Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics

Citation: Dillon, Moira R. et al. "Cognitive Science in the Field: A Preschool Intervention Durably Enhances Intuitive but Not Formal Mathematics." Science 357, 6346 (July 2017): 47–55 © 2017 The Authors

As Published: http://dx.doi.org/10.1126/science.aal4724

Publisher: American Association for the Advancement of Science (AAAS)

Persistent URL: http://hdl.handle.net/1721.1/114666

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Title: Cognitive Science in the field: A preschool intervention durably enhances intuitive but not formal mathematics

Authors: Moira R. Dillon[1]*, Harini Kannan[2], Joshua T. Dean[3], Elizabeth S. Spelke[1]*, Esther Duflo[2,3]*

Affiliations:

[1] Department of Psychology, Harvard University, Cambridge, MA USA

[2] Abdul Latif Jameel Poverty Action Lab

[3] Department of Economics, Massachusetts Institute of Technology, Cambridge, MA USA

*Correspondence to: mdillon@fas.harvard.edu, spelke@wjh.harvard.edu, or eduflo@mit.edu

Abstract: Many poor children are under-prepared for demanding primary school curricula. Research in cognitive science suggests that a preschool pedagogy in which numerate adults engage children's spontaneous, non-symbolic mathematical concepts, could improve their school achievement. To test this suggestion, we designed and evaluated a game-based preschool curriculum aimed to exercise children's emerging skills in number and geometry. In a randomized field experiment with 1,540 children (4.9-years-old, on average) in 214 Indian preschools, four months of math gameplay yielded marked and enduring improvement on the exercised intuitive abilities, relative to no-treatment and active control conditions. Math-trained children also showed immediate gains on symbolic mathematical skills, but no advantage in subsequent learning of the language and concepts of school mathematics.

One Sentence Summary: Scalable interventions durably improved young children's intuitive mathematics, but did not affect their learning of the symbolic mathematics taught in school.

Main Text:

Enrollment and attendance in primary school in developing countries has dramatically expanded over the last few decades (*1*, *2*), but children's learning outcomes remain poor. In 2014, 87% of Indian children in grade two and 52% of Indian children in grade five could not read a simple passage of text that they should have been able to read by grade two (*3*). Poorly adapted curricula may be partly to blame (*4*-*6*), as such curricula build on the verbal and mathematical skills that preschool children with educated parents gain by interacting with family members who can read, count, and calculate. But first-generation school children may be hampered by a lack of opportunities to engage, as preschoolers, with literate and numerate adults during activities that exercise basic verbal and numerical abilities (*7*-*9*).

This problem can be addressed either by dampening the level of instruction in primary school (*10*) or by bolstering children's experiences during the preschool years. Some early childhood interventions have targeted parents, training them to interact with or support their children (*11-16*). Alternatively, preschools for poor children, led by educated adults who play games that exercise their cognitive abilities, may better prepare children for school.

This idea is intuitively appealing, receiving considerable support from both academics and policy makers (*17*). Indeed, there is evidence that preschool education influences later life outcomes. In the U.S., a number of observational studies have found substantial short- and long-term impacts of the flagship preschool program, Head Start (*18*-*20*). However, a recent large-scale randomized study found only small and short-term effects of Head Start, perhaps because Head Start may not be much better than the alternative preschool choices available to poor U.S. children (*21*). In developing countries, several of the studies reviewed in Engle et al. (*22*) also find positive effects of preschool access on child development. For example, in one recent randomized trial in Mozambique, access to preschool increased children's school enrollment, fine motor skills, and problem solving, although not their later language development (*23*).

Many scholars have emphasized the importance of preschool quality (*24*), but little work has revealed what constitutes a quality program. In the U.S., even carefully designed preschool mathematics curricula based on cognitive science (such as the Building Blocks program) have produced only small effects for only a portion of the students at a portion of the measured time points (*25*). Rigorous randomized controlled trials in resource-poor settings have found no effect on children's learning of training programs for preschool teachers in Chile (*26*) or Malawi (*16*).

These results underscore how little we know about how to train teachers to prepare children for primary school: The teacher training or the curriculum they implemented might not have been intense enough, or the teaching practices and curricula themselves might not have been effective. Moreover, if such practices and curricula are ineffective, we do not know enough about what was trained to draw more general conclusions from the findings. Is the basic intuition — that exercising children's spontaneously developing cognitive abilities in preschool leads to greater school achievement — wrong? Or did the specific, chosen curricula fail to engender, in poor children with uneducated parents, the skills that develop spontaneously in preschool children in wealthy families and communities?

To address this question, we designed a game-based mathematics curriculum for poor children in the slums of Delhi, India. The curriculum is based on decades of research in cognitive science on the spontaneous development of children's numerical and spatial reasoning. We then tested the effectiveness of this curriculum in a large-scale field experiment. We demonstrate that our intervention effectively and durably improved children's spontaneously developing numerical and spatial abilities, and we were therefore able to test whether this improvement led in turn to an increase in children's learning of the symbolic mathematics taught in school. Our study is thus the first to field-test a central conjecture of contemporary basic research in psychology and cognitive science, which has, formally or informally, motivated the development of most modern preschool curricula: That children's learning of the symbolic mathematics taught in school would be facilitated by adult-led activities that exercise their intuitive cognitive abilities during the preschool years. We focus, in particular, on numerical and spatial abilities that emerge in infancy and function throughout life among people from diverse cultures (*27-37*).

Despite the importance of this conjecture, most of the evidence supporting it comes from the laboratory rather than the field. A small number of highly controlled training experiments comprise this literature. Like adults (*38*), elementary school children who are trained to add or compare arrays of dots on the basis of number show enhanced performance on the kinds of symbolic arithmetic problems presented in school, both in lab-based studies in the U.S. (*39*) and in a school-based study in Pakistan (*40*). However, these studies, and other highly controlled training experiments focusing on spatial skills, measure symbolic mathematical gains over very short time periods (from immediately after training to up to a month after training, e.g., *39, 41, 42*), providing no insight on whether any of these gains would persist or enhance learning of new mathematical concepts. A large body of longitudinal research has probed relations between early- and later-developing mathematical abilities across more diverse populations and at longer time scales (*43-46*), but these studies could be indicative of a natural correlation among different abilities, rather than of a causal relation (*47*). Individuals who are mathematically talented, or who receive rich exposure to mathematical material in their homes, may perform well on both intuitive, non-symbolic mathematical tasks as well as learned, symbolic mathematical tasks, even if the abilities underlying these skills and tasks are not causally related.

