Journal of Educational Psychology
2012, Vol. 104, No. 3, 661–672

0022-0663/12/$12.00 DOI: 10.1037/a0028995

Taking It to the Classroom: Number Board Games as a Small Group Learning Activity

Geetha B. Ramani
University of Maryland

Robert S. Siegler
Carnegie Mellon University

Aline Hitti
University of Maryland

We examined whether a theoretically based number board game could be translated into a practical classroom activity that improves Head Start children's numerical knowledge. Playing the number board game as a small group learning activity promoted low-income children's number line estimation, magnitude comparison, numeral identification, and counting. Improvements were also found when a paraprofessional from the children's classroom played the game with the children. Observations of the game-playing sessions revealed that paraprofessionals adapted the feedback they provided to individual children's improving numerical knowledge over the game-playing sessions and that children remained engaged in the board game play after multiple sessions. These findings suggest that the linear number board game can be used effectively in the classroom context.

Keywords: math, preschoolers, small group activity, interventions, board games

Supplemental materials: http://dx.doi.org/10.1037/a0028995.supp

The mathematical knowledge of young children from low-income and minority backgrounds trails behind that of peers from middle-income backgrounds even before they start school (Jordan, Kaplan, Ramineni, & Locuniak, 2008; Lee & Burkam, 2002). Many children from low-income backgrounds do not even know the relative sizes of numbers between 1 and 10 when they enter kindergarten (Starkey, Klein, & Wakeley, 2004). For children without such knowledge, there is no more reason that 4 + 4 should equal 8 rather than 2, 14, or any other number. Lack of understanding of numerical magnitudes reduces arithmetic to rote memorization, which leads to negative consequences both cognitive and motivational: Rote memorization is slow and subject to rapid forgetting, and few people are eager to do it (Reyna, 1996).

These findings point to the educational importance of developing activities that improve knowledge of numerical magnitudes of preschoolers from low-income backgrounds. In the present study, we explored whether a theory-based instructional activity that has been shown to be effective in laboratory settings also could be effective in more naturalistic settings by examining factors that would allow it to be used in a larger classroom setting. The present study focuses on three issues. One issue is whether a number board game that has been shown to increase mathematical knowledge when played with individual children also is effective when played in small groups. A second issue is whether the game is effective when the adult playing the game with the children is a paraprofessional at a Head Start Center who has been given only brief training in how to execute the instructional procedure and who then plays the game with small groups of preschoolers. The third issue is whether social interaction processes that occur during game playing are related to learning outcomes.

Linear Number Board Games

The design of the linear number board game is based on theories and empirical research regarding the mental number line (for reviews, see Ansari, 2008, and Hubbard, Piazza, Pinel, & Dehaene, 2005). The mental number line is a cognitive structure that represents numbers as increasing along a horizontal dimension (left to right in Western cultures). The form of mental number line representations can be assessed by presenting a succession of visible number lines with a fixed number at each end (e.g. 0 and 100) and a third number that varies from trial to trial above the line. The task

This article was published Online First June 18, 2012.

Geetha B. Ramani, Department of Human Development and Quantitative Methodology, University of Maryland; Robert S. Siegler, Department of Psychology, Carnegie Mellon University; Aline Hitti, Department of Human Development and Quantitative Methodology, University of Maryland.

We would like to thank the Institute of Educational Sciences, which supported this research through Grants R305A080013 and R305H050035, and the Teresa Heinz Chair at Carnegie Mellon University. We also would like to thank the administrators, parents, and children of the Head Start programs that participated in this research. Thanks also to Erica Zippert, Josh Gray, Simone Saltzman, Kristin Spurlock, Avita Jones, Shane Schweitzer, and Fei Hou for their help in collecting, coding, and entering the data.

Correspondence concerning this article should be addressed to Geetha B. Ramani, Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD 20742, or Robert S. Siegler, Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213. E-mail: gramani@umd.edu or rs7k@andrew.cmu.edu

is to estimate the location on the line of that third number (Siegler & Opfer, 2003).

Individual difference patterns illustrate the centrality of numerical magnitude understanding for mathematical knowledge. Children whose number line estimates better fit a linear function also perform better on a wide range of other numerical tasks, including magnitude comparison, memory for numbers, arithmetic, and mathematics achievement test scores. These relations are quite strong from kindergarten to fourth grade (Booth & Siegler, 2006; Geary, 2011; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Thompson & Siegler, 2010). The relation is causal as well as correlational; in one study providing randomly selected children with 0–100 number lines with the magnitudes of numbers and their sums improved the children's memory for answers to addition problems using these numbers (Booth & Siegler, 2008).

Based on this analysis, we hypothesized that providing preschool children with experience playing a board game with a linear arrangement of numbers could improve their numerical magnitude understanding. Play and other informal activities are important contexts in which children develop interests in math, support their skill development, and extend their conceptual understanding (Ginsburg, Lee, & Boyd, 2008). Game-based learning can be enjoyable, motivating, and interesting for students, while also providing immediate feedback and information about numbers from playing the game and from the other players (Papastergiou, 2009). For example, experience with board games that have linearly arranged, consecutively numbered, equal-size spaces, such as the first row of the game Chutes and Ladders, can provide kinesthetic, auditory, visuospatial, and temporal cues about numerical magnitudes (Siegler & Booth, 2004). Linear number board games also provide practice at other valuable numerical skills, notably counting and numeral identification (Ramani & Siegler, 2008).

Consistent with this analysis, playing linear numerical board games does improve children's numerical knowledge. In several studies, preschoolers have been randomly assigned to play either a linear numerical board game with squares numbered from 1 to 10 or a game that is identical except that the squares varied in color but not number. After four 15- to 20-min sessions, the children who played the numerical version of the game have shown greater improvements in number line estimation, magnitude comparison, counting, numeral identification, and ability to learn novel arithmetic problems (Ramani & Siegler, 2008; Siegler & Ramani, 2009). These gains have been shown to be stable over a 9-week period. The linear arrangement of the numbers on the board is crucial; playing on an otherwise identical circular board, where the numbers are in a circular arrangement, did not produce comparable improvements (Siegler & Ramani, 2009). Similar improvements have been found in younger children from middle-income backgrounds (Ramani & Siegler, 2011) and have been replicated in Great Britain (Whyte & Bull, 2008).

Translating Board Games to a Small Group Classroom Activity

Interventions that are effective in lab settings frequently are ineffective in classrooms (Newcombe et al., 2009; White, Frishkoff, & Bullock, 2008). Numerous factors could influence the impact of an intervention, such as the teachers' background in the subject. In the present studies, we focused on two likely reasons: (a) the 1:1 experimenter–child interactions that are typical of lab studies are often impossible in classrooms, and (b) the people executing the intervention in classrooms usually have less training in executing the interventions than do research personnel.

The logic of the present study was that rather than moving directly from laboratory demonstrations of the intervention's effectiveness to large-scale application in classrooms, we would first conduct an exploratory study to test the robustness of the intervention in the face of these two common impediments. In Experiment 1, we examined whether the game remained effective when it was played in a small group context. In Experiment 2, we examined the game's effectiveness when executed by paraprofessionals who received brief training prior to the study and then conducted the intervention in a small group context.

