Citation: Opfer, J. E., Kim, D., & Qin, J. (2018). How does the "learning gap" open? A cognitive theory of nation effects on early mathematics proficiency. In Berch, D. B., Geary, D. C., & Koepke, K. M. (Eds.), Mathematical Cognition and Learning Book Series (Vol. 4). Academic Press.

Chapter 5

How Does the "Learning Gap" Open? A Cognitive Theory of Nation Effects on Mathematics Proficiency

John E. Opfer, Dan Kim, and Jike Qin
The Ohio State University, Columbus, OH, United States

The mathematics proficiency of citizens plays a large role in the economic growth of nations (Hanushek & Woessmann, 2008), and individual differences in math proficiency predict a wide range of personal economic and educational outcomes. Among these are likelihood of completing high school, college grades, college graduation, starting salary, and salary growth (Murnane, Willett, & Levy, 1995; National Mathematics Advisory Panel, 2008; Sadler & Tai, 2007). Indeed, the dividends of high math proficiency have increased over time, due largely to secular increases in the wage premium for college attainment (Tyler, 2004).

Mathematics proficiency also differs greatly among nations (Gonzales et al., 2008; Husén, 1967; OECD, 2010). In two of the most widely cited cross-national comparisons, the Program for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS), the difference in math scores between the top-ranking nations and the bottom-ranking nations has been about three standard deviations. To grasp this difference intuitively, if men in nation A were three standard deviations taller than U.S. men (average height = 5 ft, 10 in.), then men in nation A would have an average height of 6 ft, 6 in. tall. Thus, the difference in math scores between top-ranking nations and bottom-ranking nations is like the difference between the average height of U.S. men and the average height of a professional basketball player. Of course, in economic terms, the effect of math proficiency on income is much greater than the effect of height (Mankiw & Weinzierl, 2010; Rose & Betts, 2004)—and thus a good reason to care about these very large national differences in math proficiency.

Mathematical Cognition and Learning, Vol. 4. https://doi.org/10.1016/B978-0-12-812574-8.00005-5

Although the top spots vary somewhat from year to year, there has been a very large and consistent difference in the mathematical performance of U.S. and East Asian students. The difference between U.S. and East Asian students came to the broad attention of scientists largely due to the efforts of Harold Stevenson and his colleagues, first in an article that appeared in *Science* (Stevenson, Lee, & Stigler, 1986) and more broadly publicized in his book *The Learning Gap* (Stevenson & Stigler, 1992). Not much has changed over this period of time. In the 25 years since the publication of *The Learning Gap*, East Asian students have routinely outperformed their American peers across a range of math assessments (Bailey et al., 2015; Stevenson, Chen, & Lee, 1993), with the magnitude of this "learning gap" typically appearing before the first year of formal schooling and growing larger with each successive year (Geary, Bow-Thomas, Liu, & Siegler, 1996).

In this chapter, we review two approaches for understanding these national differences in math abilities. The first approach attempts to link cross-national differences in educational input to cross-national differences in test scores. A tacit assumption of this approach is that children's performance on tests of mathematics is a direct effect of input. The second approach attempts to link cross-national differences in educational input to cross-national differences in test scores by means of indirect effects through more basic cognitive abilities. The assumption of this approach is that the specific educational inputs that differ among nations do not make much difference unless they have an effect on basic cognitive abilities that improve learning for all students. We argue that this approach has two related virtues. Because it is mechanistic, it greatly constrains the search for potential causes. Also, because it is mechanistic, it offers a clearer path to closing the learning gap. We then examine our theory in light of existing cross-national research on math-related cognitive abilities and preview results of a study (under review) that tested our idea.

THE TRADITIONAL APPROACH: DIFFERENCES IN INPUT, DIFFERENCES IN OUTPUT

Early work on cross-national differences in educational outcomes involved intensive analyses of how nations differ in their educational practices (Husén, 1967). Data on this topic came from a wide range of sources, including surveys and interviews with parents and teachers, systematic video recordings of classroom lessons, and experimental studies seeking to uncover the implicit attitudes of students living in different nations. Many potentially important national differences in educational input were uncovered in this early work, including cross-national differences in the quantity (Stevenson & Stigler, 1992) and quality of math instruction (Richland, Zur, & Holyoak, 2007; Stigler & Hiebert, 1999), quality of teacher's math knowledge (Ma, 1999), and cultural differences in math attitudes (Chen & Uttal, 1988; Stevenson & Stigler, 1992).

Quantity of Math Instruction

One of the most powerful predictors of academic achievement generally is "time on task," which is partly a function of the time that a student is available to learn (Carroll, 1989). This quantity of time differs greatly between East Asian and American schools, where the former has longer school days and more days in the school year. As a result, Chinese and Japanese children finish the sixth grade having spent between 1500 and 3000 more hours (equivalent to 1–2 more years) in school than their American peers (Stevenson & Stigler, 1992). This difference might not be as important as it seems when one considers that Chinese and Japanese children spend more time at recess and in lunch than American children. Alternatively, one must also recall that Chinese and Japanese children spend more time in teacher-directed activities, more time allocated for mathematics instruction, more time within math lessons involving teacher activity, and less time within math lessons involving "off-task" activities (Stevenson & Stigler, 1992). Thus, compared with their American counterparts, East Asian students spend more days of the year in school, more hours of the day in school, more time of the day in math class, and more time in math class encountering math instruction.

Quality of Math Instruction

All other things being equal, more time with math instruction would be expected to lead to greater math learning. Many studies, however, have shown that these potential gains may be compounded by higher quality math instruction in East Asian schools. Measuring instruction quality is a notoriously difficult task. However, experimental studies on the features of high-quality math lessons typically point to the importance of using real-world problems, connecting symbolic to concrete representations of quantity, using students' answers in discussion, and reflecting on students' errors for insight (National Research Council (NRC), 2005). These are also much more likely to be features of East Asian math classes than American ones according to observational studies in Sendai, Taipei, Beijing, and Chicago (Stevenson & Stigler, 1992; for recent reviews, see Geary, Berch, Ochsendorf, & Mann Koepke, 2017).

Another feature of high-quality instruction—one that we have used in our own studies teaching number line estimation (Thompson & Opfer, 2008)—involves providing cognitive support for connecting novel problems to familiar ones (Richland, Stigler, & Holyoak, 2012; Richland et al., 2007). Despite the potential of these cognitive supports to help young children, their use is by no means ubiquitous in the mathematics classroom—at least not in American classrooms. Support for this conclusion comes from a secondary analysis performed by Richland et al. (2007) on the video portion of the Trends in International Mathematics and Science Study (Stigler, Gonzales, Kwanaka, Knoll, & Serrano, 1999), a large-scale international comparison of mathematics instruction.

To study how often teachers provided cognitive support for analogies between familiar and novel problems, Richland et al. (2007) analyzed how analogies were used in mathematics lessons in the United States and two regions in Asia—Hong Kong and Japan—where students consistently outperform U.S. students in mathematics achievement and where the gap between high and low achievers is substantially smaller. Richland et al. specifically coded use of six cognitive supports for analogy that structure-mapping theory had identified as potentially helpful: (1) using a familiar source analog to compare with the target analog being taught, (2) presenting the source analog visually, (3) keeping the source visible to learners during comparison with the target, (4) using spatial cues to highlight the alignment between corresponding elements of the source and target (e.g., diagramming a scale below the equal sign of an equation), (5) using hand or arm gestures that signaled an intended comparison (e.g., pointing back and forth between a scale and an equation), and (6) using mental imagery or visualizations (e.g., "picture a scale when you balance an equation"). For every one of these six cognitive supports coded, teachers in the U.S. sample were less likely to promote relational learning than were teachers in Hong Kong and Japan. Thus, in addition to providing more math instruction, East Asian teachers also appear to provide better math instruction.