This study investigates the basic cognitive mechanisms promoting children's learning of mathematics by designing a curriculum that provides children who have minimal access to books, board games, or literate and numerate adults an opportunity to exercise these informal numerical and spatial skills, and by testing the curriculum’s efficacy after the first year of formal schooling. Moreover, it demonstrates an approach to develop and test a cheaply implementable intervention to improve children’s school readiness in resource-poor contexts, which if effective, could be scaled up across preschools.

Intervention and Experimental Design

The intervention took place over a 4-month period, and involved 214 preschools in Delhi, India. These preschools were run by our partner organization, Pratham, a large non-profit focused on improving and evaluating education throughout India. In poor neighborhoods of urban areas, many children now attend such preschools. They are not systematically run by the government, but are often private or, like Pratham’s, run by a non-profit organization.

Our math games curriculum was designed to be scalable and easy to implement in such a context: We used inexpensive, locally printed materials, and locally hired adults administered the games after 2-4 days of training. Children in all of the preschools were of mixed ages, but only the children who were expected by their teachers to begin primary school after the completion of the intervention were assessed and treated. The final sample included 1,540 children (mean age = 4.9 years; range = 2 years to 12 years). Almost all of the children were between 3 and 7 years of age (97.1%) and most were 4 to 5.5 years old (83.8%).

Each preschool was randomly assigned to one of the three treatment conditions. In the *math games* condition (70 schools), children played five games (**Figs. 1-2;** see SM for details) that build on intuitive numerical and geometric abilities that emerge spontaneously in the first 3 years, that are associated with achievement in school mathematics, and that encourage children to communicate using the language and symbols of primary school mathematics through social play with literate and numerate adults as well as peers. Two games tasked children to add and compare large sets of dots based on their relative numerosity: abilities that are universal (*34*), emerge in infancy (*48*), and correlate with mastery of symbolic arithmetic in children (*43, 49*) and adults (*44*). A third game required that children establish exact one-to-one correspondence relations between sets of 1-4 2D shapes and sequences of 1-4 movements on a linear board, relating numerical magnitudes to positions on a line: abilities that emerge in infancy (*50, 51*) and produce short-term enhancements in children's symbolic number concepts (*7, 52*). Finally, two games challenged children to find a geometric property (e.g., shape, parallelism, connectedness) that distinguished one figure from a group of others, or to place objects at locations indicated on a set of small-scale geometric maps: two early developing, universal abilities (*53-56*) that are believed to promote learning of a variety of mathematical concepts (*42*).

The overall cost for a group of 6 children to play the games for 4 months was $316 (**Table S22**). This figure includes the cost of materials as well as a teacher's and monitor's time and training. The materials were a significant portion of these costs, $217, which suggests that if these games were scaled up, the actual operating costs would be significantly lower since materials could be reused and produced in larger quantities.

In the *no-treatment control* condition (72 schools), children received a systematic preschool curriculum designed by Pratham. This curriculum targeted 5 main aspects of child development: physical development; language development; social and emotional development; cognitive development; and creative development. Perhaps most relevant to learning mathematics in school, children played memory games, learned about sequences and matching, and learned numbers (as words and Arabic symbols) and spatial concepts (such as small/big and near/far).

For 3 1-hour sessions per week, children in the math and social games conditions played games instead of receiving the Pratham curriculum. During gameplay, the Pratham teachers focused on the younger children in their classes (who did not participate in gameplay), thereby reducing the time devoted to the regular Pratham curriculum for the older children. We did not specify what Pratham content teachers should reduce in order to evaluate a realistic intervention in which a preschool would chose to replace part of their curriculum with ours. It was possible that the math games could have had either positive or negative effects on primary school outcomes, regardless of their mathematical content: They could have had a positive effect if the games themselves were more effective than the current practices or a negative effect if symbolic skills provided by the regular curriculum were more immediately useful.

To distinguish between these possibilities and to test the specific effects of the games' mathematical content, our experimental design included a third group of schools assigned to the *active control* condition (70 schools). Children in these schools played games that followed the same rules and procedures as the math games and were comparably challenging and engaging (see SM), but that focused on two social cognitive abilities that are critical to assessing the intentions of others: emotion reading (*57*) and gaze following (*58*; **Figs. 1-2**). Like the abilities exercised by the math games, these abilities arise in infancy (*59, 60*) and predict later cognitive skills (*61*). They are also thought to foster language development and pedagogical learning in early childhood (*62*) and may be related to future labor market success (*63, 64*). The active control games therefore were truthfully presented to teachers and children as potentially valuable for enhancing school readiness. Since these games have the same rules as the math games, they further allowed us to distinguish the general effects of gameplay (e.g., communication, language, taking turns, etc.) from the specific effects of the mathematics content.

In the math and active control conditions, each game was introduced to children with easy practice problems, and children progressed as a class through a diversity of material during regular gameplay. As the intervention progressed, classes were also presented with more difficult problems to maintain children's engagement and interest in the games (see SM). Progression to these more difficult problems was gradual, as the games were meant to encourage in children a sense of confidence and success with the game content (*65*). There was no presumption that each class group would necessarily complete all or even many levels of the games. We created several levels in order to keep children engaged throughout the duration of the intervention.

Gameplay sessions were run by intervention teachers hired by Pratham: typically, young women with a high school education but no college degree. These teachers received brief training from our research team on how to play the games with children and how to evaluate children's performance. Each intervention teacher was responsible for 2 preschools, in which she led 3 1-hour sessions per week and kept notes during each session about the game, level, and deck that was played and the individual performance of each child. To monitor the implementation of the program, a separate team of "process monitors" made unannounced visits to the preschools, collecting data on gameplay frequency, adherence to the rules, and children's attention to and facility with the game content (see SM).

Evaluation: Data Collection and Empirical Specifications

Assessments evaluating the effects of the intervention were administered at 4 time points: during the month before the intervention (baseline); 0-3 months after the intervention (endline 1), 6-9 months after the intervention (endline 2), and 12-15 months after the intervention (endline 3). For all tests, children were tested individually on a laptop computer by local non-experts trained by our team to administer assessments to young children. Assessors were unaware of children's condition assignments and had not been involved in the gameplay curricula. Unlike the games, the assessments were presented in a non-social context, and they included difficult problems, challenging time constraints, and no informative feedback.