A small group activity is an ideal context for playing a board game because teachers can guide or direct children to meet specific educational goals by extending their learning during and after playing the game (Durden & Dangel, 2008; Wasik, 2008). Teachers can also facilitate communication among the children, who can assist each other, build understanding by communicating about the game and number concepts, and provide feedback to one another (Griffin, 2004). Collaboration with peers has been found to enhance learning under many circumstances (Azmitia & Montgomery, 1993; Howe et al., 2007; Rittle-Johnson & Star, 2009).

However, collaboration with peers can also hinder learning in some circumstances (Barron, 2003; Fawcett & Garton, 2005; McCaslin & Good, 1996). Many preschoolers experience difficulty with collaborative learning, because they lack the interpersonal and linguistic skills needed to express themselves in ways that are informative and do not lead to conflict (Azmitia, 1988; Bearison, Magzamen, & Filardo, 1986). In game-playing settings, winning and losing often evoke extreme emotional reactions, and immature and variable emotion regulation skills are another problem for many preschoolers (Ramani, Brownell, & Campbell, 2010). Maintaining sustained attention is a third challenge; preschoolers often lose interest in games when they are not winning or when other children are taking their turns. Therefore, Experiment 1 was designed to determine whether preschoolers would improve their mathematics knowledge from playing the numerical board game in small groups.

Experiment 1

Method

Participants. Participants were 62 preschool children, ranging in age from 3 years 6 months to 5 years 7 months (M = 4 years 7 months, SD = 0.53; 58% female; 60% Hispanic, 32% African American, and 8% other [biracial or unknown]). Children were recruited from six Head Start classrooms in a mid-Atlantic metropolitan area. The families met the income requirement for entrance into the Head Start program as indicated by the U.S. government for the academic year 2008–2009 (income limit of $17,600 for a family of three). All classroom activities were conducted in English. Parental reports indicated that 37% of children were monolingual English speakers; the remaining 63% were bilingual. The primary language was English in 47% of homes and Spanish in 52%. Children who were bilingual English and Spanish speakers were given the choice of the pretest and posttest sessions being in English or Spanish. All children chose English.

Children in each classroom were randomly assigned to the number board game or the color board game condition. The number board game condition included 34 children (M_{age} = 4 years 7 months, SD = 0.51; 62% female; 56% Hispanic, 32% African American, and 12% other). The color board game condition included 28 children (M_{age} = 4 years 5 months, SD = 0.57; 53% female; 63% Hispanic, 31% African American, and 6% other). There were a somewhat more children in the number board condition because in cases where the center had an odd number of children, the extra participant was assigned to the number condition. An additional eight participants did not complete the study; five did not respond to the questions they were asked, and three were absent for extended periods. Their data are not included in the analyses.

Here and in Experiment 2, participants within a condition were randomly divided into groups of three children, unless the number of children in a classroom was not divisible by three, in which case two children played the game together. The number condition included 10 groups of three children and two groups of two; the color condition included eight groups of three children and two groups of two. Children remained in the same group in all sessions. The experimenters were a female of Indian descent (the first author), a female White graduate student, and a female White undergraduate student.

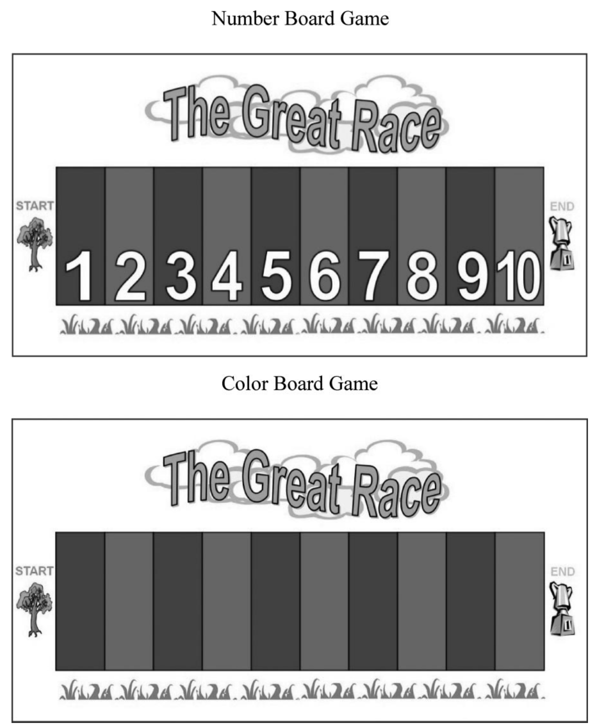

Figure 1. The number board game and color board game. A color version of this figure is available online as supplemental material (http://dx.doi.org/10.1037/a0028995.supp).

Overview of procedure. The study included six 20- to 25-min sessions within a 3- to 4-week period. The sessions were somewhat longer than in previous studies because of the additional children playing the game. Sessions were held either in the children's classroom or in an unoccupied room nearby. During Session 1, children were administered pretest measures of numerical knowledge and the Peabody Picture Vocabulary Test (4th ed.; Dunn & Dunn, 2007), a measure of receptive language skills. The same measures of numerical knowledge were presented as a posttest in Session 6. During Sessions 2 through 5, children played the number or color board game.

Board games. The boards used in the two conditions were the same size (50.60 cm × 27.80 cm) and included 10 spaces of identical size in an alternating pattern of red and blue (Figure 1). The number version had the numbers 1–10 listed consecutively in each space, whereas the color version did not include numbers. The spinner for the number version of the game had a "1" half and a "2" half, whereas the spinner for the color version had a red half and a blue half. Children were given animal character tokens to mark their progress on the board.

At the beginning of Session 2, the experimenter explained the rules of the game named *The Great Race*. Children playing the number board game would spin the spinner and move their token the number of spaces indicated on it. Children playing the color board game were told they would move to the next square that matched the color they spun. Children were also told to say the number (color) that they spun and to say the numbers (colors) in the squares through which they moved their token. For example, if a child playing the number board game was on the square with the number 6 and spun a 2, the child would say "2" and then say "7, 8" as he or she moved the token through those squares. If a child in the color board group was on a blue square and spun a blue, the child would say "red, blue." Each child took a practice turn before the game began. The experimenter reviewed the rules of the game at the beginning of each game-playing session. Each child had an opportunity to go first during each session. Children played on average 12 games over the four sessions, with each game lasting 5–7 min.

The experimenter provided positive feedback to motivate the children, prompted them to say the numbers/colors if they forgot, and told them the numbers/colors to say when necessary. Children were also encouraged to help their friends in the group play the game correctly to promote peer assistance and motivation.

Measures of numerical knowledge. Children were presented four tasks—counting, number line estimation, numerical magnitude comparison, and number identification—in the same order on the pretest and posttest. See Ramani and Siegler (2008) for a full description.

Counting. Children were asked to count from 1 to 10 and were given a score equal to the highest number reached before their first error (10 if they counted perfectly).

Number line estimation. Children were presented 18 sheets of paper, each with a 25-cm line, with "0" at the left end and "10" at the right end. A number from 1–9 inclusive was printed above the center of the line, with each number printed on two of the 18 sheets. On each trial, after asking the child to identify the number at the top of the page, the experimenter asked the child to mark the location of a number on the line.

Numerical magnitude comparison. Children were presented a booklet with 20 pairs of numbers between 1 and 9 inclusive and asked to choose the bigger number. The first two pairs were used for practice with feedback. On the 18 experimental problems, half of the children within each condition were presented a given pair in one order and half in the opposite order.