Quality of Teacher Knowledge

Teachers' ability to provide high-quality math instruction partly depends on their own fluency with mathematical content, including their ability to envision real-world math problems, to link the elements of the problem to mathematical language and to discuss the mathematical ideas implied by their students' errors. Unfortunately, American elementary school teachers struggle with mathematical concepts much more than their East Asian peers (e.g., Lo & Luo, 2012; Ma, 1999). For example, in one cross-cultural study of American and Chinese early mathematics teachers, nearly all of the 23 American teachers who were interviewed struggled to accurately solve and generate word problems about fraction division (Ma, 1999). In contrast, not only all of the Chinese teachers could accurately solve fraction division problems, but also they were able to describe multiple ways of conceptualizing fraction division and multiple strategies for approaching this topic with their students.

Cultural Differences in Math Attitudes

A final difference in educational input concerns students' willingness to spend the time required to learn math, that is, to expend effort. Stevenson and Stigler (1992) have documented many ways that attitudes toward effort differ greatly between American and East Asian parents, teachers, and students. American mothers were found to rate the importance of effort for academic achievement as lower than did their East Asian counterparts, and American mothers rated the importance of ability much higher than East Asian mothers. Also, American students were found to rate the importance of ability as higher than did Chinese students. Finally, when given a difficult math problem, American students spent much less time trying to solve it than their East Asian counterparts.

Beyond Habits of Highly Successful Educational Systems

The traditional approach of documenting differences in educational input has an obvious and intuitive appeal; for children in a given nation to be doing better at mathematics, those children must have been exposed to something in their environment that would lead to better outcomes. Thus, the scientist's job is simply to find what those environmental factors are.

At first glance, the only alternative approach to the traditional one would posit that children in different nations have different intrinsic math ability (e.g., Lynn, 1982; Rushton, 1992). Such an essentialist approach, however, does not appear to be supported by the evidence. First, the "learning gap" seems to be a recent phenomenon. That is, although contemporary Chinese students outperform American students in arithmetic, elderly Chinese individuals perform no better than elderly Americans (Geary et al., 1996). This would not be expected if Chinese students were intrinsically better at math than Americans. Yet, another reason to reject the essentialist approach is that within-nation differences in many basic cognitive abilities—such as IQ, perceptual speed, working memory, and spatial skill—are as large or larger than between-nation differences, at least when good translations of the tests of these abilities are used (Geary et al., 1996; Stevenson et al., 1985). Thus, with the essentialist approach having been disconfirmed, the traditional approach seems rather sensible.

The traditional approach, however, has two major problems. The first is that—in the absence of a good theory of how math proficiency develops—it is tempting to consider any and every national difference in child socialization as potentially causing national differences in math proficiency. The problem is not simply a matter of inferring causation from correlations (often a reasonable source for our hypotheses). Rather, the problem is chasing down one dependent variable (math scores) with too many independent variables, which is bound to lead to some spurious correlations and wrong-headed educational reforms.

Another limitation of the traditional approach is that it is nonmechanistic. That is, even if national differences in the quantity or quality of math instruction do reliably lead to national differences in math ability, we would want to know *how*. Does increasing the quantity of math instruction lead directly or indirectly to greater math ability? Distinguishing between direct and indirect effects is important both theoretically and practically. The issue is important theoretically because it allows us to identify the mechanisms by which nations produce better math students. The issue is also important practically because it allows us to identify novel targets for interventions.

To illustrate, consider a problem with many parallels to the “learning gap” that Stevenson identified: the gap between children of high and low socioeconomic status. Like the difference between East Asian and American children, children in low-income families achieve lower math scores than peers in higher-income families (Case, 1975; Jordan, Huttenlocher, & Levine, 1992; Saxe, Guberman, & Gearhart, 1989), despite going to schools with a similar number of days in the school year, a similar number of hours in the school day, a similar cultural beliefs about effort and ability, and so on. Intriguingly, children from low-income families were also found to lag considerably behind their middle-income peers in their knowledge of fundamental aspects of number itself, thereby impairing their ability to compare the size of Arabic numbers, to order a group of four numbers from smallest to largest, or to determine which of two numbers was closer to a target number (Case & Griffin, 1990; Case & Sandieson, 1987; Griffin, Case, & Sandieson, 1992). Put another way, the performance of low-income children suggested they lacked a “mental number line” that associated ordered numbers with a linearly increasing series of magnitudes. Hypothesizing that this “central conceptual structure” might explain much of the learning gap, Griffin, Case, and Siegler (1994) developed a program to improve low-income children’s numerical magnitude judgments, such as having them play games where they had to locate numbers on a number line. The results were quite dramatic, with the children in the experimental training group catching up to their middle-income peers in tests of arithmetic ability. Thus, rather than simply documenting all the ways that children from low- and high-income families differ, the investigators made progress in closing the learning gap by identifying a crucial *cognitive* difference between the groups and targeting it for intervention.

A COGNITIVE THEORY ON HOW THE LEARNING GAP OPENS

Our theory is that national differences in early math proficiency emerge for the same reason that socioeconomic differences lead to differences in math proficiency. That is, what mediates the effect of socioeconomic and national differences on early math proficiency is the same cognitive feature—the quality of numerical magnitude judgments, especially symbolic numerical magnitude judgments.

As we will argue, numerical magnitude judgments are important for mathematical thinking at any given point in time—a point that is consistent with

Siegler, Thompson, and Schneider's (2011) theory about fraction learning and recent work on math achievement in at-risk preschoolers (Geary & vanMarle, 2016). Additionally, numerical magnitude judgments help to foster *learning* from mathematical instruction (Booth & Siegler, 2008; Siegler & Ramani, 2008). In this way, high-quality numerical magnitude judgments can be thought of as a kind of malleable math aptitude. As this aptitude rises, the amount of time that a student needs in math instruction to reach a given level of competence will decline. If true, this theory is important because it suggests that East Asian students' math learning is not likely to be matched simply by increasing American students' time in math class, the quality of their math instruction, or their willingness to spend time on mathematics. At the same time, it suggests that the learning gap between East Asian and American children may be closed—perhaps substantially—by boosting American students' numerical magnitude judgments.

Why Numerical Magnitude Judgments?

Math proficiency—and quantitative reasoning more broadly—entails the ability to encode magnitudes, and the development of math proficiency builds and expands on the kinds of magnitudes that can be accurately encoded at any given age (Butterworth, 2005; Dehaene, 1997). What kinds of information-processing capacities must develop to encode these relations among magnitudes? Two kinds—ones that correlate with culturally invariant vs culturally variant abilities—are processing of nonsymbolic and symbolic numeric magnitudes, respectively. Proficiency at processing nonsymbolic numeric magnitudes (or "number sense acuity") (Halberda, Mazzocco, & Feigenson, 2008) involves facility with the ability to represent and compare nonsymbolic quantities accurately, such as comparing the relative quantities of two collections of items. Proficiency at processing symbolic numeric magnitudes (or "numeracy") (Hofstadter, 1982) involves the ability to accurately translate between magnitudes and symbolic numbers (e.g., "9," "nine," "1/9") and to encode quantitative relations among the numbers (e.g., 10 is closer to 1 than 100; 1/10 is closer to 1/100 than 1).