Tests of children's concepts and skills built on Pratham's experience evaluating children's learning of mathematics throughout India (*66, 67*). Tests of non-symbolic numerical and geometric abilities were based on research in cognitive science assessing these abilities in children and adults (*43*) in diverse cultures, including remote cultures with minimal education (*34, 53*). School-relevant assessments focused on comparing and adding numbers presented as words and Arabic numerals (*68*) and answering verbal questions about shape properties, similarity, and symmetry (*69*). Social skills were measured by evaluating children's sensitivity to gaze direction (*70*) and emotional expressions (*71*). Standardized measures of intelligence were not given, both because of resource limitations and because the aim of the intervention was to enhance children's learning of school mathematics. Because mathematics learning relates to children's mastery of language, to their developing executive functions, and to their motivation to tackle challenging problems, we also presented children with assessments of language and reading based on Pratham's tests of these abilities, and we adapted tests of executive function (*72*) and motivation for school learning (*73*) from tests that are widely used in cognitive science laboratories. Children in the math games, social games, and no-treatment preschools exhibited similar baseline achievement, as well as similar characteristics across basic demographic measures (**Table 1**).

While the baseline assessment and, for some children, the first endline assessment were presented to children in their preschools, the remaining endline assessments were presented to children in their homes. We surveyed 94%, 87%, and 84% of the original sample at the three successive endlines; 80% of the children were surveyed at all three endlines. There were no significant differences in the baseline test scores, demographic variables, or treatment statuses of those who dropped out of the study and those who completed the assessments at all of the time points (**Table S1A-B**). At endlines 2 and 3 respectively, 83% and 91% of the tested children were enrolled in primary school, and the proportion was similar in all treatment conditions (**Table S13**). Following a pre-specified "intention to treat" design, we included children in all assessments whether or not they were enrolled in primary school (see SM for analyses comparing children who did and did not progress to primary school) or received the intervention assigned to their preschool.

Due to the randomized design, analysis is straightforward. The main results are apparent by comparing the descriptive statistics for each test across the different conditions at all three endlines. We perform joint Fisher randomization inference tests of statistical significance on these basic comparisons, evaluating the hypothesis that children in the math games classrooms perform better on math questions than children in the social or no-treatment control classrooms (see SM text and **Table S2**). We also present analyses based on a regression specification, which was pre-registered, along with a complete pre-analysis plan, on socialscienceregistry.org (see SM). We used the following specified regression framework:

$$y_{i,j} = \beta_1 \text{math}_j + \beta_2 \text{social}_j + \beta_3 \text{age}_{i,j} + \beta_4 \text{gender}_{i,j} + \beta_5 \text{baseline}_{i,j} + \varepsilon_{i,j}$$

Where $y_{i,j}$ represents the endline value of an outcome for child *i* in school *j*; math_j is an indicator variable for whether school *j* was treated with the math games; social_j is an indicator variable for whether school *j* was treated with the social games; $\text{age}_{i,j}$ is age in months of child *i* in school *j*;

$\text{gender}_{i,j}$ is an indicator variable for gender of child *i* in school *j*; and $\text{baseline}_{i,j}$ is the baseline value of an outcome for child *i* in school *j*. Because the treatment was administered to all children in a given school, the standard errors are clustered at the school level. Our analysis plans also called for a specification without a baseline control. The two specifications revealed largely the same findings (see SM).

Four measures comprise our primary outcomes. Each outcome is based on a composite Z-score (computed by taking an average of the Z-scores for each individual on each test, relative to the mean and standard error of the control group's average baseline performance on that test), and so the coefficients for these outcomes can be interpreted as effect sizes in terms of standard deviations. The "math composite" includes all the math tests, the "non-symbolic composite" includes the math tests of approximate numerical comparison and of finding a deviant shape, and the "symbolic composite" includes the math tests assessing knowledge of number words and shape names, abilities to compare and add numbers presented as words and/or symbols, and at endlines in which they were presented, facility in answering verbal questions concerning relations of shape similarity and symmetry. At the baseline and first endline, the symbolic tests probed abilities that develop spontaneously in children living in educated families, prior to the start of schooling (for example, children's mastery of ordinary terms for shapes, such as "egg"). At the two later endlines, these tests focused primarily on abilities that are taught in school (for example, children's mastery of geometric terms for shapes, such as "rectangle"). The "social composite" includes a test probing sensitivity to gaze direction and, at endlines in which it was administered, a test probing knowledge of emotion words.

Results

Based on the data collected by the teachers and process monitors during the game play, we first asked whether the two game-based interventions were implemented, engaged children's interest, and led to improved performance over the course of the intervention. Both the math and social games were played regularly and most children attended to the game play. In the preschools where the math and social games were played, all the children attended to the games on 52% and 53% of the observed sessions. Most schools progressed through all of the materials included in the first level of play in at least one of the five games (93% in the math condition and 89% in the social condition), and most classrooms remained engaged with the games through the materials of the first 2 levels (**Table S6**). Children performed well in the first level of each game (between 61% correct and 87% correct, depending on the game, in the first two rounds of play with those materials), and their scores improved 7 percentage points, on average, between the first two and last two times that a level was played (see **Table S6**). Thus, we successfully designed a scalable preschool math games curriculum that was implemented as intended, led to progress within the game itself, and engaged children with its content.

The mean percentages of correct responding for each test, treatment group, and endline are reported in **Table 2**, and they tell a clear story. At endline 1, children in the math games group had a higher proportion of correct responding on all math tests compared to the children in the two other groups. For example, they scored 36% correct on the test of geometric sensitivity (chance = 17%), while the no-treatment and social control groups scored 25% and 29%. In contrast, and as expected, children in the social games group had a higher proportion of correct responding on gaze sensitivity. At endlines 2 and 3, children in the math games group still performed best on the non-symbolic math tests, but not on the symbolic measures targeting the concepts taught in school, which were very similar across the three groups. Children in the social games group still performed better on the test of gaze sensitivity.