Numeral identification. Children were presented 10 randomly ordered cards, one at a time, each with a number between 1 and 10 inclusive on it. The experimenter asked the children to identify the number on each card.

Language measure. The Peabody Picture Vocabulary Test (4th ed.; PPVT; Dunn & Dunn, 2007), a widely used measure of children's receptive language skills, was administered during Session 1. If the test could not be completed during Session 1, the remainder of the test was completed during Session 6. On each PPVT item, the experimenter said a target word, and children pointed to the picture that corresponded to it.

Data analysis. The main goal of the analyses was to compare the effects of the number board and color board games on children's numerical knowledge. Our central prediction was that there would be greater gains on the numerical knowledge measures from pretest to posttest for children in the number board game condition.

For the number line estimation task, the accuracy of estimates on each item was measured by percentage of absolute error (PAE):

$$\left|\frac{\text{Estimate} - \text{Presented Number}}{\text{Scale of Estimates}}\right| \times 100$$

For example, if a child marked the location of 5 on a 0–10 number line at the position that corresponded to 9, the PAE would be 40%: $[(9 - 5)/10] \times 100$. The dependent measures on the other tasks were number of correct magnitude comparisons and number of numerals correctly identified. Standardized scores were used for these three variables whose distribution approximated a normal form. Scores on the counting task fit a bimodal distribution; therefore, a binary variable was created where 0 included one or more counting errors and 1 indicated errorless performance.

Because children played the games in small groups, we calculated intraclass correlations to test for nonindependence between members of each group (Grawitch & Munz, 2004; Kenny & la Voie, 1985; Kenny, Mannetti, Pierro, Livi, & Kashy, 2002). A hierarchical linear model (HLM) was used to calculate the intraclass correlations using HLM 7 (Scientific Software International, Lincolnwood, IL), which measures the proportion of total variance in each outcome that is between groups. Fully unconditional two-level models were tested with the child specified at Level 1 and the small groups specified at Level 2 (Raudenbush & Bryk, 2002). Models for each numerical knowledge measure were tested separately.

The intraclass correlation in posttest scores for the groups playing the board game together was significant for only number line PAE. On that measure, 23% of the variance in PAE was accounted for by group membership ($p < .01$). There was no obvious explanation for why scores on this one measure of numerical knowledge were not independent of group. The overall pattern led to the decision to use single-level analyses to examine the effects of condition on unstandardized pretest and posttest performance for number line estimation, magnitude comparison, and numeral identification. To check the appropriateness of this decision, we conducted multilevel modeling of the PAE data; the results were identical to those from the single-level analyses reported in the following section. We analyzed counting separately using the binary variable (0 included one or more errors in counting and 1 indicated errorless performance).

Results and Discussion

Preliminary analyses of covariate. PPVT scores were significantly related to seven of the eight pretest and posttest measures of numerical knowledge and therefore were included as a covariate in the analyses of numerical knowledge. The PPVT scores did not differ between children in the number and color board game conditions (*M*s = 65 and 59), $t(60) = 1.12$, $p = .27$, $d = 0.29$.

Multivariate analyses. We examined multivariate effects of condition and session on number line PAE, magnitude comparison, and numeral identification accuracy. A 2 (condition: number or color board game) × 2 (session: pretest or posttest) repeated-measures multivariate analysis of covariance (MANCOVA) was conducted on the three measures, with raw PPVT scores as a covariate, and session as the repeated measure. Main effects emerged for session, $F(3, 57) = 4.38$, $p < .01$, $\eta_p^2 = .19$; for the Session × Condition interaction, $F(3, 57) = 4.31$, $p < .01$, $\eta_p^2 = .19$; and for the covariate, PPVT scores, $F(3, 57) = 14.48$, $p < .001$, $\eta_p^2 = .43$.

To better understand the results, we conducted a repeated-measures analysis of covariance (ANCOVA) for each task. Means reported have been adjusted for the covariate, so that they represent projections of what scores would have been if all children performed equally on the PPVT. Given that the main prediction of the study is that there would be greater gains in numerical knowledge from pretest to the posttest for children who played the number board game, we conducted planned comparisons examining gains in the outcome measures in both conditions. Paired samples *t* test comparisons were calculated for the three measures in the MANCOVA using the means and standard errors adjusted for the covariates in the ANCOVA models. A Bonferroni correction for the four numerical knowledge measures indicates that the *p* values for the planned comparisons should be evaluated at an adjusted critical value of $p < .0125$.

Number line estimation. Individual children's number line PAEs varied by condition, $F(1, 59) = 4.52$, $p < .01$, $\eta_p^2 = .12$, and for the Session × Condition interaction, $F(1, 59) = 8.00$, $p < .01$, $\eta_p^2 = .12$. The interaction resulted from the number game producing greater improvements than the color game. After PPVT scores were controlled, PAE decreased from pretest to posttest among children who played the number board game from 32% to 25%, $t(33) = 3.94$, $p < .001$, $d = 0.69$, but not among children who played the color board game (from 33% to 34%).

Numerical magnitude comparison. Number of correct magnitude comparisons varied with session, $F(1, 59) = 5.98$, $p < .05$, $\eta_p^2 = .09$, and with PPVT performance, $F(1, 59) = 26.49$, $p < .001$, $\eta_p^2 = .31$. Across conditions, magnitude comparison accuracy improved, from 65% correct on the pretest to 70% correct on the posttest, $t(60) = 2.65$, $p < .05$, $d = 0.32$. Planned comparisons of within-group differences revealed that among children in the number board game, percentage of correct magnitude comparisons improved from pretest to posttest, from 65% to 73% correct, $t(33) = 2.85$, $p < .01$, $d = 0.46$. In contrast, there was no significant improvement among children in the color board game (from 66% to 68% correct), $t(27) = 1.18$, $p = .54$, $d = 0.17$.

Numeral identification. Number of correctly identified numerals varied by session, $F(1, 59) = 6.38$, $p < .05$, $\eta_p^2 = .10$, by the Session × Condition interaction, $F(1, 59) = 4.73$, $p < .05$, $\eta_p^2 = .07$, and by PPVT performance, $F(1, 59) = 27.18$, $p < .00$, $\eta_p^2 = .32$. The Session × Condition interaction again resulted from greater gains by children in the number board condition. Mean correct number identifications of children who played the number board game improved from 6.3 to 7.3, $t(33) = 4.47$, $p < .001$, $d = 0.34$, but no significant changes were present among children who played the color board game (from 6.5 correct on the pretest to 6.7 correct on the posttest), $t(27) = 1.15$, $p = .75$, $d = 0.09$.

Counting. To analyze counting, we conducted a logistic regression that included pretest counting scores, PPVT performance, and condition to predict posttest counting performance (did not count correctly vs. counted correctly). Overall, the model significantly predicted the odds of correctly counting to 10 at posttest, $\chi^2(3) = 33.06$, $p < .001$. Wald criterion showed both counting at pretest ($p = .015$) and PPVT performance ($p = .03$) predicted counting performance at posttest. Counting correctly at pretest increased the odds of counting correctly at posttest by a factor of 27.22, 95% confidence interval (CI) [1.90, 389.361]. An increase of 1 point in PPVT scores was associated with an increase in counting correctly at posttest by a factor of 1.13, 95% CI [1.011, 1.26]. However, the type of game children played did not predict posttest counting performance.