The number acuity/numeracy distinction provides a useful way to think about the varieties of numerical magnitude judgments that might lead to national differences in math proficiency. Numbers were only invented in the last 10,000 years of human history; the human brain is unlikely to have evolved specifically to handle numbers (Geary, 1995). Instead, numbers must be learned by modifying the brain through socialization and other interactions with the environment. As we will see, number acuity and numeracy differ in many ways that are consistent with this theoretical perspective. Number acuity emerges in infancy; numeracy develops much later. Number acuity is approximate and ratio-dependent; numeracy allows precision and encoding of mathematically important (linear) quantitative relations. Finally, numeracy varies widely among children growing up in different nations, whereas number acuity shows less variation and (where it exists) may be diagnostic of neurocognitive disorders (Piazza et al., 2010). Therefore, just as the distinction between processing of speech sounds (/kæt/) and signs (cat) is important for reading, the distinction between number acuity and numeracy is likely to be an important one for arithmetic generally (Vanbinst, Ansari, Ghesquière, & De Smedt, 2016) and in cross-cultural studies of arithmetic in particular.

Varieties of Numerical Magnitude Judgments

In the following section, we review findings on three major types of numerical magnitude judgments: comparing sets and numbers, estimating the location of a set or number on a number line, and approximately adding two sets or numbers (Fig. 1).

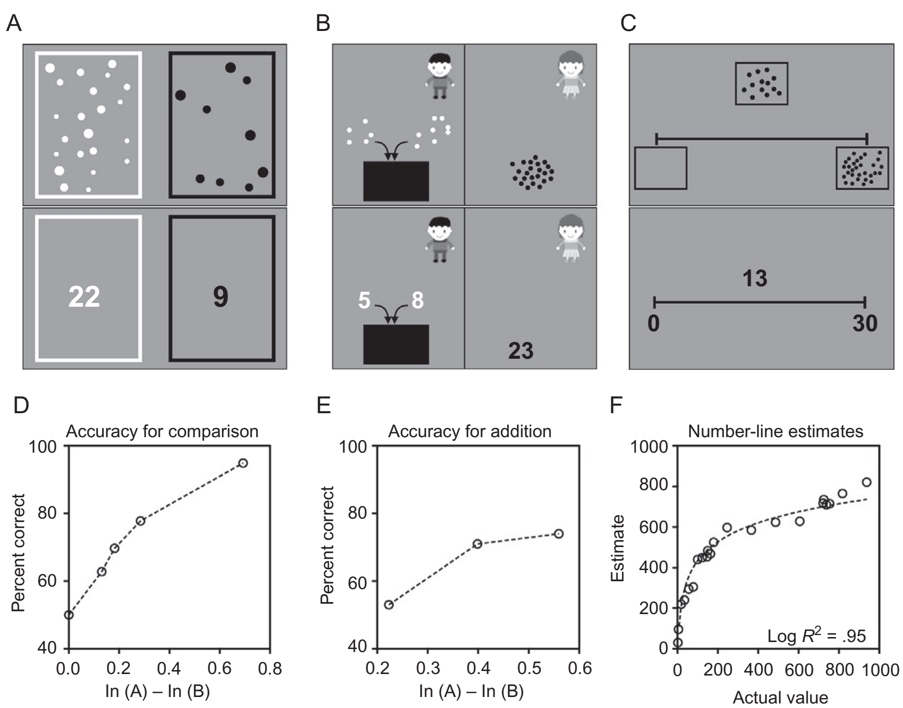

FIG. 1 Illustration of six numerical magnitude tasks and evidence for compressive representation of numbers. (A) Nonsymbolic and symbolic number comparison. (B) Nonsymbolic and symbolic approximate addition. (C) Nonsymbolic and symbolic number line estimation. (D) Ratio effect in comparison (Halberda et al., 2008) (A > B). (E) Ratio effect in approximate addition (Barth, La Mont, Lipton, & Spelke, 2005) (A > B, on half of trials A was the sum of two addends). (F) Logarithmic responses in number line estimates (Opfer & Siegler, 2007).

Comparing Sets and Numbers

A breakthrough in research on the development of mathematical abilities came from developmental psychologists' discovery that human infants notice changes in set sizes, typically sets of 2 vs 3 objects (Starkey & Cooper, 1980; Strauss & Curtis, 1981) and prefer to look at the larger set when the two sets differ numerically, for example, 128 vs 32 elements (Fantz & Fagan, 1975), 32 vs 16 (Xu, Spelke, & Goddard, 2005), or 16 vs 8 (Lipton & Spelke, 2003). In a typical study, investigators repeatedly presented babies with N_1 objects until looking times decreased and then presented N_2 objects. Infants' recovery of attention (suggesting they noticed a difference) to N_2 was evident only when the ratio between N_1 and N_2 was 3:2 or 2:1 (2 vs 3, Strauss & Curtis, 1981; 8 vs 4, Xu, 2003), a finding that appeared even in 21- to 144-hour-old neonates (Antell & Keating, 1983).

Research on infants' and children's ability to compare sets led to three broad generalizations. First, at any given age, the number required to discriminate two sets of large numbers (i.e., 4 or more) does not depend on an absolute difference in number. Rather, as in the Weber-Fechner psychophysical function, *the probability of discrimination is proportional to the difference in the logarithms of the numbers*, where $\ln(32)-\ln(16)=\ln(16)-\ln(8)=\ln(8)-\ln(4)>\ln(12)-\ln(8)=\ln(6)-\ln(4)$. (The term "ln" refers to the natural logarithm.) Second, the difference in logarithms required to discriminate numbers decreases with age. Thus, older infants were found to discriminate small ratios (e.g., 4:6 or 8:10) that younger infants could not (Cordes & Brannon, 2008). Third, for very small numbers (i.e., 3 or less), the probability of discrimination was uniformly high, higher than would be predicted from considering the differences in logarithms alone. Thus, discriminating 2 vs 3 is easier for infants than discriminating 4 vs 6. These three generalizations also hold for a wide range of animals' ability to compare sets (honeybees, Dacke & Srinivasan, 2008; fish, Piffer, Petrazzini, & Agrillo, 2013; pigeons, Scarf, Hayne, & Colombo, 2011; rats, Mechner, 1958; and monkeys, Brannon & Terrace, 2000) making it likely that number acuity is the product of evolution and not schooling.

The time required for older children and adults to compare sets parallels findings with infants. When prevented from counting and asked to select the larger of two sets, the time required to select the larger set is generally proportionate to the difference in logarithms, with the time required to select the larger of 4–6 being less than 6–7 (Birnbaum, 1980; Buckley & Gillman, 1974; Temple & Posner, 1998). Also, the difference in logarithms necessary for discriminating sets decreases with age. That is, a small difference in set size takes longer for younger children to discriminate than it does for older children and longer for older children than for adults (Temple & Posner, 1998). To quantify this development, a useful formalism is the "internal Weber fraction" (Halberda & Feigenson, 2008; Piazza & Izard, 2009;

Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004). For example, both schooled and unschooled adults can reliably discriminate numbers that are only within 11%–15% of each other (Halberda & Feigenson, 2008; Halberda et al., 2008; Piazza et al., 2010), which would be measured as an internal Weber fraction of 0.11–0.15. Regardless of cultural experiences, this internal Weber fraction changes greatly from infancy to adulthood (Piazza & Izard, 2009; Pica, Lemer, Izard, & Dehaene, 2004), starting from an average of 0.34 for kindergarteners, down to 0.25 for 10-year olds, and 0.15 for adults. Finally, the speed of older children's and adults' discrimination of very small numbers of objects is consistently greater than would be expected from considering the ratio in isolation. For example, the time required to judge the greater of two sets of 3 or fewer objects is uniformly low, much lower than would be expected given their ratios alone (Chi & Klahr, 1975; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008).