Results from Fisher permutation tests (**Table S2**) confirm the statistical significance of these findings. Compared to the control condition, we reject the hypothesis of no effect of the math games treatment on all the math assessments, the symbolic assessments, and the non-symbolic assessments for all endlines taken together and for endline 1 individually. At endlines 2 and 3, we reject the hypothesis of no effect overall and no effect on the non-symbolic assessments, but not on the symbolic assessments. In contrast, we do not reject the hypothesis of no effect of the social games compared to the no-treatment control, for all endlines taken together or for each endline individually.

To summarize these results effectively, we used our pre-registered regression framework. We first tested whether, immediately after the intervention, children who had exposure to the math games curriculum had higher scores on the math assessments than those who did not. Consistent with the descriptive statistics, at the first endline, the math games led to a significant increase in the overall math composite: 0.25 standard deviations compared to the no-treatment control group ($t(213) = 5.88$, $P < 0.001$). There was also an impact of the social games on the overall math composite compared to the no-treatment control group, but this effect was smaller than that of the math games (**Fig. 3**, **Table 3**). Playing the math games therefore had a positive effect on children's subsequent performance on tests evaluating their sensitivity to number and geometry, relative both to children who received only the regular preschool curriculum and to children who played games with similar rules and materials but with no mathematical content.

The math games led to a particularly large increase in the non-symbolic math composite: 0.42 standard deviations above the no-treatment control group ($t(213) = 7.34$, $P < 0.001$). The social games also led to an increase on the non-symbolic math composite, but that increase was smaller than that of the math games (**Fig. 3**, **Table 3**). The social games had a similarly large impact on the social skills measure (0.44 standard deviations), while the impact of the math games on this measure was smaller (**Table 3**). These findings suggest that the difference between the impact of the two treatments was due to their different content, replicating and extending prior evidence that children's early-developing sensitivity to mathematical and social information improves with experience and exercise (*42*, *74*, *75*).

Do the gains in preschool children's intuitive, non-symbolic numerical and geometric skills lead to improvements in their knowledge of the symbols and language of formal mathematics? Consistent with this possibility, the children who played the math games outperformed the children in the no-treatment control group by 0.13 standard deviations on the symbolic mathematics composite ($t(213) = 2.70$, $P = 0.007$) at the first endline, whereas the children who played the social games did not. Nevertheless, the children in the math games condition showed only a relatively small advantage over those in the social games condition on this composite measure; the difference was significant only at the 10% level (**Fig. 4**, **Table 3**).

Further analyses of the first endline focused on children's performance on the individual assessments. Relative to the no-treatment control group, the math-games treatment individually impacted the tests probing non-symbolic numerical and geometric abilities as well as the tests probing knowledge of number words, Arabic numerals, and shape names, although not the test of simple verbal arithmetic (**Table S4**). There was no impact of the math or social games on the test of executive function, although performance on that test showed moderate test-retest reliability and strong effects of age (**Table S3, S8, S12, S14**). The measure of motivation also showed no impact of the math or social games, but children performed poorly and inconsistently on this measure (**Table S15**). The effects of the math games are therefore not attributable to changes in executive function, but we cannot determine whether they depend on changes to children's motivation. According to prior research, the effects are unlikely to be rooted in the Hawthorne effect, for example, in children's expectations about the effects of non-symbolic mathematical training (*76*).

As in laboratory-based studies (*38-40*, *42*, *77*), we thus observed that non-symbolic mathematical training caused short-term gains in symbolic mathematical outcomes. It is worth emphasizing that these gains were evaluated over the three months that followed the end of the intervention, a substantially longer follow up than that of a typical laboratory study in this domain. Moreover, the gains in mathematical skills after the intervention were due to the specific math games training, as opposed to the rule-based structure of the games or the increased attention of children in that treatment group. This training intervention, implemented in a field environment with minimally trained teachers and assessors, captured and sustained children's interest over months of gameplay. Its findings thus suggest that it is possible to translate the findings from basic cognitive science research into field experiments in children's everyday environments.

In light of these findings, we asked whether children's mathematical gains persisted in the longer term. The benefits of most educational interventions are short-lived, even when initial gains are significant (*78*) and especially when training is not reinforced by "booster" sessions (*79*). Remarkably, this is not the case in this study. At the two later endlines, 6 months and a year later, the overall math composite remained significantly improved in the math games group, and the gains were stable in magnitude between endlines 2 and 3 (compared to the no-treatment control group, 0.12 standard deviations at endline 2, $t(208) = 2.74$, $P = 0.007$, and 0.14 standard deviations at endline 3, $t(213) = 2.77$, $P = 0.006$, **Fig. 4**, **Table 3**). Though the effect of the math games treatment was smaller in endlines 2 and 3, compared to endline 1, the increase in math ability that the children in the social games group experienced in endline 1 vanished by endline 2 (0.01 standard deviations compared to the no-treatment control group). Thus, the differential impact of the math versus social games on the overall math composite was remarkably constant over the endline assessments (**Table 3**).

The gains on the non-symbolic composite by children in the math games group proved enduring through the first year of primary school: 0.29 standard deviations compared the no-treatment control group at the second endline ($t(208) = 4.59$, $P < 0.001$) and 0.32 standard deviations at the third endline ($t(213) = 4.32$, $P < 0.001$). The persistence of these improvements is striking because children had no access to the game materials after the intervention ended, and their homes and schools provided no opportunities to engage in related game activities or anything resembling them.

Do such enduring gains in preschool children's non-symbolic mathematical skills also improve their readiness to learn new mathematical content in primary school? The answer here was a decisive "no." By mid-way through the first year of primary school, the effect of the math-games training on the symbolic composite measure had disappeared. Although the math games caused persistent gains in children's non-symbolic mathematical abilities, they failed to enhance children's readiness for learning the new symbolic content presented in primary school. To better interpret these negative findings, we first asked whether the symbolic math assessments used in the later endlines were unreliable or invalid for this population. Contrary to these concerns, there were strong intertemporal relations (i.e., test-retest reliability) for each test for children in the no-treatment control group, indicating that our tests were highly reliable (**Table S5)**. Moreover, older children in the no-treatment control group performed better than younger children on each of the math tests, suggesting that the tests indeed were sensitive to developmental changes in these abilities (**Table S16**). Third, although the difficulty of the assessments increased across the three endlines, children showed stable performance on the symbolic math tests (**Table 3** and **Table S4**). These findings suggest that the assessments were valid measures of children's symbolic mathematical knowledge and that children were in fact learning mathematics during the first year of primary school.