For the planned comparisons, we conducted a McNemar's test to examine whether the number of children who counted correctly at pretest changed at posttest from playing the number and color board games. In the number board game condition, there was a decrease in the number of children who made counting errors from pretest to posttest (27% vs. 6%), $\chi^2(1, N = 34) = 5.90$, $p = .016$, which is marginally significant at the adjusted critical value. In the color board game, the decrease in the number of children who made counting errors from pretest to posttest was not significant (32% vs. 25%, $p = .63$).

Thus, children who played the number board game in small groups improved on all four measures of numerical knowledge. Overall, the findings indicate that children learn from playing the game in small groups, just as they do from playing the game one-on-one with an adult.

Experiment 2

The effectiveness of interventions often falls off when the interventions are conducted by school rather than research personnel. For example, when graduate students read stories one-on-one to children from low-income backgrounds and asked the children questions about the stories' content while doing so, the children's verbal skills increased considerably (Valdez-Menchaca & Whitehurst, 1992), but when day care staff members were trained to read the same stories to their children, the intervention was much less effective (Whitehurst et al., 1994). This suggests scaling up any instruction for use in preschools will require that the game produce gains when executed by teachers and paraprofessionals in classrooms.

Therefore, in Experiment 2, we tested whether the board game would produce positive effects on students' numerical knowledge when the adult playing the game with the children was a paraprofessional who had received only brief training. Previous studies of the number board game have all used extensively trained experimental personnel. This improves experimental control of the way in which the intervention is executed, but would be impractical in real-world settings. Therefore, we tested whether paraprofessionals working at Head Start could, given brief training, play the board game with small groups of children in the intended way, and whether the experience would increase the children's numerical knowledge.

A second goal was to examine the children's and paraprofessionals' behavior during the group board game play and to relate this interaction to the children's learning. The effectiveness of small group interventions depends on what adults do but also on interactions among children and between children and adults (van de Pol, Volman, & Beishuizen, 2010). Learning tends to be greater when adults provide feedback that is contingent on students' level of understanding (Chiu, 2004; Meloth & Deering, 1999), when adults gradually transfer responsibility for problem solving to children (Maloch, 2002), and when children become engaged in the activity and engage other children in it (Azmitia, 1988; Wasik, 2008).

In addition to examining changes in children's numerical knowledge, we also examined variations in children's engagement as a function of the game they were playing and the session. To measure engagement, we examined children's time spent focusing on the board game and their comments during the session. We predicted that children's focus and comments would be equal for both conditions during the first session playing the game, because both activities were likely to be novel, but that engagement in the number game would be higher than in the color game during the final session, due to the number game seeming likely to be more challenging.

We also examined the paraprofessionals' behavior while playing the game. As with the experimental personnel who conducted previous experiments, the paraprofessionals were trained to respond to children's errors by providing a series of prompts and instructive statements, with each statement being more directive than its predecessor. If the child continued to err, the paraprofessionals were to model how to move correctly and have the child repeat the actions. We hypothesized that the paraprofessionals would execute both board games as intended and that the assistance they provided would vary with the board game they were playing, the session, and the children's numerical knowledge. We further expected that children playing the number board game would receive more instruction, because that game was more challenging. In addition, prompts were expected to decrease from the first game-playing session to the last session, as children became better able to play the game, and more prompts were expected to be given to children with less advanced numerical knowledge.

Method

Participants. Participants were 105 preschool children, ranging in age from 3 years 3 months to 5 years 8 months ($M_{age} = 4$ years 5 months, $SD = 0.59$; 51% female; 52% Hispanic, 42% African-American, and 6% other [biracial or unknown]). The children were recruited from seven Head Start classrooms in a metropolitan area in the mid-Atlantic region. The families met the income requirement for entrance into the Head Start program as indicated by the U.S. government for the 2009–2010 academic year (income limit of $18,310 for a family of three). All lessons in the classrooms were conducted in English. Among the participants, 40% were monolingual English speakers, and 60% were bilingual. The primary language spoken at home by participants was 30% Spanish, 68% English, and 2% other languages (e.g., Thai).

Children in each classroom were randomly assigned to the number board game condition or the color board game condition. The number board condition included 54 children (M_{age} = 4 years 4 months, $SD = 0.60$; 52% female; 50% Hispanic, 44% African American, and 6% other). The color board condition included 51 children (M_{age} = 4 years 5 months, $SD = 0.58$; 51% female; 55% Hispanic, 39% African American, and 6% other). One additional participant was recruited but did not complete the study because of an experimenter error.

The number condition included 19 groups (16 groups of three and three groups of two children), and the color condition included 18 groups (15 groups of three and three groups of two children). The experimenters were three White females (one research assistant and two undergraduate students).

Playing the game with the children were seven female paraprofessionals from their classrooms (one Hispanic and six African American). The paraprofessionals had worked in early childhood classrooms for an average of 8.1 years and had worked in their current classroom for an average of 5.8 years. Each was paid $200 for participating in the study.

Paraprofessional training. Paraprofessionals met individually with the first author or a trained research assistant for approximately 1 hr prior to the start of the study. They were presented the board game materials, an opportunity to practice with them, and a short manual that included the rules for the games, scripts for how to explain them, and standard prompts for correcting errors, which were all identical to those used in Experiment 1(manual available at http://www.education.umd.edu/EDHD/faculty2/Ramani/index.html). The paraprofessionals also watched a demonstration video of a group of children playing both games.

The first author or a trained research assistant watched the paraprofessionals' first session and the beginning of their second session with the children to be sure they executed the procedure correctly. The research assistant watched an additional 2 days of each paraprofessional's game-playing sessions after they had played the board game with children for about a week. On very rare occasions, the first author or the research assistant provided feedback or answered questions. All game-playing sessions were videotaped.

Measures. The measures were identical to those used in Experiment 1. Again, all children chose to play the board game in English.

Coding. Children's board game play during Sessions 2 and 5 (i.e., the first and last game playing session) was coded from videos with computer-based observation software (Observer XT 9.0 release; Noldus Information Technology, Wageningen, the Netherlands). Only task-related behavior and communication were coded. Interobserver reliability was established among three independent observers. Two of the observers established reliability with a set of tapes coded by a master coder. Each observer independently coded 10 sessions, roughly equally distributed over sessions, condition, and schools. We calculated interobserver reliability separately using Cohen's kappa for each coding scheme, except for time focused on the game, for which interobserver correlations were calculated. All discrepancies between coders were resolved by discussion. Kappas for the codes are reported in the following sections. Kappas between .6 and .8 indicate substantial reliability (Landis & Koch, 1977).

Time focused on game. Percentage of time during which children attended to the board game was coded for each session. This includes time children were taking their turns, talking to their peers and the paraprofessional, listening to the paraprofessional, and watching the other children take their turns. The interobserver correlations between the master coder and two other coders for total time engaged was $r = .95$, $p < .001$, and $r = .92$, $p < .001$.

Children's comments. Children's game-related statements were coded. Examples of comments included statements related to numbers or colors (e.g., "I have two more to go," and "I am on the blue space"), general game-related comments (e.g., "You are ahead of me," and "I am going to win"), and peer help or motivation ("It's your turn"). Kappas between the master coder and each independent coder were .67 and .66.