Unlike the early ability to compare sets of objects, the ability to compare numbers develops much later. Young children—and some adults—who reliably notice the difference between collections of 3 and 4 objects are hopelessly innumerate when asked to compare the magnitudes of numerals like "3" and "4," "30%" and "40%," or "3/10" and "2/5." Unlike the ability to compare sets of objects, these skills must be learned and show substantial variability. For example, 3-year olds from middle-income backgrounds or 4- and 5-year olds from ow-income backgrounds perform near-chance levels when asked to compare numbers between 1 and 9 (Ramani & Siegler, 2008). By kindergarten, most (but not all) children are generally accurate at comparing numerals, but their accuracy and solution times—like those of older children and adults—are subject to the same Weber-Fechner law that characterizes comparisons of nonsymbolic quantities (Dehaene, 1989; Laski & Siegler, 2007; Moyer & Landauer, 1967; Sekuler & Mierkiewicz, 1977).

The parallels between the ability of babies and other animals to compare sets and the ability of children and other humans to compare numbers suggests that once the meaning of numeric symbols are learned, numeric symbols automatically activates the brain's system for representing *magnitude*, thereby leading to the Weber-Fechner pattern in response times. In support of this idea, presentations of either sets or numerals evoke number-related activation in the intraparietal cortex of educated adults (Naccache & Dehaene, 2001). Also, after adaptation with a set of 17, 18, or 19 dots, blood oxygenation recovery is observed when the number 50 is presented, but not when the number 20 is presented. These results suggest that—at least in educated human adults—populations of neurons in parietal and prefrontal cortex are jointly activated by nonsymbolic and symbolic quantities (Naccache & Dehaene, 2001). This is a powerful result in that it provides a link between the mechanism of magnitude representation and the symbols that evoke that mechanism.

It also illustrates why all human children may possess some mathematical aptitude, but why culture is needed to fully develop that aptitude.

Estimating the Placement of Sets and Numbers on a Number Line

A still more direct method of assessing numerical magnitude judgments asks children to estimate the placement of a number on a number line (Siegler & Opfer, 2003). Like comparisons of sets and numbers, children's early number line estimates follow the Weber-Fechner law, with children—like babies, rats, and pigeons comparing sets—estimating the differences among small numbers to be greater than the differences among large numbers (as on a logarithmic ruler). With education, however, children's estimates typically follow a more linear function.

This "logarithmic-to-linear shift" was first observed on a number line task in which a symbolic number was estimated on a line flanked by another symbolic number at each end and with no intermediate anchors provided. On this task, young children's placement of numbers typically follows an approximately logarithmic function, but this logarithmic pattern changes to a linear one with age and experience, with the timing of the shift coinciding with the magnitude of the number ranges tested and children's experiences with numbers in school (Booth & Siegler, 2006; Kim & Opfer, in press; Laski & Siegler, 2007; Opfer & Martens, 2012; Opfer & Siegler, 2007; Opfer, Siegler, & Young, 2011; Opfer & Thompson, 2008; Siegler & Booth, 2004; Siegler & Opfer, 2003; Siegler, Thompson, & Opfer, 2009; Thompson & Opfer, 2008). The logarithmic-to-linear shift has since been observed across a wide range of other number line tasks, including ones in which a midpoint anchor is provided for children (Booth & Siegler, 2006; Opfer, Thompson, & Kim, 2016; Siegler & Booth, 2004), one in which there is no upper boundary marked (Kim & Opfer, in press), and ones in which the numbers are replaced by sets of dots (Opfer, Qin, & Kim, submitted for publication).

An interesting feature of number line estimation is that the magnitude of the logarithmic-to-linear shift is much greater when subjects are asked to estimate the placement of numbers instead of sets of dots. On this nonsymbolic version of the task, estimates of nonschooled adults and children are more logarithmic than would be expected from their age alone (Dehaene, Izard, Spelke, & Pica, 2008; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Sella, Berteletti, Lucangeli, & Zorzi, 2015), and the estimates of even educated adults show a substantial degree of logarithmicity (Anobile, Cicchini, & Burr, 2012; Dehaene et al., 2008). These findings suggest that providing linear number line estimates—where the distance between any two successive estimates (e.g., 5 and 6 vs 155 and 156) is equal—depends on schooling and culture, where children gain proficiency with symbolic quantities.

Approximately Adding Sets and Numbers

A mental number line makes basic addition and subtraction trivial: traveling four spaces forward from four registers the sum of four and four, traveling four spaces back from eight registers the difference between eight and four, and so on. Thus, if infants encode the approximate numerical value of a set and possess something akin to a mental number line (for at least nonsymbolic numbers), they should be able to register sums and differences of numeric quantities (at least approximately).

The hypothesis that infants could approximately add sets led Wynn (1992) to conduct a series of surprising experiments on infants' arithmetic abilities. In one condition, Wynn tested babies' ability to compute $1+1$: she recorded infants' looking times as they watched one object appear to be placed behind an opaque screen and then another object added to it behind the screen. When the screen dropped, seeming to reveal the arithmetically impossible event $1+1=1$, infants looked longer than when the screen dropped and revealed an outcome consistent with the arithmetically realistic event $1+1=2$. Additionally, infants' ability to approximately add sets is not limited to numbers less than 4 (McCrink & Wynn, 2004). Consistent with Wynn's interpretation that infants were surprised by the numerical outcome, when 9-month-old babies were confronted with arithmetic transformations of large sets (e.g., $5+5=10$ vs $5+5=5$ or $10-5=5$ vs $10-5=10$), babies also looked longer at the arithmetically impossible events than the arithmetically possible ones. Thus, it seems that nonsymbolic, approximate arithmetic may not require any schooling at all.

How are babies able to add and subtract sets of objects (at least approximately)? One possibility is that nonsymbolic, approximate addition relies on the same logarithmically compressed number line representation that we reviewed above. In this view, adding $n2$ to $n1$ involves traveling $n2$ spaces forward from $n1$ on the mental number line. If this is right, then babies would actually arrive at the position $\log(n1)+\log(n2)$, which would be a considerable overestimation of the actual result. From this perspective, babies would have no doubt found $5+5=5$ surprising because they would have experienced it as $\log(5)+\log(5)$ and thus expected to see $\log(25)$! Conversely, subtraction through traversing a logarithmically spaced number line would yield a considerable underestimation of the actual results.

