadratic regression curve was necessary to describe this relation for the mixed-age group. This difference can be seen in Figure 2; the same-age group curve is a straight line, but the mixed-age group curve bows up between the ages of 22 to about 40 months and then bows down after the age of about 40 months. Descriptively, these curves suggest that children in the two classes look similar at 2 years of age, that mixed-age children may have an advantage over same-age children between the ages of 2 and 4, but that same-age children catch up and may surpass the mixed-age children by entry to kindergarten. Post hoc analysis indicated that amount of time in treatment was not a significant predictor, given age and treatment condition. Also, using the child's initial BDI total score as a covariate did not change the findings obtained.

Statistical comparisons confirmed that different patterns of development were observed in the two types of classroom. The degree of quadratic curvature ($\chi^2 = 14.44$, $df = 1$, $p = .0001$) was greater for the mixed-age group, and the intercept (i.e., mean level at 36 months) was higher for the mixed-age group ($\chi^2 = 6.83$, $df = 1$, $p = .009$). However, no significant group differences were detected when we used the confidence band approach to determine at which ages the two groups significantly differed. These seemingly discrepant results indicate we may have lacked sufficient power to identify the specific ages at which the two groups differed with our modest sample size, while we could address the more general question of whether the curves were identical across time.

Similar results were observed for the Cognitive, Communication, and Motor Development scales. The same general patterns of development were observed in these three domains as in the overall index; quadratic growth characterized development in the mixed-age classes and linear growth best described growth in the same-age classes (see Figs. 4, 5, and 6). The quadratic slope parameters were significantly different for the Cognitive scale ($\chi^2 = 7.68$, $df = 1$, $p = .006$) and the Motor scale ($\chi^2 = 9.83$, $df = 1$, $p = .002$). In addition, the regression curves describing motor development for the two groups showed significantly different linear slopes ($\chi^2 = 4.90$, $df = 1$, $p = .027$) and intercepts ($\chi^2 = 10.19$, $df = 1$, $p = .0014$). Post hoc analysis indicated that these treatment effects were still observed when time-in-treatment was used as a covariate. Amount of treatment exposure did show a significant main effect on communication development ($\chi^2 = 3.86$, $df = 1$, $p = .049$), and a marginal interaction with treatment condition was found in communication ($\chi^2 = 3.38$, $df = 1$, $p = .066$). These time effects indicated that the longer a child was in the study, the better their communication development tended to be, and this time-in-treatment effect on communication development may have been stronger for mixed-age children. Again, however, using time-in-treatment as a covariate did not change the significant findings observed.

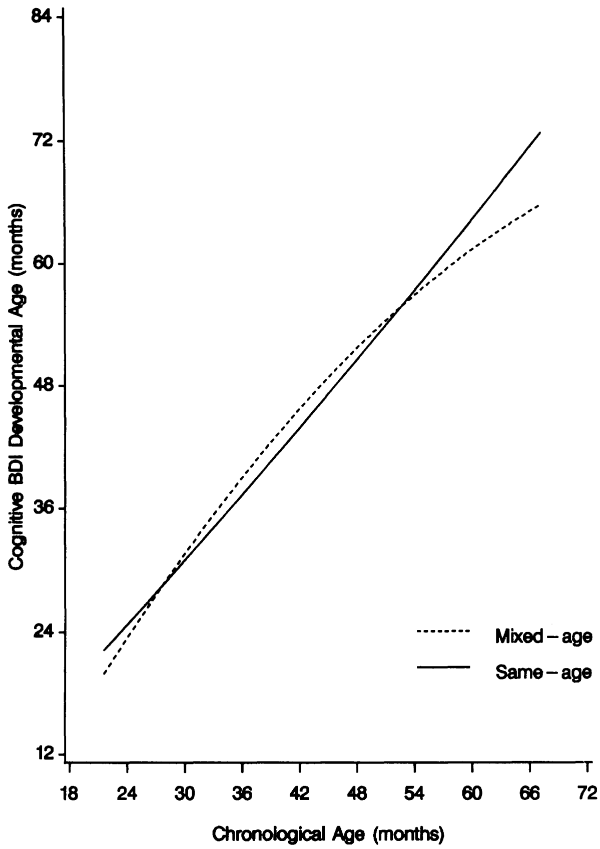



FIG. 4.—Estimated group growth curves for cognitive development of children in same-age and mixed-age groups.

Also, using initial BDI scores as covariates did not change the results.