Children's errors. Children's turns were coded into whether they included an error. Three types of errors were coded: errors associated with identifying the number or color on the spinner, moving a game token (e.g., child moved two spaces and should have moved only one space), or identifying a number or color while moving a token (e.g., child incorrectly identified a number on a square). For the three errors, kappas between the master coder and each independent coder were .75 and .70. Errors across the three types of errors were aggregated for the analyses.

Assistance. Codes for the type and amount of assistance the children received while playing the game were adapted from Bjorklund, Hubertz, and Reubens (2004) and Wood and Middleton (1975). Table 1 presents the definitions of the assistance behaviors. Scores ranged from Level 0, indicating no assistance from the paraprofessional, to Level 4, indicating that the paraprofessional modeled how to move the token or completed the turn for the child. The highest level of assistance observed during a turn was the score for that turn. For example, if children received a general prompt, such as a request to move their token (Level 1) and the paraprofessional also modeled how to move the token (Level 4), the assistance on that turn was coded as Level 4. Assistance was scored regardless of whether it occurred prior to or following a child error. Because Level 2 occurred relatively infrequently, it was collapsed with Level 1. Similarly, Level 3 was collapsed in the analyses with Level 4. Kappas for the collapsed categories between the master coder and each independent coder were .73 and .72. The main dependent measure was percentage of each child's turns on which the child received no help (Level 0), Level 1 or 2 help, or Level 3 or 4 help.

Data analysis. As in Experiment 1, because children played the board games in small groups, we calculated intraclass correlations to test for nonindependence on posttest scores. In addition, given that groups were nested within the paraprofessional who guided the game playing, we also calculated intraclass correlations to test for nonindependence at that level. As in Experiment 1, fully unconditional HLM models were used to calculate the intraclass correlations. A three-level model was conducted with child specified at Level 1, small group at Level 2, and paraprofessional at Level 3. Standardized posttest scores for number line estimation, magnitude comparison, and numeral identification were used as

Table 1
Levels of Paraprofessional Assistance During Game Playing

Behaviors	Definition
Level 0: No help	Child receives no assistance from paraprofessional and completes the turn independently.
Level 1: Prompt	Child receives a general prompt as a reminder of how to play the game (e.g. "Move your character").
Level 2: Guidance	Child receives verbal guidance on completing the turn, usually reiterating how to move given what he or she spins (e.g., "Move one space" or " Move to a blue space").
Level 3: Instruction	Child receives verbal and/or gestural instruction on moving the token or naming the spaces, usually receiving a portion of the information needed to complete the turn (e.g., paraprofessional says a number for the space while the child is moving the token).
Level 4: Model	Child is told how to complete the turn by the paraprofessional (e.g. paraprofessional moves child's token and says "Say '3' as you move to this space." Paraprofessional then returns token to original space for child to complete the turn).

the outcome measures. Counting, in which the data approximated a bimodal distribution, was scored 0 if there were errors and 1 for errorless performance.

The intraclass correlations were not significant at the small group level. At the paraprofessional level, intraclass correlations of 10% and above were found for counting (10.3%; $p < .05$), suggesting that part of the variance in these scores could be accounted for by differences among paraprofessionals. A significant intraclass correlation for numeral identification was also found, with between-paraprofessional differences accounting for 7% ($p < .05$) of the variance. However, intraclass correlations that account for less than 10% of the variance are considered trivial (Lee, 2000). Given that the intraclass correlations at the paraprofessional level were significant and nontrivial for only one of the four measures, that there was no theoretical explanation for why scores on only this one measure of numerical knowledge would be affected by which paraprofessional was playing the game, and that multilevel modeling analyses yielded the same results as single-level analyses, we report only single-level analyses of the numerical knowledge measures.

Results and Discussion

Multivariate analyses. A 2 (condition: number or color board game) × 2 (session: pretest or posttest) repeated-measures MANCOVA was conducted on number line estimation, magnitude comparison, and numeral identification with raw PPVT scores as a covariate. Main effects emerged for the Session × Condition interaction, $F(3, 100) = 4.22$, $p < .01$, $\eta_p^2 = .11$, and for PPVT scores, $F(3, 100) = 21.36$, $p < .001$, $\eta_p^2 = .39$. As in Experiment 1, a repeated-measures ANCOVA was conducted for each task. Means reported have been adjusted for the covariate. Planned comparisons of pretest–posttest differences on the outcome measure for the two conditions were conducted. We calculated paired samples *t* test comparisons for the three measures in the MANCOVA using the means and standard errors adjusted for the covariates in the ANCOVA models. A Bonferroni correction for the four outcome measures indicates that the *p* values for the planned comparisons should be evaluated at $p < .0125$.

Number line estimation. PAEs of children's number line estimates varied with the Session × Condition interaction, $F(1, 102) = 8.98$, $p < .01$, $\eta_p^2 = .08$, and with PPVT performance, $F(1, 102) = 8.30$, $p < .01$, $\eta_p^2 = .08$. The Session × Condition interaction resulted from the number board game producing greater pretest–posttest improvements than the color condition. After PPVT scores were controlled, PAEs of children who played the number board game decreased from 34% on the pretest to 29% on the posttest, $t(53) = 4.00$, $p < .001$, $d = 0.48$. In contrast, PAEs of children who played the color board game did not change (from 30% to 31%).

Numerical magnitude comparison. Number of correct magnitude comparisons varied with PPVT performance, $F(1, 102) = 30.68$, $p < .001$, $\eta_p^2 = .23$, and with the PPVT × Session interaction, $F(1, 102) = 4.38$, $p < .05$, $\eta_p^2 = .04$. Planned comparisons of within-group differences revealed that magnitude comparison accuracy of children who played the number board game improved from 63% correct on the pretest to 68% correct on the posttest, $t(53) = 2.41$, $p = .019$, $d = 0.28$, which is marginally significant at the adjusted critical value. There were no significant improvements in accuracy among children who played the color board game (from 64% to 67% correct), $t(50) = 1.39$, $p = .42$, $d = 0.16$.

To analyze the PPVT × Session interaction, we performed a median split on the scores. Magnitude comparison accuracy of children with above-median PPVT scores improved from 68% to 74% correct from pretest to posttest, $t(51) = 3.24$, $p < .01$, $d = 0.28$, but accuracy of children with below-median PPVT scores was unchanged (from 60% to 61% correct).

Numeral identification. Number of correctly identified numerals varied with the Session × Condition interaction, $F(1, 102) = 5.21$, $p < .05$, $\eta_p^2 = .05$, and with PPVT performance, $F(1, 102) = 39.47$, $p < .001$, $\eta_p^2 = .28$. After the effect of language skills was controlled, the Session × Condition interaction resulted from greater gains by children in the number board game condition. Children who played the number board game improved from a mean of 5.9 to 7.0 correct number identifications from pretest to posttest, $t(53) = 4.26$, $p < .01$, $d = 0.35$. There were no significant changes among children who played the color board game (from 6.1 to 6.3 correct).