The idea that the approximate addition relies on a logarithmically spaced number line is consistent with the mistakes that babies make when adding sets. Specifically, babies' expectations do overestimate the results of additive operations and underestimate the results of subtractive operations (McCrink & Wynn, 2009). Thus, when babies were initially shown a sequence of events equivalent to $6+4$, they looked significantly longer when the raised screen revealed 5 objects than when it revealed 20 objects. Similarly, when babies were shown a sequence equivalent to $14-4$, they looked significantly longer when the raised screen revealed 20 objects than when it revealed 5. In the same study, the infants looked equally long at 20 objects as 10 at the end of the first example and looked as long at 5 objects as at 10 at the end of the second example.

Approximately adding sets of objects appears quite similar in infants, young children, and adults (Barth, Beckmann, & Spelke, 2008; Barth et al., 2006; McCrink, Dehaene, & Dehaene-Lambertz, 2007; Park, Bermudez, Roberts, & Brannon, 2016; Pica et al., 2004). In a representative study, McCrink et al. (2007) showed adults several hundred videos of two successive sets of dots and asked them to approximate their sum or difference by choosing one of two sets of dots. As with infants, adults almost always overshot the correct outcomes on addition problems, whereas they almost always undershot the correct outcomes on the subtraction problems. Researchers found that from adults' modal response, the distribution of other responses tapered off as a function of the ratio of the true and alternative quantities, just as would be predicted by Weber-Fechner's law. To illustrate the magnitude of this error, the presented subtraction problem of $32-16=8$ was judged to be correct approximately 60% of the time, which is quite a radical departure from moving along a linearly scaled mental number line.

Before formal arithmetic training, children's ability to add numbers is very similar to adults' ability to add sets of objects. One source of evidence comes from Gilmore, McCarthy, and Spelke's (2007) study of preschoolers' estimates of answers to arithmetic problems that they had not yet encountered in school. The investigators presented 5- and 6-year olds with problems such as "Sarah has 21 candies; she gets 30 more; John has 34 candies—who has more?" To ensure that the preschoolers understood the symbols that were being used, the problems were simultaneously presented both orally—as spoken numerals—and in writing, as Arabic numerals. Despite the fact that the preschoolers had received no training with numbers of that size, they spontaneously performed better than chance. Performance was still approximate, however, and depended on the ratio of the two sums that the children were choosing between, a signature of the Weber-Fechner law.

Connections Between Numerical Magnitude Judgments and Math Proficiency

In the previous section, we reviewed evidence that three types of numerical magnitude judgments—comparison, estimation, and approximate addition—bear striking psychophysical similarities and may rely on highly overlapping cortical resources. In this section, we examine the potential importance of this cognitive feature on math proficiency by reviewing the literature that examines relations between numerical magnitude judgments, math proficiency, and aptitude for math learning.

Comparing Sets and Numbers

Given that arithmetic proficiency is usually measured by participants' speed and accuracy with symbolic problems, it was a real surprise to many investigators to learn that individual differences in the ability to compare sets of objects (thereby tapping the ANS) could (retrospectively) predict individual differences in mathematics ability (Halberda et al., 2008). The finding that number acuity predicts math skill, however, has now been tested repeatedly in preschoolers and kindergartners (Bonny & Lourenco, 2013; Libertus, Feigenson, & Halberda, 2011, 2013), older children (Holloway & Ansari, 2009; Mundy & Gilmore, 2009), and adults (Castronovo & Göbel, 2012; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Lyons & Beilock, 2011). The breadth of these studies allowed Fazio, Bailey, Thompson, and Siegler (2014) to conduct a metaanalysis of 19 studies examining the relation between number acuity and math proficiency. Their general finding was that there is a small but reliable connection between number acuity and math proficiency, with the ability to compare sets accounting for about 4% of the variation in math scores. One way to make sense of these findings is that math ability partly develops as an elaborated, conceptual understanding of numerosity—the property shared by sets of *N* objects or *N* tones or *N* events (Butterworth, 2005), which is often impaired in a subpopulation of individuals with dyscalculia (Butterworth, 2005).

Children's ability to compare symbolic numbers also has a positive relation to their arithmetic proficiency (Castronovo & Göbel, 2012; De Smedt, Verschaffel, & Ghesquière, 2009; Fazio et al., 2014; Holloway & Ansari, 2009; Vanbinst, Ghesquière, & De Smedt, 2012), with the magnitude of the correlation between number comparison and math scores typically being much larger than the correlation between ability to compare sets and math scores (for a metaanalysis of 45 studies, see Schneider et al., 2016). For example, in the study by Fazio et al. (2014), the speed of number comparison explained 33% of the variation in standardized math scores (compared with 4% from their metaanalysis on number acuity and math proficiency). Additionally, individual differences in the magnitude of the numerical distance effect explain about 11% of the variation in math fluency (Holloway & Ansari, 2009). Thus, if Chinese children have greater facility with comparing numbers than U.S. children, one might expect them to also become more fluent in math and score better on standardized tests.

Estimating the Placement of Sets and Numbers on a Number Line

As with magnitude comparison, estimating the placement of sets and numbers on a number line shows a strong and consistent relation to math skills, with the magnitude of the relation again being larger for symbolic than nonsymbolic number line estimation. For example, in the Fazio et al. (2014) study, accuracy of nonsymbolic whole number line estimation accounted for about

10% of the variation in standardized math scores. In our own studies (Opfer, Qin, et al., submitted for publication), the logarithmicity of nonsymbolic number line estimation explains about the same proportion of variance in addition and subtraction scores after controlling for all other variables in the model (such as age) and experiment-wise error.

The relation between ability to place numbers on a number line and math proficiency is also a very consistent finding. Individual differences in the accuracy (Fazio et al., 2014; LeFevre et al., 2013; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Muldoon, Towse, Simms, Perra, & Menzies, 2013; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013; Schneider, Grabner, & Paetsch, 2009; Siegler & Ramani, 2008, 2009) and logarithmicity of estimates of a number on a number line (Kim & Opfer, in press; Opfer, Qin, et al., submitted for publication) correlates positively and strongly with speeded arithmetic and standardized math scores. In studies that compare placement of numbers with placement of sets, the proportion of variance in math scores explained by symbolic number line estimates is typically around 30% (Fazio et al., 2014; Kim & Opfer, in press; Opfer, Qin, et al., submitted for publication) and roughly twice that of nonsymbolic number line estimation.

Probably, the best evidence that linearity of number line estimates improves math proficiency comes from experimental studies that have studied the effects of number line training (Booth & Siegler, 2008; Siegler & Ramani, 2009). These studies find that presenting randomly chosen school aged children with accurate number line representations of the magnitudes within arithmetic problems improves children's learning of the answers to the problems (Booth & Siegler, 2008). Similarly, playing a linear board game with randomly chosen preschoolers improves their ability to learn answers to arithmetic problems (Siegler & Ramani, 2009). Thus, rather just linking number line estimate accuracy to arithmetic performance, these studies suggest that accurate representations of numerical magnitude make learning math easier for children. This is an important finding because it suggests that Chinese children might gain more from their experiences in school than American children if they have more linear numerical magnitude representations (Fig. 2).