Estimated confidence bands about the group regression curves indicated significant differences at specific ages only for motor development. Between the ages of 36 and 48 months, children in the mixed-age classes tended to outscore children in the same-age classes on the Motor scale.

Different patterns of results were observed for the two scales assessing social and adaptive development (see Figs. 7 and 8). Differences in intercepts, but not in slopes, characterized the patterns of development observed in these two domains. The children in the mixed-age group tended to perform better than children in the same-age groups at the younger ages, with the size of the difference decreasing over age. Post hoc analysis indicated that amount of time in treatment was not a significant predictor for adaptive development; for personal-social adjustment, a positive association ($\chi^2 = 5.02$, $df = 1$, $p = .025$) was observed, such that the longer the child participated in the study, the higher the Personal-Social scale score tended to be. However, no time × treatment interaction was found. Again, using time in treatment or BDI initial scores as covariates did not change the results of these analyses.

No significant group x time differences emerged in the analysis of the Personal-Social scale scores. However, the intercept for the mixed-age group was significantly higher than the intercept for the same-age group ($\chi^2 = 5.34$, $df = 1$, $p = .021$), suggesting a modest advantage for the mixed-age grouping. No differences at any age were detected when confidence bands were estimated. Neither time-in-treatment nor initial BDI scores substantively changed the results of these analyses.

The patterns of development of adaptive behaviors for the two groups showed significant differences in intercept ($\chi^2$ =

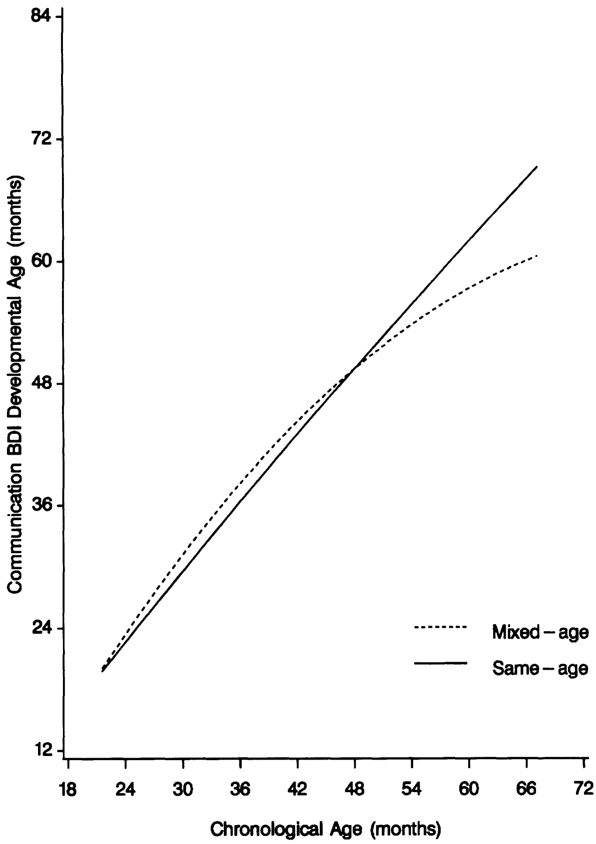



FIG. 5.—Estimated group growth curves for communication development of children in same-age and mixed-age groups.

15.38, $df = 1, p < .0001$) and in linear slope ($\chi^2 = 5.85, df = 1, p = .016$). Significant differences in favor of mixed-age classes were observed between 22 and 36 months of age when confidence bands were computed. Again, neither time-in-treatment nor initial BDI scores substantively changed the results of these analyses.

## Discussion

This study investigated the extent to which the development of young children is influenced by placement in same-age or mixed-age child care groups. Previous research has suggested that mixed-age groupings may be beneficial for younger but not older children; no studies, however, have addressed the longitudinal impact of same-age versus mixed-age groups. Data from 59 children who participated in the study for 1 to 4 years were examined using growth curve analyses to determine differences in growth patterns.

The first major finding of the study is that classroom type appears to affect the developmental trajectories of children for communication, motor, cognitive, and overall development. In these four developmental areas, children in the mixed-age class tended to show a quadratic trajectory between 22 and 67 months, while children in the same-age classes tended to show a linear trajectory. Children in the mixed-age classes tended to score higher than children in same-age classes between 2 and 4 years of age, but this trend appeared to be reversed at the oldest ages, suggesting that mixed-age classes may be better for younger children and that same-age classes are better for kindergarten preparation at age 4 and beyond. However, these conclusions must be viewed with caution. While mean differences favoring younger children in mixed-age groups were statistically significant, the differences favoring older children in same-age classes were not significant. Thus, our data show trends that we may have lacked power