Counting. We conducted a logistic regression that included pretest counting scores, PPVT performance, and condition to predict posttest counting performance. The overall model significantly predicted counting performance, $\chi^2(3) = 17.93$, $p < .001$. Significant effects were found for pretest counting performance ($p = .004$) based on the Wald criterion, but no effects were found for PPVT or condition. Counting without any errors at pretest increased the odds of counting correctly at posttest by a factor 8.23, 95% CI [1.93, 35.07].

A McNemar's test showed there was a decrease from pretest to posttest in the percentage of children making counting errors among those who played the number board game, from 35% to 13%, $\chi^2(1, N = 54)$, 10.97, $p < .01$, and among those who played the color board game, from 35% to 16%, $\chi^2(1, N = 51)$, 6.55, $p = .013$, which is marginally significant at the adjusted critical value.

Engagement in board game play. We hypothesized that engagement would be better maintained among children who played the number board game from the first to the last game-playing session. However, the percentage of time children were engaged in the first and last board game-playing sessions was high for both conditions (number board condition: 95% and 92% of the session; color board game: 96% and 91% of the session; see Table 2).

Children's comments during the game provided another measure of engagement. For children's comments, a 2 (condition: number or color board game) × 2 (session: first or last board game session) repeated-measures analysis of variance (ANOVA) was conducted. Main effects emerged for session, $F(1, 103) = 7.44$, $p < .01$, $\eta_p^2 = .07$, and was marginally significant for condition, $F(1, 103) = 3.42$, $p = .07$, $\eta_p^2 = .03$. Across both conditions, children's comments increased from the first to the last game playing session, $M = 6.04$ versus 8.46, $t(104) = 2.75$, $p = .01$, $d = 0.54$. Among children in the number board game, mean number of comments increased from the first to the last game session, $M = 7.09$ versus 9.94, $t(53) = 2.01$, $p = .05$, $d = 0.30$. Among children in the color board game, number of comments also tended to increase, but the increase was marginally significant, $M = 4.92$ versus 6.88, $t(50) = 1.93$, $p = .06$, $d = 0.27$.

Errors during game playing. We conducted a 2 (condition: number or color board game) × 2 (session: first or last board game session) repeated-measures ANOVA was conducted on children's errors. Main effects emerged for session, $F(1, 103) = 9.91$, $p < .01$, $\eta_p^2 = .09$, and for condition, $F(1, 103) = 14.16$, $p < .001$, $\eta_p^2 = .12$. Children's errors were more frequent when playing the number board game than when playing the color board game during both the first game playing session, 9.02 versus 4.43 errors, $t(103) = 3.97$, $p < .001$, $d = 0.78$, and the last one, 7.22 versus 3.63 errors, $t(103) = 3.07$, $p < .01$, $d = 0.60$. However, number of errors decreased from the first to the last session for the number board condition, from 9.02 to 7.22, $t(53) = 2.65$, $p < .05$, $d = 0.26$, but not for the color board condition, from 4.43 to 3.63, $t(50) = 1.72$, $p = .09$, $d = 0.17$.

Assistance during game play. To examine the assistance the paraprofessionals provided during the board game play, we conducted a 2 (condition: number or color board game) × 2 (session: first or last board game session) repeated-measures MANOVA on the three composite measures of paraprofessional assistance: no help, prompt and guidance (Level 1 or 2 help), and instruction and modeling (Level 3 or 4 help). Main effects emerged for session, $F(2, 102) = 68.53$, $p < .001$, $\eta_p^2 = .57$, and for condition, $F(2, 102) = 10.57$, $p < .001$, $\eta_p^2 = .17$. To better understand the results, we conducted a repeated-measures ANOVA for each measure.

The percentage of turns during which the paraprofessional did not provide any help varied by session, $F(1, 103) = 116.58$, $p < .001$, $\eta_p^2 = .53$, and by condition, $F(1, 103) = 19.17$, $p < .01$, $\eta_p^2 = .16$. If the paraprofessionals appropriately responded to the children's game playing, they would provide more assistance during the more difficult number board game than during the easier color board game. As predicted, the percentage of turns the children completed without assistance was lower for children playing the number board game than the color board game during both the first session, (7% vs. 20%), $t(103) = 3.67$, $p < .001$, $d = 0.72$, and during the last session (31% vs. 53%), $t(103) = 3.85$, $p < .001$, $d = 0.76$. Also, as predicted, the paraprofessionals provided less assistance as the children gained experience. Planned comparisons of within-group differences revealed the percentage of turns completed without assistance increased over sessions for children in the number condition, (7% vs. 31%), $t(53) = 6.28$, $p < .001$, $d = 0.98$, and the color condition, (20% vs. 53%), $t(50) = 9.04$, $p < .001$, $d = 1.32$.

For the specific types of assistance, the percentage of turns during which the paraprofessional provided instruction and modeling varied by session, $F(1, 103) = 98.05$, $p < .001$, $\eta_p^2 = .49$, and by condition, $F(1, 103) = 20.35$, $p < .001$, $\eta_p^2 = .17$. As predicted, the paraprofessionals provided the most intensive level of assistance on a higher percentage of turns to children playing the number board game than those playing the color board game during both the first session, 72% versus 54% of turns, $t(103) = 3.38$, $p < .01$, $d = 0.66$, and the last session, 50% versus 25%, $t(103) = 4.49$, $p < .001$, $d = 0.88$. However, planned comparisons of within-group differences revealed the percentage of turns during

Table 2
Children's and Paraprofessionals' Behavior During the First and Last Board Game Session

	Number board game		Color board game	
Variable	First session *M*	Last session *M*	First session *M*	Last session *M*
Percentage of time engaged	95	92	96	91
Total game-playing errors	9.02	7.22	4.43	3.63
Child comments	7.09	9.94	4.92	6.88
Paraprofessional assistance level (%)				
0 (No help)	7	20	31	53
1 (Prompt)	14	13	9	9
2 (Guidance)	7	13	10	13
3 (Instruction)	7	7	8	5
4 (Model)	65	42	47	20

which the paraprofessionals instructed the children and modeled the correct procedures decreased from the first session to the last game-playing session in the number condition (72% vs. 50%), $t(53) = 6.28$, $p < .001$, $d = 0.78$, and in the color condition, (54% vs. 25%), $t(50) = 7.66$, $p < .001$, $d = 1.11$. There were no significant effects for the prompt and guidance feedback the paraprofessionals provided.

To examine whether the paraprofessionals provided differential feedback based on the children's numerical knowledge, we created a measure of children's numerical knowledge—the sum of standardized scores on the four numerical knowledge measures. We then correlated these scores with the assistance children received during the first and last game-playing sessions (Table 3). For the number board group, during both the first and last board game-playing sessions, pretest numerical knowledge was positively correlated with the paraprofessionals providing the less directive type of assistance (prompts and guidance) and negatively correlated with them providing the more directive assistance (instruction and modeling). During the fourth session, numerical knowledge was also positively correlated with children completing more turns without any help. A similar pattern was found in the color board condition.

In sum, playing the number board game in a small group of peers supervised by a paraprofessional from their classroom improved from pretest to posttest children's numerical estimation, magnitude comparison, and numeral identification; in contrast, playing the color board game did not. Observations of board game play also provided information on the kinds of assistance children received while playing the game. The paraprofessionals adjusted their feedback in response to the game the children were playing, the session in which the game was being played, and the children's numerical knowledge.

General Discussion

In this concluding section, we summarize the results of the study in terms of its three main goals. We then consider issues raised by the findings.