Cross-National Differences in Numerical Magnitude Judgments

Given the early emergence of competence in nonsymbolic numerical magnitude judgments and the relatively late emergence of competence in symbolic numerical magnitude judgments, it seems reasonable to expect a certain pattern to cross-national differences in numerical magnitude judgments. Specifically, because proficiency with symbols depends on experiences in early childhood, cross-national differences with symbolic numbers would be expected to be larger than cross-national differences in proficiency with nonsymbolic numbers. This hypothesis has been tested indirectly with Western

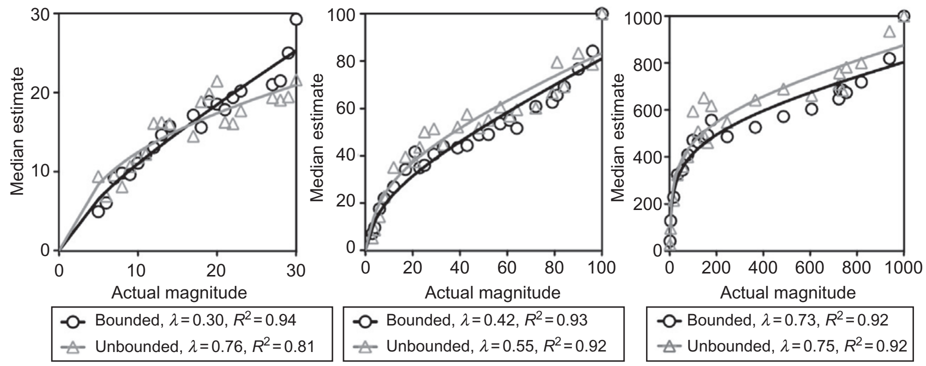

FIG. 2 Median estimates for American kindergartners' estimates on 0–30 number lines, first graders' estimates on 0–100 number lines, and second graders' estimations on 0–1000 number lines. On unbounded versions of the number line task, estimates were generally more logarithmic than on the bounded versions studied in Siegler and Opfer (2003).

and East Asian children in a number of studies (Rodic et al., 2015; Siegler & Mu, 2008; Xu, Chen, Pan, & Li, 2013; Zhou et al., 2007). These studies have examined cross-national differences in comparing sets and numbers and estimating the placement of sets and numbers on a number line. However, no studies have yet examined cross-national differences in approximate addition.

Comparing Sets and Numbers

When comparing two sets of dots for the greater quantity, Chinese children perform somewhat better than their peers in Western countries. For example, Rodic et al. (2015) tested ethnic and national Chinese children and their Western peers on a large test battery, including comparing sets of dots. Chinese children's judgments were reliably more accurate than children growing up in the United Kingdom, but the judgments were typically made just as quickly. This is important because the standard positive relation between nonsymbolic number comparison and math performance typically uses a composite measure of speed and accuracy, suggesting that this component of numerical magnitude judgments may not mediate the effect of nation on math proficiency.

When comparing two numbers for the greater quantity, Chinese children also performed better than their Western peers, and this difference was larger and better supported by a range of evidence than evidence regarding set size comparison (Rodic et al., 2015). More strikingly, Chinese children were found to automatically process symbolic magnitudes as early as 5 years old (Zhou et al., 2007). In their study, Zhou et al. (2007) asked children to estimate the *physical* size of numbers but ignore their magnitude information. In this Stroop-like task, 5-year-old Chinese children showed a significant size congruity effect, suggesting magnitude information of numbers interfered with their physical size comparison. This automatic processing of symbolic magnitude ability is striking because it emerges at a much older age for children in Western nations (Berch, Foley, Hill, & Ryan, 1999; Girelli, Lucangeli, & Butterworth, 2000; Rubinsten, Henik, Berger, & Shahar-Shalev, 2002).

Estimating the Placement of Sets and Numbers on a Number-Line

When estimating the placement of sets on a number line, Chinese kindergartners' estimates were also found to be more accurate and linear than their Western peers. However, both groups showed a similar developmental pattern (a logarithmic-to-linear shift), and both showed a linear pattern of estimates on 0–10 number lines and a logarithmic pattern on 0–100 number lines (Sasanguie et al., 2012; Sella et al., 2015; Xu et al., 2013). Given the small but reliable relation between nonsymbolic number line estimation and math proficiency, generally, these results should lead us to expect that this component of numerical magnitude judgments will mediate a small but reliable effect of nation on math proficiency.

When estimating the placement of numbers on number lines, a somewhat similar pattern emerges. Specifically, Chinese children have shown the same logarithmic-to-linear shift as their Western peers, but they typically show an earlier onset of attaining linear representations (Siegler & Mu, 2008; Xu et al., 2013). For example, Xu et al. (2013) found the estimates of Chinese preschoolers (3- to 7-year olds) were more accurate and linear on the 0–10, 0–100, and 0–1000 number lines than their Western peers. Moreover, the linearity indexes of estimates for Chinese *kindergartners* were similar to those for American *first and second graders* (Booth & Siegler, 2006; Opfer & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003, Xu et al., 2013). Given the previously observed association between symbolic number line estimation and math proficiency and math learning, this is an important finding. It suggests that a given hour of math instruction in China might have a larger effect on children than the same hour of math instruction in the United States because children in the former group have a linear representation of numerical magnitude—and thus a stronger foundation for symbolic magnitude representations than their Western peers.

DO NUMERICAL MAGNITUDE JUDGMENTS MEDIATE NATION EFFECTS ON MATH PROFICIENCY?

In the previous sections, we have reviewed evidence that (1) East Asian children typically show greater math proficiency than their American peers, (2) numerical magnitude judgments (particularly symbolic ones) are generally correlated with greater math proficiency, and (3) East Asian children typically show more accurate numerical magnitude judgments (particularly symbolic ones) than their American peers. Together, these three lines of work suggest a simple model in which the effect of nation on math proficiency—especially early, but possibly all math proficiency (see below)—is mediated by the quality of numerical magnitude judgments (Fig. 3).

A test of this model has recently been conducted by Opfer, Kim, Siegler, Fazio, and Zhou (submitted for publication). Although these results have not been published yet, we can preview a narrative description of their findings. At present, these preliminary results provide the only test of our cognitive theory of nation effects on mathematics proficiency.

In their study, Opfer et al. examined kindergarteners recruited at the end of the school year from four schools in Beijing, China, and Columbus, OH. All children completed a mathematics assessment (Geary et al., 1996), a subtest of the WISC-III with a numeric response code (Stevenson et al., 1985), and six numeric magnitude judgment tasks. The numeric magnitude tasks varied in the format of numeric information (symbolic numerals vs dot arrays) and type of magnitude judgment (comparison, approximate addition, and number line estimation). Time for each task was roughly equivalent; each took less than 5 min.

Consistent with previous reports (Geary et al., 1996), Opfer et al. found that arithmetic scores of Chinese kindergartners were roughly three times higher than those of U.S. kindergartners. Each element of the arithmetic score differed by nationality: U.S. kindergartners provided fewer accurate answers than Chinese kindergartners, and U.S. children provided more wrong answers than Chinese children. Finally, when wrong answers were given, the difference between the child's answer and the correct one was significantly smaller among Chinese than U.S. children.