Finally, we asked whether the failure to show persistent gains in symbolic mathematics, despite enduring gains in non-symbolic abilities, resulted from inherent differences between the ways that Indian and Western children think about or learn mathematics. Contrary to this possibility, the Indian preschool children in the no-treatment condition exhibited the characteristic profile of correlations across skills found in Western children. Among children in the U.S., there are strong correlations between non-symbolic and symbolic numerical abilities (*43, 80*) as well as correlations between sensitivity to the shapes of geometric forms and abilities to interpret geometric maps (*30*). Similarly, the performance of the Indian children on the non-symbolic math composite strongly correlated with performance on the symbolic math composite, both within and across time points. We also replicated the more specific correlations in the mathematical cognition literature, relating non-symbolic numerical acuity to symbolic number abilities, and relating sensitivity to geometric forms to performance on the tests of verbal geometric reasoning. Indeed, each of these correlations survived controls for performance on the tests in the other domain, both within and across time points (**Table 4A-B**). These parallels between our findings and those of laboratory studies of children in the U.S. and other developed countries, suggest that the negative findings of our intervention do not stem either from differences between the mathematical concepts of children in poor and wealthy countries or from differences in the ways that those concepts were measured in the lab and in the field. On the contrary, the preschool mathematical abilities revealed by laboratory studies of Western children proved to be both generalizable and robust, but an intervention exercising those abilities failed to enhance poor Indian children's learning of school mathematics.

Discussion

This study underscores the importance of field experiments to elucidate universal cognitive mechanisms underlying children's learning of mathematics. Previous research, based on robust correlations and laboratory studies using short-term training, raised the strong possibility that (*a*) universal, early emerging mathematical abilities would improve with exercise over the preschool years, and (*b*) such exercise would enhance children's subsequent learning of primary school mathematics. Our study demonstrates that the first part of the conjecture was correct, but not the second. Children's readiness for learning formal mathematics in India appears to require something more than improvement in non-symbolic numerical and geometric skills through games that make mathematics fun and show children that they can and do improve in this domain.

On the positive side, our results show that it is possible to translate the subtle manipulations of the laboratory into implementable interventions in the field. Children learned, played, and enjoyed the games. Their intuitive mathematical abilities improved with practice, and these improvements persisted more than a year after the completion of a game-based intervention that exercised them, despite the removal of the games and the absence of any similar resources to sustain children’s gains. The assessments of children’s cognitive gains, based not on standardized tests of intelligence but on laboratory-based measures of sensitivity to number and geometry, yielded findings that are highly similar to the findings from laboratory experiments in developed countries, despite large differences in the conditions under which the tests were administered and in the lives and environments of the children who took them. They revealed strong correlations between poor preschool children’s early emerging and intuitive numerical and geometric abilities and the symbolic mathematical content of primary school, just as in other populations from developed countries. Finally, parallel to the short-term results found in lab-based training experiments, the improvements in children’s intuitive mathematical abilities had a positive impact on their simultaneous learning of numerical and spatial language and symbols, which were used in the preschools where children played the games. Nevertheless, the preschool intervention had no evident effect on children’s subsequent learning of mathematics in primary school.

We conclude that a preschool intervention that effectively fosters an attunement of intuitive mathematical skills, in social and communicative contexts, is not sufficient to promote children’s later learning of school mathematics, at least as that learning is measured at the end of the first year of primary school and in the Indian context. This finding echoes the negative findings from other randomized control trials in developing countries, and it suggests a possible explanation. Preschool interventions may fail unless they are designed to complement a central feature of primary school in these settings, i.e., a strictly symbolic curriculum. Indeed, exploratory analyses show that the children who returned to preschools after the intervention showed more enduring effects of the math intervention than those who went on to primary school (**Table S13**).

These findings suggest two ways to redesign the intervention to make it more successful. First, a math games intervention might be more effective at fostering school readiness if the games were presented in a way that connects their non-symbolic mathematical content directly to the mathematical language and symbols used in school. For example, children could be introduced to mathematical language and symbols along with the card and board games that mainly exercise their intuitive abilities, or play versions of the games that alternate between pictorial materials and materials presenting words and symbols. Second, non-symbolic math games training might be more effective if training coincided with children’s learning of formal mathematics rather than preceding that learning. Future field experiments could test these and other possibilities.

Our findings underscore both the promise and the necessity of rigorous testing of reforms to school curricula inspired by basic science, using scalable programs over extended timeframes in the environments in which those curricula will be implemented. Laboratory-based experiments provide the most sensitive setting for discovering the cognitive and neural underpinnings of children’s learning, but they alone do not reveal the causal factors that produce knowledge over long time spans, or the most effective means for enhancing that knowledge in school. For those questions, cognitive science and public policy may advance in tandem, through research in homes, street corners, and classrooms, testing interventions that combine the diverse processes that together allow children to master new cognitive challenges.

References and Notes:

1. P. Glewwe, M. Kremer, “Schools, teachers, and education outcomes in developing countries” in *Handbook of the Economics of Education*, F. Hanushek, F. Welch, Eds. (North-Holland, 2006), pp. 945-1017.

2. World Education Forum, “Incheon Declaration” (2015; http://en.unesco.org/world-education-forum-2015/incheon-declaration).

3. ASER Centre, “Annual Status of Education Report 2014” (2015; http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/fullaser2014mainreport_1.pdf)

4. A. Banerjee, E. Duflo, *Poor economics: A radical rethinking of the way to fight global poverty* (PublicAffairs, 2012).

5. K. Muralidharan, “Field experiments in education in developing countries” in *Handbook of Field Experiments*, E. Duflo, A. Banerjee, Eds. (Elsevier, 2017), pp. 323-385.

6. UNESCO, “Education for all 2000-2015: achievements and challenges” (2015; http://unesdoc.unesco.org/images/0023/002322/232205e.pdf).

7. S. Griffin, R. Case, Evaluating the breadth and depth of training effects when central conceptual structures are taught. *Monogr Soc Res Child Dev* **61**, 83-102 (1996).

8. N. Hurtado, V. A. Marchman, A. Fernald, Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Dev Sci* **11**, F31-9 (2008).

9. B. N. Verdine, R. M. Golinkoff, K. Hirsh-Pasek, N. S. Newcombe, Finding the missing piece: Blocks, puzzles, and shapes fuel school readiness. *Trends Neurosci Educ* **3**, 7-13 (2014).