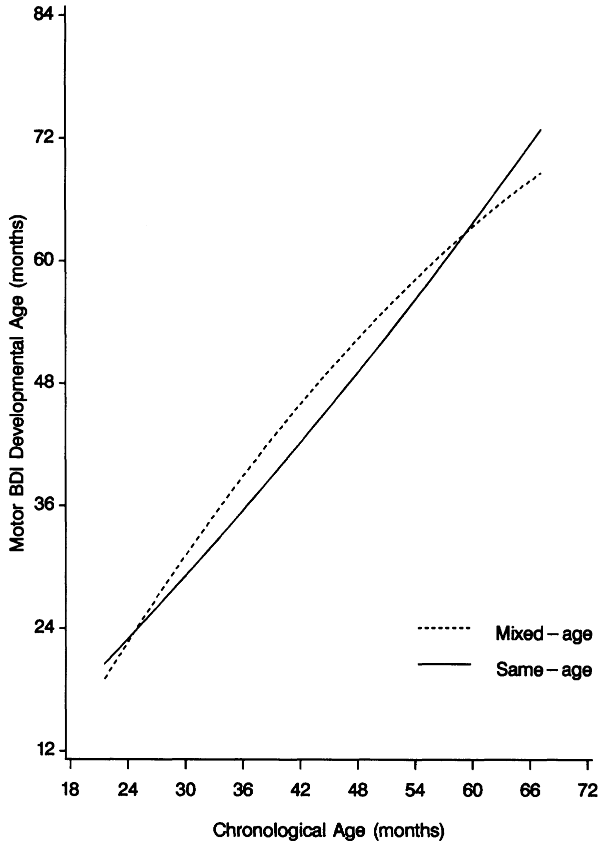



FIG. 6.—Estimated group growth curves for motor development of children in same-age and mixed-age groups.

to confirm statistically. Given our moderate sample size ($n = 59$), our power to detect differences the magnitude of those observed after 4 years of age was small, ranging from .01 to .20.

The other major finding is that for adaptive behavior and personal-social development, children in both classes tended to show similar types of patterns over time. The mixed-age children tended to score higher than the same-age children at the younger ages, but these average differences decreased over time and had disappeared by age 5 years.

The data from this study need to be interpreted cautiously for at least four reasons. First, the sample size is relatively small, especially for children with 4 years of data. We addressed this concern by using analytic procedures appropriate for the design and sample size and by computing confidence bands with respect to each analysis. Second, the fact that children participated in the study for different lengths of time and that variability in participation time was confounded with age may have influenced the pattern of results in some fashion. While post hoc analyses indicated that amount of treatment exposure showed neither a significant main effect nor a significant interaction with treatment condition, modest setting × time-in-setting interactions may have been too small to be detected. Third, the presence of children with disabilities added to the heterogeneity of groups, especially the same-age groups. Although this important limitation must be acknowledged, we should note that this would most likely serve to reduce any differences observed in same-age versus mixed-age children. The fact that differences emerged anyway suggests the possibility of larger differences had children with disabilities not participated in the study. Finally, because of the study design, many of the children served as both older and younger subjects in the mixed-age groups. For example, a child could, in 1 year,

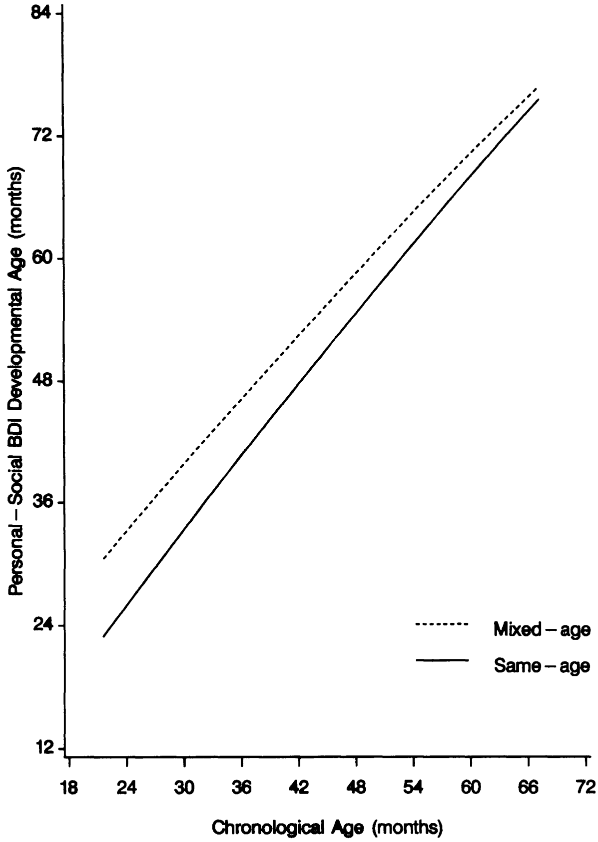



FIG. 7.—Estimated group growth curves for personal-social development of children in same-age and mixed-age groups.