What might account for these differences in math proficiency? Opfer et al. first reexamined the essentialist hypothesis that East Asian kindergartners are simply smarter than American ones. To do this, they used a coding task (WISC-III; also Digit-Symbol Substitution task, WAIS-III) that is widely used to measure children's and adults' performance IQ (Ryan & Lopez, 2001; Wechsler, 1991). The task is thought to index the overall speed of information processing (Lezak, Howieson, Loring, Hannay, & Fischer, 2004; Stevenson

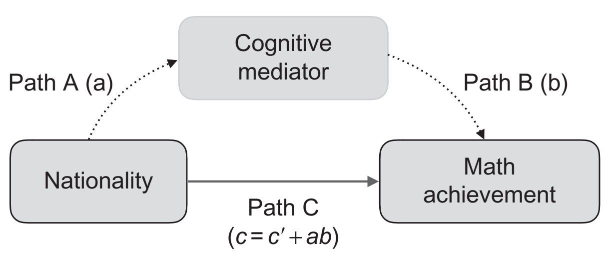

FIG. 3 Mediation model of the effects of nationality on math proficiency. The total effects of nationality (c) are divided into two effects: the direct effect (c') and the indirect effect via a cognitive mediator (ab).

et al., 1985) and correlates more highly with arithmetic performance than any other IQ subscale (Osborne & Lindsey, 1967). Additionally, because the test relies on a numeric response code, it provides a simple test of the idea that Chinese children are simply more familiar with numbers overall.

Consistent with Stevenson et al. (1985) findings with first graders (Geary et al., 1997; Stevenson et al., 1985), Opfer et al. found that average scores on this coding task did not differ between Chinese and U.S. kindergarteners, suggesting their samples were equivalent to previous studies and national norms. Additionally, variability in scores did not differ between the two groups, suggesting the samples were equivalently homogenous. Although IQ scores did correlate highly with mathematics scores, the lack of difference between the American and Chinese children in IQ meant that nationality could not be having an effect on math proficiency via IQ.

Comparing Sets and Numbers: U.S. Vs Chinese children

Opfer et al. next examined the effect of nationality on three measures of nonsymbolic and symbolic number comparison. These measures included overall accuracy (Lourenco, Bonny, Fernandez, & Rao, 2012; Mazzocco, Feigenson, & Halberda, 2011), Weber fraction (*w*; Halberda et al., 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011), and numeric distance effect (Holloway & Ansari, 2009; Price, Palmer, Battista, & Ansari, 2012). Across all three indexes of nonsymbolic number comparison, a remarkably consistent pattern emerged. Overall accuracy scores, the average Weber fraction, and the numeric distance effect were similar for Chinese and U.S. children. As we have argued, comparing sets is likely to be a universal cognitive ability, and the similarity of Chinese and U.S. children should not be surprising. Also, like IQ, accuracy of nonsymbolic number comparison was somewhat correlated with mathematics scores, but the lack of difference between the Chinese and U.S. children's ability to compare sets meant that numerical magnitude judgments could not have mediated the effect of nation on mathematics proficiency.

Chinese and U.S. children's ability to compare numbers was examined similarly. Unlike set comparison, number comparison differed markedly between the U.S. and Chinese children. In comparing numbers, U.S. children provided less accurate answers and required a larger numeric difference to accurately discriminate two numbers. Additionally, because accuracy of symbolic number comparison was correlated with mathematics scores and because U.S. children provided less accurate symbolic number comparisons, symbolic number comparison could partially mediate the effect of nationality on arithmetic performance. Consistent with this conjecture, Opfer et al. found that nearly half the effect of nationality on arithmetic performance was an indirect effect that occurred via symbolic number comparison.

Estimating the Placement of Sets and Numbers on a Number Line: U.S. Vs Chinese Children

Children's ability to estimate the placement of sets and numbers on a number line was also evaluated in terms of overall accuracy and by assessing the degree of logarithmicity in the responses. Chinese children's estimates where a set of dots would fall on the number line were surprisingly more accurate than U.S. children's estimates. Additionally, the estimates of Chinese children were less logarithmic than estimates of U.S. children. Because erroneous nonsymbolic number line estimation was somewhat negatively correlated with mathematics scores and because U.S. children provided less accurate nonsymbolic number line estimates than Chinese children, nonsymbolic number line estimation would be expected to partially mediate the effect of nationality on arithmetic performance. Consistent with this conjecture, Opfer et al. found that almost half of the effect of nationality on arithmetic performance was an indirect effect that occurred via nonsymbolic number line estimation.

Chinese and U.S. children's ability to place numbers on number lines was evaluated and compared similarly (see also Chapter 4 by Okamoto for a related discussion of cross-cultural differences in number line estimation). Consistent with previous reports, accuracy of Chinese children's symbolic number line estimates was greater than accuracy of U.S. children's estimates, and the estimates of Chinese children were less logarithmic than the estimates of U.S. children (Siegler & Mu, 2008). Because erroneous symbolic number line estimation was negatively correlated with mathematics scores and because U.S. children were less accurate at symbolic number line estimation, it was expected that symbolic number line estimation ability might mediate the effect of nationality on arithmetic performance. Consistent with this conjecture, Opfer et al. found that nearly half of the effect of nationality on arithmetic performance was an indirect effect that occurred via symbolic number line estimation.

Approximately Adding Sets and Numbers: U.S. Vs Chinese Children

Children's ability to approximately add sets of dots and numbers (e.g., $a+b>$ or $<c$?) was evaluated in both the U.S. and Chinese samples. When approximately adding sets of dots, Chinese and U.S. children were remarkably similar in accuracy and the average Weber fraction. In both groups, accuracy improved as a function of the difference in the logarithms of the numbers compared, and the magnitude of this distance effect was similar as well. Thus, although accuracy of nonsymbolic approximate addition was somewhat correlated with mathematics scores, there was very little indirect effect of nation on math through this route.

Children's ability to approximately add numbers, somewhat unsurprisingly, did differ between Chinese and U.S. children. Overall accuracy scores were higher for Chinese than for U.S. children, and the average Weber fraction was also lower for Chinese than for U.S. children. Because accuracy of symbolic approximate addition was correlated with mathematics scores and because U.S. children were less accurate at symbolic approximate addition, symbolic approximate addition would be expected to partially mediate the effect of nationality on arithmetic performance. Consistent with this conjecture, Opfer et al. found that a little more than half the effect of nationality on arithmetic performance was an indirect effect that occurred via symbolic approximate addition.

Cognitive Mechanisms of Nation Effects on Mathematics Proficiency

In the previous three sections, independent mediation analyses indicated that the effect of nation on mathematics proficiency was partially mediated by a number of symbolic numerical magnitude judgments. These cognitive abilities, however, were largely correlated with one another. As a result, the total indirect effect of nation on mathematics via these cognitive abilities cannot be a sum of these partial effects (e.g., more than two different mediators accounted for roughly half the indirect effect, which is mathematically impossible). Additionally, each independent mediation analysis is associated with the risk of experiment-wise error. To deal with these issues, Opfer et al. conducted a multiple mediation analysis to measure how much the total effect of nationality on mathematics occurred via an indirect effect on numerical magnitude judgments and which of these were the mechanisms by which nation affected mathematics proficiency (Fig. 4). Overall, they found that only a very small proportion of the total effects of nation on arithmetic performance occurred directly via nationality itself, whereas almost all of the effect occurred through the numerical magnitude judgment mediators. Put another way, if the total effect of nationality on mathematics was the size of the "learning gap," almost all of this learning gap was reduced by accuracy of numerical magnitude judgments.

Although the multiple mediators accounted for almost all of the effects of nation on arithmetic scores, contribution of individual mediators to the mediated effects varied considerably. Generally, indirect effects of nationality through the three nonsymbolic numeric judgments were not significant, whereas the symbolic numerical judgments mediated effects of nationality significantly. In contrast, effects of nationality on arithmetic scores were significantly mediated by all symbolic numeric judgment tasks. These results suggest that the effects of nation on mathematics proficiency are mediated by symbolic numeric performance significantly more than by nonsymbolic numeric judgments. Additionally, of these putative mediators, only number line estimation showed a direct effect of nation controlling for differences in math achievement (Fig. 5).