10. K. Muralidharan, A. Singh, A. J. Ganimian, “Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India” (National Bureau of Economics Research Working Paper No. 22923, 2016).

11. O. P. Attanasio, C. Fernández, E. O. Fitzsimons, S. M. Grantham-McGregor, *et al.*, Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: Cluster randomized controlled trial. *BMJ* **349**, g5785 (2014).

12. R. Banerji, J. Berry, M. Shotland, “The Impact of Maternal Literacy and Participation Programs: Evidence from a Randomized Evaluation in India” *Am Econ J Appl Econ* (2016).

13. S. M. Chang, S. M. Grantham-McGregor, C. A. Powell, M. Vera-Hernández, *et al.*, Integrating a parenting intervention with routine primary health care: A cluster randomized trial. *Pediatrics* **136**, 272-80 (2015).

14. L. C. Fernald, R. M. Kagawa, H. A. Knauer, L. Schnaas, *et al.*, Promoting child development through group-based parent support within a cash transfer program: Experimental effects on children's outcomes. *Dev Psychol* **53**, 222-36 (2017).

15. P. Gertler, J. Heckman, R. Pinto, A. Zanolini, *et al.*, Labor market returns to an early childhood stimulation intervention in Jamaica. *Science* **344**, 998-1001 (2014).

16. B. Ozler, L. C. Fernald, P. K. Kariger, C. McConnell, *et al.*, “Combining Preschool Teacher Training with Parenting Education: A Cluster-Randomized Controlled Trial” (World Bank Policy Research Working Paper No. 7817, 2016).

17. J. J. Heckman, Schools, skills, and synapses. *Econ Inq* **46**, 289 (2008).

18. J. Currie, D. Thomas, Does Head Start make a difference. *Am Econ Rev* **85**, 341-64 (1995).

19. D. Deming, Early childhood intervention and life-cycle skill development: Evidence from Head Start. *Am Econ J Appl Econ* **1**, 111-34 (2009).

20. E. Garces, D. Thomas, J. Currie, Longer-term effects of Head Start. *Am Econ Rev* **92**, 999-1012 (2002).

21. P. Kline, C. R. Walters, Evaluating public programs with close substitutes: The case of Head Start. *Q J Econ* **131**, 1795-848 (2016).

22. P. L. Engle, L. C. Fernald, H. Alderman, J. Behrman, *et al.*, Strategies for reducing inequalities and improving developmental outcomes for young children in low-income and middle-income countries. *Lancet* **378**, 1339-53 (2011).

23. S. Martinez, S. Naudeau, V. Pereira, “The promise of preschool in Africa: A randomized impact evaluation of early childhood development in rural Mozambique” (The World Bank, 2012; http://repositorio.minedu.gob.pe/bitstream/handle/123456789/3614/The%20Promise%20of%20Preschool%20in%20Africa%20A%20Randomized%20Impact%20Evaluation%20of%20Early%20Childhood%20Development%20in%20Rural%20Mozambique.pdf?sequence=1).

24. J. J. Heckman, S. H. Moon, R. Pinto, P. A. Savelyev, A. Yavitz, The rate of return to the highscope perry preschool program. *J Public Econ* **94**, 114-28 (2010).

25. T. W. Watts, G. J. Duncan, D. H. Clements, J. Sarama, What is the long-run impact of learning mathematics during preschool? *Child Dev* (2017; 10.1111/cdev.12713).

26. H. Yoshikawa, D. Leyva, C. E. Snow, E. Treviño, *et al.*, Experimental impacts of a teacher professional development program in chile on preschool classroom quality and child outcomes. *Dev Psychol* **51**, 309-22 (2015).

27. S. Carey, *The origin of concepts* (Oxford UP, 2009).

28. S. Dehaene, *The number sense: How the mind creates mathematics* (Oxford UP, 2011).

29. S. Dehaene, V. Izard, E. Spelke, P. Pica, Log or linear? Distinct intuitions of the number scale in western and amazonian indigene cultures. *Science* **320**, 1217-20 (2008).

30. M. R. Dillon, Y. Huang, E. S. Spelke, Core foundations of abstract geometry. *Proc Natl Acad Sci U S A* **110**, 14191-5 (2013).

31. L. Feigenson, S. Dehaene, E. Spelke, Core systems of number. *Trends Cogn Sci* **8**, 307-14 (2004).

32. R. Gelman, C. R. Gallistel, *The child's understanding of number* (Harvard UP, 1978).

33. S. F. Lourenco, J. Huttenlocher, The representation of geometric cues in infancy. *Infancy* **13**, 103-27 (2008).

34. P. Pica, C. Lemer, V. Izard, S. Dehaene, Exact and approximate arithmetic in an amazonian indigene group. *Science* **306**, 499-503 (2004).

35. R. S. Siegler, J. L. Booth, Development of numerical estimation in young children. *Child Dev* **75**, 428-44 (2004).

36. E. S. Spelke, "Natural number and natural geometry" in *Space, time and number in the brain*, S. Dehaene, E. Brannon, Eds. (Elsevier, 2011), pp. 287-318.

37. D. H. Uttal, Seeing the big picture: Map use and the development of spatial cognition. *Dev Sci* **3**, 247-64 (2000).

38. J. Park, E. M. Brannon, Training the approximate number system improves math proficiency. *Psychol Sci* **24**, 2013-9 (2013).

39. D. C. Hyde, S. Khanum, E. S. Spelke, Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition* **131**, 92-107 (2014).

40. S. Khanum, R. Hanif, E. S. Spelke, I. Berteletti, D. C. Hyde, Effects of non-symbolic approximate number practice on symbolic numerical abilities in Pakistani children. *PLoS One* **11**, e0164436 (2016).

41. J. Park, V. Bermudez, R. C. Roberts, E. M. Brannon, Non-symbolic approximate arithmetic training improves math performance in preschoolers. *J Exp Child Psychol* **152**, 278-93 (2016).

42. D. H. Uttal, N. G. Meadow, E. Tipton, L. L. Hand, *et al.*, The malleability of spatial skills: A meta-analysis of training studies. *Psychol Bull* **139**, 352-402 (2013).

43. J. Halberda, M. M. Mazzocco, L. Feigenson, Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* **455**, 665-8 (2008).