Playing Number Board Games in Small Groups of Peers

The present results demonstrate that The Great Race can be successfully translated from a one-on-one activity into a small group learning activity for use in classrooms. When playing the number board game as a group, children improved their number line estimation accuracy, magnitude comparison, numeral identification, and counting. These findings replicated those from Ramani and Siegler (2008), a study in which children played the game one-on-one with trained research personnel. Similar to that previous study, children who played a color version of the game did not show comparable improvements in their numerical knowledge. The one exception was for counting in Experiment 2, which is a finding that differed between the present study and the previous one for unknown reasons.

The small group format seems to be an effective context for young children to learn math through game playing. Playing the board game with a group of peers could have been challenging and ineffective at promoting numerical knowledge, since it requires relatively advanced social and emotional skills. Children had to wait their turn, manage emotions related to winning and losing, and interact with peers, all of which could have distracted them and reduced their learning. However, they remained engaged with the game throughout the four game-playing sessions, and their numerical knowledge benefited. The board game might have been even more enjoyable and exciting to play with other children than one-on-one with an adult. Others have found, for example, when 5-year-old children work on a computer activity together, they show more positive affect and rate their affect higher than children working alone on the task (Perlmutter, Behrend, Kuo, & Muller, 1989).

The small group format may also have been effective because children learned from each other while playing the game. Children may have learned how to play the game, to identify the numbers on the board, and how to count while watching their peers move, as well as on their own turns (Azmitia, 1988; Verba, 1998). Children in the present study were randomly assigned to the small groups but more strategic assignment of children with more and less

Table 3
Relations Between Children's Behavior During the First and Last Board Game Session and Numerical Knowledge at Pretest and Posttest

Variable	Number board game: Pretest number knowledge	Number board game: Posttest number knowledge	Color board game: Pretest number knowledge	Color board game: Posttest number knowledge
First board game-playing session feedback				
No help	.07	.12	$.32^{*}$	$.38^{**}$
Prompt/guidance	$.42^{**}$	$.40^{**}$	$.34^{*}$	.25
Instruction/model	$-.37^{**}$	$-.39^{**}$	$-.43^{**}$	$-.42^{**}$
Last board game-playing session feedback				
No help	$.46^{***}$	$.38^{**}$	$.46^{**}$	$.49^{**}$
Prompt/guidance	$.32^{*}$	$.29^{*}$	$-.42^{**}$	$-.51^{***}$
Instruction/model	$-.62^{***}$	$-.53^{***}$	$-.29^{*}$	$-.26$

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

advanced numerical knowledge may increase learning from playing the game with classmates even further.

In general, small group activities seem to be an effective way to promote early math skills. Effective preschool math curricula that include both direct instruction and games often include small group learning activities. These curricula include Number Worlds (Griffin, 2007), Building Blocks (Clements & Sarama, 2007), and Big Math for Little Kids (Ginsburg, Greenes, & Balfanz, 2003).

Classroom Implementation and Early Childhood Education

The second goal of the study was to examine whether the number board game improves children's numerical knowledge when a paraprofessional from the classroom, rather than a trained researcher, leads the group. Our results suggest that the number board game can be easily implemented by paraprofessionals in classrooms and that children's numerical knowledge can benefit under these conditions. When children played the number board game with the paraprofessionals, their numerical knowledge improved on all four numerical knowledge tasks.

This result was especially promising because the paraprofessional training was only 1 hr. In this brief training, paraprofessionals learned how to deliver the instructions, have the children follow the rules of the game, and provide specified types of assistance in such a way that children's numerical knowledge increased. The efficiency of the training might be improved further by having paraprofessionals who had been trained subsequently training other paraprofessionals, either individually or in groups. It is also possible that simply providing teachers with the instruction manual would be sufficient to have them implement the game.

More broadly, providing teachers with training and access to number-related activities would likely strengthen math education in early childhood classrooms. Although professional development training can be costly and time-consuming, it can provide teachers the skills to implement math curricula, reduce anxiety related to teaching math, and promote understanding of math standards for young children (Bowman, Donovan, & Burns, 2001; Sarama & DiBiase, 2004; Starkey et al., 2004). Math-related activities also would offer teachers more opportunities to talk about numbers and engage in math-related play with students (Ginsburg et al., 2008; Klibanoff, Levine, Huttenlocher, Vasilyeva, & Hedges, 2006).

Small Group Social Dynamics

The third goal of the study was to examine the small group social interactions, both the feedback and assistance that children receive from the paraprofessionals and children's behavior during the board game play. The observations suggested that during the game, the paraprofessionals adapted their support to the children's changing knowledge. This type of individualized support seems especially important for promoting learning in small group activities (Durden & Dangel, 2008). The feedback provided by the paraprofessionals varied with the game the children were playing, the session in which the game was being played, and (most important) the children's numerical knowledge. Children with poorer numerical knowledge received more intense instruction, such as modeling of correct counting, and children with better numerical knowledge received less intense guidance, such as general prompts, or no help at all.

Although the paraprofessionals adapted their guidance to the children's changing knowledge, they also provided help even when children did not make an error. It is likely that they did this in anticipation of children making an error or to prevent an error from occurring. Since the paraprofessionals worked with the children on a daily basis, they might be able to give anticipatory support more effectively than an experimenter could. We do not know, however, whether this type of support helped or hurt children's learning. It would be interesting for future research to examine the effects of such anticipatory support.

The results also showed that even though children played the board games over multiple sessions, they remained focused and attentive. The fact that the children's errors while playing the number board game decreased over the sessions, and the fact that their comments increased over the course of the games, also provides evidence that children remained engaged in the game. Thus, the number board game can be used as a learning tool over several sessions.

Limitations

Several limitations of the current study should be noted. First, ceiling effects on the counting measure limited our ability to analyze improvements in children's counting skills. In future research, additional or expanded counting measures could be included to allow investigators to assess how playing the number board could improve counting skills. Second, although experimenters remained uninvolved during the board game play, paraprofessionals were aware that they were being videotaped, which could have influenced the feedback they provided. Future research should examine the paraprofessionals' implementation and effectiveness without the presence of an experimenter. Finally, despite the promising results, children's board game play was scaffolded by the paraprofessionals, limiting our ability to generalize how children would benefit without such guidance. Future research should investigate ways to improve the impact of playing the number board game by varying the amount and type of assistance provided by the adult.

Conclusions

In sum, the current exploratory studies lay the necessary groundwork for examining how the number board game could be scaled-up for use more widely in Head Start classrooms. The results showed that a number board game when played in small groups with paraprofessionals from the classrooms can promote the numerical knowledge of young children from low-income backgrounds. Overall, the positive benefits and greater efficiency of playing the game with peers suggest that small group activities can be useful for scaling up this research-based intervention for widespread use in classrooms.

References

Ansari, D. (2008). Effect of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience, 9,* 278–291. doi:10.1038/nrn2334

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59,* 87–96. doi:10.2307/1130391

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development, 2,* 202–221. doi:10.1111/j.1467-9507.1993.tb00014.x

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12,* 307–359. doi:10.1207/S15327809JLS1203_1

Bearison, D. J., Magzamen, S., & Filardo, E. K. (1986). Socio-cognitive conflict and cognitive growth in young children. *Merrill-Palmer Quarterly, 32,* 51–72.