be a 2-year-old in a group of 2's and 4's, and the next year be a 3-year-old in a group of 1's and 3's. Thus our analysis is confounded by this necessary transition between 2 and 3 years of age.

Despite these limitations, several aspects of the study suggest that it is internally valid. These include random assignment to groups; equal distribution of gender, race, and socioeconomic status; provision of a standard curriculum; constant group size; blind assessments; and use of sophisticated analytic techniques. Nevertheless, replication studies are needed to verify the findings of the present study.

The results extend previous studies suggesting that mixed-age groups may be beneficial for younger children, but that this advantage is not maintained when the children are 4–5 years of age, at least within the basic domains of cognitive, adaptive, motor, social, and communication development. The factors contributing to the observed effects are likely to be complex and are not easily determined on the basis of the present study. Several hypotheses are offered, however. With respect to younger children, the observed benefits of mixed-age settings may be explained in several ways. First, the older children may have provided competent models for more advanced behavior. Bandura's (1969) research suggests that children are more likely to imitate individuals whom they perceive to be more competent or have higher status. Second, the older children may have created an environment that was more stimulating or exciting for the children, resulting in greater interest and involvement in activities, as suggested by other studies of the effects of age grouping on play and social skills (e.g., Brownell, 1990; Howes & Farver, 1987; Lougee et al., 1977). Third, the older children may have played the role of tutor or teacher for the younger children by instructing them to engage in certain activities or by giving them feedback on certain behavior. Finally, the teachers probably pro-

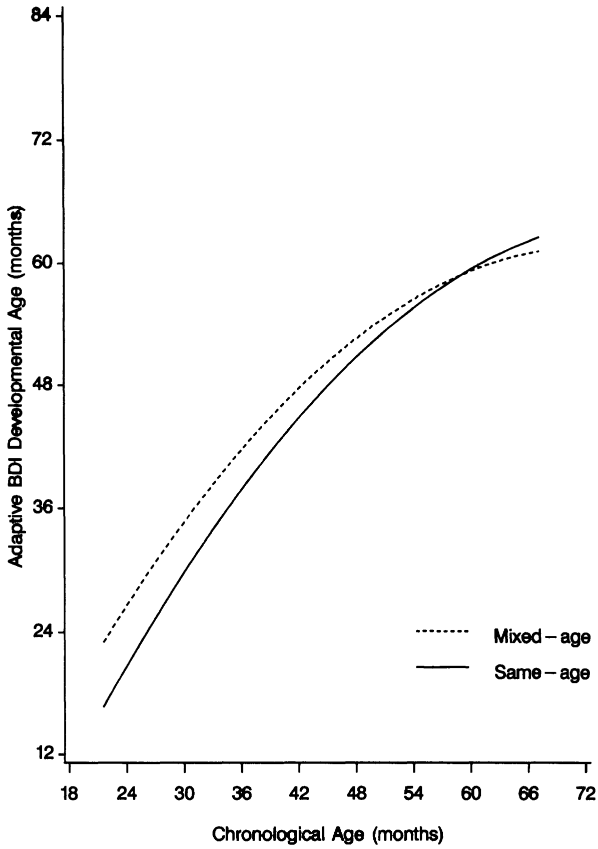



FIG. 8.—Estimated group growth curves for adaptive development of children in same-age and mixed-age groups.

vided a greater variety of activities in mixed-age settings, exposing younger children to experiences that otherwise might not have been available until later.

With respect to the oldest children, the lack of differences of same-age groups also may be explained by several factors. One possibility is that the effects were limited to the early age period and that placement in same-age groups at older ages would not have maintained these advantages. Another possibility is that the presence of younger children resulted in fewer opportunities for developmentally beneficial stimulation than otherwise would have been available. A third possibility is that the younger children may have interfered with the activities of older children to such an extent that conversations, elaborate games, or other more complex interactions were interrupted. Dunn and Shatz (1989), for example, found that an average of 22% of young children's conversational turns were intrusions. Finally, the teacher may have spent more time attending to the caregiving demands of the younger children, thus limiting opportunities for interaction with older children.