44. J. Halberda, R. Ly, J. B. Wilmer, D. Q. Naiman, L. Germine, Number sense across the lifespan as revealed by a massive internet-based sample. *Proc Natl Acad Sci U S A* **109**, 11116-20 (2012).

45. J. E. Lauer, S. F. Lourenco, Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychol Sci* **27** 1291-1298 (2016).

46. K. S. Mix, S. C. Levine, Y. L. Cheng, C. Young, *et al.*, Separate but correlated: The latent structure of space and mathematics across development. *J Exp Psychol Gen* **145**, 1206-27 (2016).

47. D. H. Bailey, T. W. Watts, A. K. Littlefield, D. C. Geary, State and trait effects on individual differences in children’s mathematical development. *Psychol Sci* **25**, 2017-26 (2014).

48. F. Xu, E. S. Spelke, Large number discrimination in 6-month-old infants. *Cognition* **74**, B1-B11 (2000).

49. H. Barth, K. La Mont, J. Lipton, E. S. Spelke, Abstract number and arithmetic in preschool children. *Proc Natl Acad Sci U S A* **102**, 14116-21 (2005).

50. M. D. de Hevia, V. Izard, A. Coubart, E. S. Spelke, A. Streri, Representations of space, time, and number in neonates. *Proc Natl Acad Sci U S A* **111**, 4809-13 (2014).

51. L. Feigenson, S. Carey, Tracking individuals via object-files: Evidence from infants’ manual search. *Devl Sci* **6**, 568-84 (2003).

52. R. S. Siegler, G. B. Ramani, Playing linear number board games but not circular ones improves low-income preschoolers numerical understanding. *J Educ Psychol* **101**, 545 (2009).

53. S. Dehaene, V. Izard, P. Pica, E. Spelke, Core knowledge of geometry in an amazonian indigene group. *Science* **311**, 381-4 (2006).

54. J. Huttenlocher, N. Newcombe, M. Vasilyeva, Spatial scaling in young children. *Psychol Sci* **10**, 393-8 (1999).

55. J. S. DeLoache, Rapid change in the symbolic functioning of very young children. *Science* **238**, 1556-7 (1987).

56. N. Winkler-Rhoades, S. C. Carey, E. S. Spelke, Two-year-old children interpret abstract, purely geometric maps. *Dev Sci* **16**, 365-76 (2013).

57. M. C. Moulson, K. Shutts, N. A. Fox, C. H. Zeanah, *et al.*, Effects of early institutionalization on the development of emotion processing: A case for relative sparing? *Dev Sci* **18**, 298-313 (2015).

58. M. J. Doherty, J. R. Anderson, L. Howieson, The rapid development of explicit gaze judgment ability at 3 years. *J Exp Child Psychol* **104**, 296-312 (2009).

59. B. M. Hood, J. D. Willen, J. Driver, Adult's eyes trigger shifts of visual attention in human infants. *Psychol Sci* **9**, 131-4 (1998).

60. A. S. Walker-Andrews, Emotions and social development: Infants' recognition of emotions in others. *Pediatrics* **102**, 1268-71 (1998).

61. T. Gliga, M. Elsabbagh, K. Hudry, T. Charman, M. H. Johnson, Gaze following, gaze reading, and word learning in children at risk for autism. *Child Dev* **83**, 926-38 (2012).

62. G. Csibra, G. Gergely, Natural pedagogy. *Trends Cogn Sci* **13**, 148-53 (2009).

63. D. J. Deming, "The growing importance of social skills in the labor market" (National Bureau of Economics Research Working Paper No. w21473, 2015).

64. J. J. Heckman, T. Kautz, "Fostering and measuring skills: Interventions that improve character and cognition" (National Bureau of Economics Research Working Paper No. w19656, 2013).

65. J. J. Wang, D. Odic, J. Halberda, L. Feigenson, Changing the precision of preschoolers' approximate number system representations changes their symbolic math performance. *J Exp Child Psychol* **147**, 82-99 (2016).

66. ASER Centre, "Annual Status of Education Report 2011" (2012; http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER_2011/aser_2011_report_8.2.12.pdf)

67. S. B. Vagh, "Validating the ASER testing tools: Comparisons with reading fluency measures and the Read India measures" (2010; http://img.asercentre.org/docs/Aser%20survey/Tools%20validating_the_aser_testing_tools__oct_2012__3.pdf.

68. H. Ginsburg, A. J. Baroody, *The test of early mathematics ability* (Pro-Ed, 2003).

69. Everyday Mathematics, (2016; http://everydaymath.uchicago.edu).

70. A. Frischen, A. P. Bayliss, S. P. Tipper, Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychol Bull* **133**, 694 (2007).

71. X. Gao, D. Maurer, Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *J Exp Child Psychol* **102**, 503-21 (2009).

72. M. C. Davidson, D. Amso, L. C. Anderson, A. Diamond, Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* **44**, 2037-78 (2006).

73. C. S. Dweck, *Self-theories: Their role in motivation, personality, and development* (Psychology Press, 2000).

74. V. Izard, P. Pica, S. Dehaene, D. Hinchey, E. S. Spelke, “Geometry as a universal mental construction” in *Space, time and number in the brain*, S. Dehaene, E. Brannon, Eds. (Citeseer, 2011), pp. 319-332.

75. L. S. Scott, A. Monesson, The origin of biases in face perception. *Psychol Sci* **20**, 676-80 (2009).

76. M. R. Dillon, A. C. Pires, D. C. Hyde, E. S. Spelke, Children's expectations about training the approximate number system. *Br J Dev Psychol* **33**, 411-8 (2015).

77. J. Park, E. M. Brannon, Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition* **133**, 188-200 (2014).

78. T. Andrabi, J. Das, A. I. Khwaja, T. Zajonc, Do value-added estimates add value? Accounting for learning dynamics. *Am Econ J Appl Econ* **3**, 29-54 (2011).

79. A. Banerjee, S. Cole, E. Duflo, L. Linden, “Balsakhi” (2007; http://hdl.handle.net/1721.1/39123).

80. Q. Chen, J. Li, Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychol (Amst)* **148**, 163-72 (2014).

81. J. Halberda, and R. Ly, “Panamath [computer software]” (2011; http://panamath.org/testyourself.php).

82. N. Tottenham, J. W. Tanaka, A. C. Leon, T. McCarry, *et al.*, The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Res* **168**, 242-9 (2009).