Bjorklund, D. F., Hubertz, M. J., & Reubens, A. C. (2004). Young children's arithmetic strategies in social context: How parents contribute to children's strategy development while playing games. *International Journal of Behavioral Development, 28,* 347–357. doi:10.1080/01650250444000027

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42,* 189–201. doi:10.1037/0012-1649.41.6.189

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79,* 1016–1031. doi:10.1111/j.1467-8624.2008.01173.x

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). *Eager to learn: Educating our preschoolers.* Washington, DC: National Academy Press.

Chiu, M. M. (2004). Adapting teacher interventions to student needs during cooperative learning: How to improve student problem solving and time on-task. *American Educational Research Journal, 41,* 365–399. doi:10.3102/00028312041002365

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks Project. *Journal for Research in Mathematics Education, 38,* 136–163.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). San Antonio, TX: Pearson.

Durden, T., & Dangel, J. (2008). Teacher-involved conversations with young children during small group activity. *Early Years, 28,* 251–266.

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology, 75,* 157–169. doi:10.1348/000709904X23411

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47,* 1539–1552. doi:10.1037/a0025510

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78,* 1343–1359. doi:10.1111/j.1467-8624.2007.01069.x

Ginsburg, H. P., Greenes, C., & Balfanz, R. (2003). *Big math for little kids.* Parsippany, NJ: Dale Seymour.

Ginsburg, H. P., Lee, J. S., & Boyd, J. S. (2008). Mathematics education for young children: What it is and how to promote it. *Social Policy Report, 22,* 1–24.

Grawitch, M. J., & Munz, D. C. (2004). Are your data nonindependent? A practical guide to evaluating nonindependence and within-group agreement. *Understanding Statistics, 3,* 231–257. doi:10.1207/s15328031us0304_2

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly, 19,* 173–180. doi:10.1016/j.ecresq.2004.01.012

Griffin, S. (2007). *Number Worlds: A mathematics intervention program from Grades preK–6.* Columbus, OH: SRA/McGraw-Hill.

Howe, C., Tolmie, A., Thurston, A., Topping, K., Christie, D., Livingston, K., . . . Donaldson, C. (2007). Group work in elementary science: Towards organizational principles for supporting pupil learning. *Learning and Instruction, 17,* 549–563. doi:10.1016/j.learninstruc.2007.09.004

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between numbers and space in parietal cortex. *Nature Reviews Neuroscience, 6,* 435–448. doi:10.1038/nrn1684

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2008). Development of number combination skill in the early school years: When do fingers help? *Developmental Science, 11,* 662–668. doi:10.1111/j.1467-7687.2008.00715.x

Kenny, D. A., & la Voie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology, 48,* 339–348. doi:10.1037/0022-3514.48.2.339

Kenny, D. A., Mannetti, L., Pierro, A., Livi, S., & Kashy, D. A. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology, 83,* 126–137. doi:10.1037/0022-3514.83.1.126

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., & Hedges, L. V. (2006). Preschool children's mathematical knowledge: The effect of teacher "math talk." *Developmental Psychology, 42,* 59–69. doi:10.1037/0012-1649.42.1.59

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159–174. doi:10.2307/2529310

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist, 35,* 125–141. doi:10.1207/S15326985EP3502_6

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

Maloch, B. (2002). Scaffolding student talk: One teacher's role in literature discussion groups. *Reading Research Quarterly, 37,* 94–112. doi:10.1598/RRQ.37.1.4

McCaslin, M., & Good, T. L. (1996). The informal curriculum. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology.* (pp. 622–670). New York, NY: Prentice Hall International.

Meloth, M. S., & Deering, P. D. (1999). The role of the teacher in promoting cognitive processing during collaborative learning. In A. M. O'Donnell & A. King (Eds.), *Cognitive perspectives on peer learning.* (pp. 235–255). Mahwah, NJ: Erlbaum.

Newcombe, N. S., Ambady, N., Eccles, J., Gomez, L., Klahr, D., Linn, M., . . . Mix, K. (2009). Psychology's role in mathematics and science education. *American Psychologist, 64,* 538–550. doi:10.1037/a0014813

Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers & Education, 53,* 603–622. doi:10.1016/j.compedu.2009.04.001

Perlmutter, M., Behrend, S. D., Kuo, F., & Muller, A. (1989). Social influences on children's problem solving. *Developmental Psychology, 25,* 744–754. doi:10.1037/0012-1649.25.5.744

Ramani, G. B., Brownell, C. A., & Campbell, S. B. (2010). Positive and negative peer interaction in 3- and 4-year-olds in relation to regulation and dysregulation. *Journal of Genetic Psychology, 171,* 218–250. doi:10.1080/00221320903300353

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79,* 375–394. doi:10.1111/j.1467-8624.2007.01131.x

Ramani, G. B., & Siegler, R. S. (2011). Reducing the gap in numerical knowledge between low- and middle-income preschoolers. *Journal of Applied Developmental Psychology, 32,* 146–159. doi:10.1016/j.appdev.2011.02.005

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

Reyna, V. F. (1996). Conceptions of memory development, with implications for reasoning and decision making. *Annals of Child Development, 12,* 87–118.

Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101,* 529–544. doi:10.1037/a0014224

Sarama, J., & DiBiase, A. M. (2004). The professional development challenge in preschool mathematics. In D. H. Clements & J. Sarama (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 415–448). Mahwah, NJ: Erlbaum.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75,* 428–444. doi:10.1111/j.1467-8624.2004.00684.x

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237–243. doi:10.1111/1467-9280.02438

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101,* 545–560. doi:10.1037/a0014239

Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly, 19,* 99–120. doi:10.1016/j.ecresq.2004.01.002

Thompson, C. A., & Siegler, R. S. (2010). Linear numerical magnitude representations aid children's memory for numbers. *Psychological Science, 21,* 1274–1281. doi:10.1177/0956797610378309

Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican. *Developmental Psychology, 28,* 1106–1114. doi:10.1037/0012-1649.28.6.1106

van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational Psychology Review, 22,* 271–296. doi:10.1007/s10648-010-9127-6

Verba, M. (1998). Tutoring interactions between young children: How symmetry can modify asymmetrical interactions. *International Journal of Behavioral Development, 22,* 195–216. doi:10.1080/016502598384577

Wasik, B. A. (2008). When fewer is more: Small groups in early childhood classrooms. *Early Childhood Education Journal, 35,* 515–521. doi:10.1007/s10643-008-0245-4

White, G., Frishkoff, G., & Bullock, M. (2008). Bridging the gap between psychological science and educational policy and practice. In S. Thurman, C. A. Fiorello, S. Thurman, & C. A. Fiorello (Eds.), *Applied cognitive research in K–3 classrooms* (pp. 227–263). New York, NY: Routledge/Taylor & Francis Group.

Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30,* 679–689. doi:10.1037/0012-1649.30.5.679

Whyte, J. C., & Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Developmental Psychology, 44,* 588–596. doi:10.1037/0012-1649.44.2.588

Wood, D. J., & Middleton, D. (1975). A study of assisted problem-solving. *British Journal of Psychology, 66,* 181–191. doi:10.1111/j.2044-8295.1975.tb01454.x

Received January 6, 2012
Revision received May 11, 2012
Accepted May 16, 2012 ■