The data presented in this article suggest that opportunities for mixed-age interaction may have some developmental benefits for younger children, but that this advantage is not maintained when children are 4–5 years of age. However, nonsignificant trends suggest that older children may profit from same-age experiences. If future research maintains these trends, child-care providers may want to consider implementing flexible scheduling to provide both types of groupings. For example, outdoor free play, meals and snacks, field trips, and similar activities may be provided for mixed-age groups, whereas same-age groups may be optimal for story time, art, and specific curriculum activities.

The results and hypotheses suggested by these data are consistent with findings and hypotheses generated from previous observational research. However, they also point to the need for further research investigating the effects of age of peers on children's development. Needed are (*a*) studies with variations in group size to test whether increasing the number of children in the group provides more same-age opportunities for older children, (*b*) further observational studies to identify factors contributing to the findings, (*c*) follow-up studies to determine kindergarten readiness, and (*d*) replication studies to verify these findings.

## References


Bandura, A. (1969). *Principles of behavior modification.* New York: Holt, Rinehart & Winston.

Brownell, C. A. (1990). Peer social skills in toddlers: Competencies and constraints illustrated by same-age and mixed-age interaction. *Child Development,* **61,** 838–848.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin,* **101,** 147–158.

Dunn, J., & Shatz, M. (1989). Becoming a conversationalist despite (or because of) having an older sibling. *Child Development,* **60,** 399–410.

Furman, W., Rahe, D. F., & Hartup, W. W. (1979). Rehabilitation of socially withdrawn preschool children through mixed-age and same-age socialization. *Child Development,* **50,** 915–922.

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

Graziano, W., French, D., Brownell, C. A., & Hartup, W. W. (1976). Peer interaction in same- and mixed-age triads in relation to chronological age and incentive condition. *Child Development,* **47,** 707–714.

Guralnick, M. J., & Paul-Brown, D. (1984). Communicative adjustments during behavior request episodes among children at different developmental levels. *Child Development,* **55,** 911–919.

Guralnick, M. J., & Paul-Brown, D. (1986). Communicative interactions of mildly delayed and normally developing preschool children: Effects of listener's developmental level. *Journal of Speech and Hearing Research,* **29,** 2–10.

Harris, P. L. (1983). Infant cognition. In M. M. Haith & J. J. Campos (Eds.), P. H. Mussen (Series Ed.), *Handbook of child psychology: Vol. 2. Infancy and developmental psychobiology* (pp. 689–782). New York: Wiley.

Hartup, W. W. (1976). Cross-age versus same-age peer interaction: Ethological and cross-cultural perspectives. In V. Allen (Ed.), *Children as tutors: Theory and research on tutoring* (pp. 41–55). New York: Academic Press.

Hocking, R. R. (1985). *The analysis of linear models.* Monterey, CA: Brookes/Cole.

Hollingshead, A. B. (1975). *Four-Factor Index of Social Status.* Unpublished manuscript, Yale University, New Haven, CT.

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

Laird, N. M., & Ware, J. H. (1982). Random effects models for longitudinal data. *Biometrics,* **38,** 963–974.

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

McCall, R. B. (1984). Developmental changes in mental performances: The effect of the birth of a sibling. *Child Development,* **55,** 1317–1321.

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development.* New York: Academic Press.

Newborg, J., Stock, J. R., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). *The Battelle Developmental Inventory.* Allen, TX: DLM/Teaching Resources.

Rodgers, J. L. (1988). Birth order, SAT, and confluence: Spurious correlations and no causality. *American Psychologist,* **43,** 476–477.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological Bulletin,* **92,** 726–748.

Shatz, M., & Gelman, P. (1973). The development of communication skills: Modification in the speech of young children as a function of listeners. *Monographs of the Society for Research in Child Development,* **38**(5, Serial No. 152).

Sparling, J., & Lewis, I. (1979). *Learningames for the first three years.* New York: Berkley.

Sparling, J., & Lewis, I. (1984). *Learningames for threes and fours.* New York: Berkley.

Stewart, P. W. (1987). Line-segment confidence bands for repeated measure. *Biometrics,* **43,** 629–640.

Suomi, S. J., & Harlow, H. F. (1972). Social rehabilitation of isolate-reared monkeys. *Developmental Psychology,* **6,** 487–496.

Zajonc, R. B. (1986). The decline and rise of scholastic aptitude scores: A prediction derived from the confluence model. *American Psychologist,* **41,** 862–867.