Acknowledgments:
The data and analyses presented in this paper are available at: doi:10.7910/DVN/LCLKDT. This work was funded by grants from The Union Bank of Switzerland, Optimus Foundation to J-PAL South Asia, from the Mind, Brain, and Behavior Initiative at Harvard University to E. Spelke, from the Foundations of Human Behavior Program at Harvard University to E. Spelke, and from the National Science Foundation to M. Dillon (DGE-1144152) and to J. Dean (DGE-1122374). The project received IRB approval from MIT and IFMR and was registered in the social science registry (https://www.socialscienceregistry.org/trials/27/history/15141).

We thank C. Cooper, K. Goffin, L. Schiller, and E. Schmidt for assistance in generating game materials; K. Dua, M. Duhon, A. Gupta, and A. Jha for assistance with intervention piloting, implementation, and evaluation; X. Gao and D. Mauer for providing images for the social games; N. Tottenham and C. Nelson for advice on the use of these images; J. Halberda for assistance in generating the images for the numerical games and assessments; Y. Huang for assistance in calibrating the math and social games; B. Reilly for assistance in the design and functioning of the math and social games; and K. Shutts for suggestions on elements of the social games. We thank P. Kakaji, M. Sharma, S. Sharma, Pratham, and the Pratham team of teachers for generously providing us with the opportunity to present these games to their preschool children in the warm and engaging environments that they have created for young learners in India and around the world. Finally, we want to thank R. Banerji, Pratham CEO, who was initially due to be an investigator, for many inspiring discussions throughout the course of this project.

Supplementary Materials

Materials and Methods

Supplementary Text

Figure S1-S20

Tables S1-S22

References (*83-91*)

Fig. 1. Materials from three math games (left column) and the corresponding social games (right column). The math games focused on either comparison of numerical magnitudes (top left), categorization of different shapes (middle left), or symbol reading based on an analysis of the features of a geometric form (bottom left). The corresponding social games focused on either comparison of emotional intensities (top right), categorization of different emotional expressions (middle right), or symbol reading based on an analysis of a face's gaze direction (bottom right). One pair of corresponding math and social games (top) involved sorting cards into one of two piles depending on the color of the larger number or greater expression of happiness. Another pair of corresponding games (middle) involved finding the figure that did not belong with the other figures based on its shape or expression. A third pair of corresponding games (bottom) involved using the shape of a figure or the gaze direction of a face on a 20 cm x 20 cm map to find a corresponding location on a 1 m x 1 m mat, which appeared at varied orientations; children placed an object on the location on the mat that was indicated on the map. The dot arrays in the top-left math game were created with Panamath (*81*), a free software used for generating numerical stimuli. The faces in the top two social games were obtained from Gao and Maurer (*71*), who adapted them from the face battery created by Tottenham *et al.* (*82*). The middle-right faces have been pixelated for display purposes only; children played this game with the non-pixelated faces.

Fig. 2. Children in the intervention playing the math (top) and social (bottom) versions of the linear board game. In the math game, there was one deck of face-down cards; children spun a spinner, whose arrow indicated how many cards they could choose from the deck. When turned over, the cards displayed either 1-4 small figures or an "X"; children moved their token forward on the board by one space for each figure on their card(s). In the social game, there were two decks of face-down cards, each with a different color on their back; the spinner depicted a face; when spun, its gaze indicated which colored deck(s) children could choose from. When turned over, each card displayed another face looking at a colored dot; children moved their token forward according to the colors on the board indicated by the gaze direction on the cards. Children therefore used either number or gaze to establish correspondences between the spinner, the cards, and the movements of a token on a board.

Fig. 3. Z-scores of children's performance on the three primary math outcome measures after the intervention, but before the start of primary school. Colored bars show the impact of the math and social treatments on each outcome measure. Error bars represent standard errors clustered at the school level; coefficient estimates indicate differences from the no-treatment control group. On the coefficients, stars indicate a rejection of the null hypothesis of no difference compared to the no-treatment condition (omitted category) $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

Fig. 4. Z-scores of children's performance on the three math composite measures at endlines 2 and 3. Colored bars show the impact of the math and social treatments on each measure mid-way through the first year of primary school (EL2) and after one year of primary school (EL3). Error bars represent standard errors clustered at the school level and coefficient estimates indicate differences from the no-treatment control group. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$

Table 1. Demographic information and baseline scores for children randomized to the three conditions in this study. Individual tests of joint equality of the Math Games, Active Control, and No-Treatment Control conditions (with standard errors clustered at the school level) for each measure (see SM for detailed descriptions of each measure) revealed no differences between groups. A Chi-Squared Test of joint equality across all measures also revealed no difference. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$

Table 2. Descriptive statistics for each test, for each treatment group, and for each endline. Mean percentages of correct responding are listed in each cell above their standard deviations (in parentheses).

Table 3. Coefficients from a linear regression model estimated using Ordinary Least Squares, controlling for age, gender, and baseline tests scores for each of the four main outcomes. Assessment time points consist of 3-month intervals beginning immediately after the intervention (before the start of primary school), 6 months after the intervention (mid-way through the first year of primary school), and 12 months after the intervention (after 1 year of primary school). The first two rows for each endline panel compare math and social treatments to no treatment (respectively), the third row indicates the results of a two-sided test of equality between the math and social coefficients, and the fourth row presents the no-treatment control group's mean performance. Standard errors (in parentheses) are clustered at the school level. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$

Table 4. (**A**) Coefficients from a linear regression model estimated using Ordinary Least Squares and controlling for age and gender illustrate the relation between non-symbolic numerical discrimination (*43*) and a symbolic numerical composite score (including tests probing knowledge of number words and simple arithmetic) calculated with separate regressions at each contemporaneous time point (baseline [BL] and three endlines [EL]) as well as across time points. All regressions control for the effects of non-symbolic and symbolic geometric abilities. (**B**) Coefficients from a linear regression model estimated using Ordinary Least Squares and controlling for age and gender illustrate the relation between non-symbolic geometric discrimination and a symbolic geometric composite score (including tests probing knowledge of shape words and judgments about shape properties) calculated with separate regressions at each contemporaneous time point (baseline [BL] and three endlines [EL]) as well as across time points. All regressions control for the effects of non-symbolic and symbolic numerical abilities.