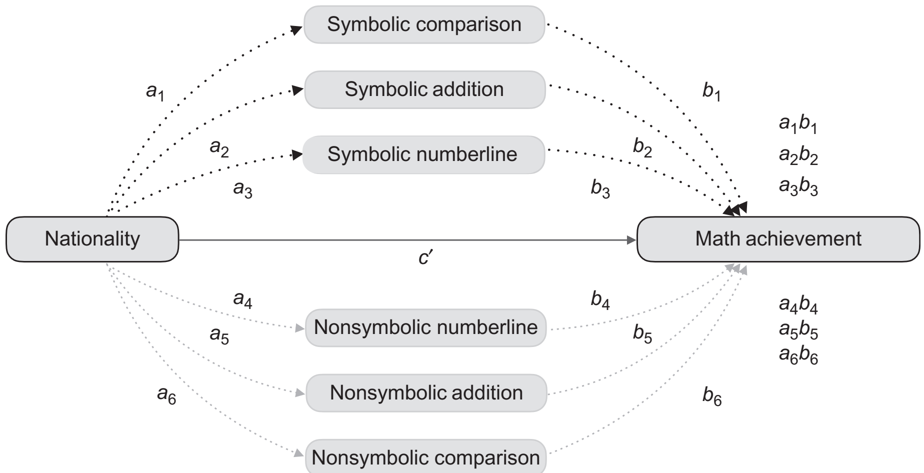

FIG. 4 Multiple mediation analysis. Nearly all of the total effects of nationality occurred via the six tasks (indirect effects, $a_1b_1+a_2b_2+a_3b_3+a_4b_4+a_5b_5+a_6b_6$), whereas nationality itself had no remaining proportion of the total effects (direct effect, c'). The indirect effects through individual symbolic number abilities were significant.

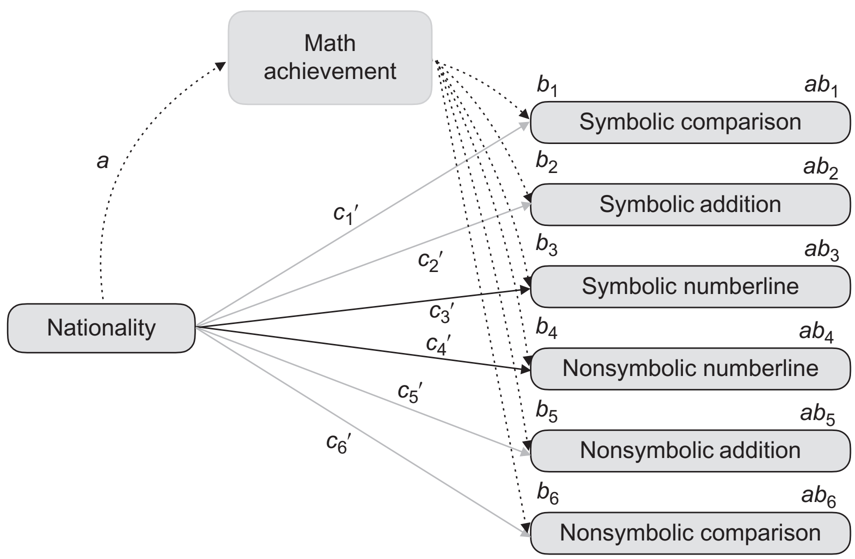

FIG. 5 Mediation analyses with math achievement as a mediator. Only number line estimation presented direct effects of nationality controlling for math achievement (c_3' and c_4').

CONCLUSIONS

For many decades, American children have lagged behind their East Asian peers in math proficiency. Efforts to explain this "learning gap" have traditionally looked for potential differences in input, and these efforts have not come up short: East Asian children have been found to have more opportunities to spend time on math, have math instructors that make their time better spent, and (maybe for these reasons) are more willing to spend time learning math. Emulating the East Asian model seems promising. However, reforms to lengthen the American school day or school year, to improve the quality of American teachers and instruction, and to try to convince American children that math is worthwhile have not narrowed the gap between East Asian and American students' math proficiency in the 25 years since the learning gap first came to broad attention.

In this chapter, we proposed a very different approach to the problem of the "learning gap." Specifically, we hypothesized that the cause of the learning gap may not be simply a matter of differences in schooling, but a *cognitive* difference in representations of numerical magnitude. This cognitive difference is superficially similar to IQ in that representations of numerical magnitude also increases learning from a given unit of input and emerges before formal schooling. Unlike IQ, however, these differences in numerical magnitude representations almost certainly come from differences in informal teaching at home. Thus, the learning gap may result from a kind of aptitude difference (i.e., a greater ability to learn) between East Asian and American children, but this rests on domain-specific knowledge—namely, representations of numeric magnitude—rather than a general cognitive ability, and these representations of numeric magnitude are known to be malleable (Opfer & Siegler, 2007; Opfer & Thompson, 2008; Ramani & Siegler, 2008; Siegler & Ramani, 2008; Thompson & Opfer, 2008).

High-quality numerical magnitude representations—assessed by children's ability to compare, estimate, and approximately add sets and numbers—generally correlate positively with math proficiency and have been shown to increase the ability of children to learn arithmetic. Additionally, evidence suggests that these numerical magnitude representations also differ between East Asian and U.S. children. These facts led us to hypothesize that American children with high-quality numerical magnitude representations might almost be as proficient at mathematics as East Asian children.

Our preliminary results from a study of Beijing and Columbus students provided much support for our cognitive theory. Specifically, we found that accurate numerical magnitude representations were generally associated with greater math proficiency, and we found that accuracy of numerical magnitude representations were generally higher among Chinese students than American ones. Mediation analyses also showed that nearly the entire effect of nation on math proficiency came via the indirect route of improving numerical magnitude representations. Thus, when American children had numerical magnitude representations that were equal to those of their Chinese peers, they were also just as proficient at mathematics—despite being in schools with American-length school years and school days, American-trained teachers, and presumably the same American attitudes toward mathematics.

These results are broadly consistent with several lines of previous work. First, after controlling for Americans' ethnicity (including Asian and European heritage) and a broad array of general cognitive abilities (such as working memory), symbolic number knowledge in first grade has also been found to predict performance on economically relevant math skills in adolescence (Geary, Hoard, Nugent, & Bailey, 2013). This relation is not unique to early math learning: adolescents who show better symbolic number knowledge for fractions also do better in high school math classes, after controlling for a very wide range of other predictors (Siegler et al., 2012). Thus, rather than symbolic number knowledge simply providing a foundation to early math proficiency, it appears to play a supporting role for proficiency throughout development.

Results from our study of East Asian and American students closely parallel findings about socioeconomic differences in math ability. Like American students, children low in SES also show deficits in numerical magnitude representations and math proficiency. Moreover, experimental lessons that have intervened specifically on low SES children's numerical magnitude representations were also able to eliminate the differences between low and high SES children's math proficiency. Combined with our cross-national findings, previous efforts suggest that much of the learning gap between U.S. and American children could be closed if educational interventions focused on American children's relatively poor-quality numerical magnitude representations.

ACKNOWLEDGMENTS

This chapter was supported in part by IES grant R305A160295. We gratefully acknowledge the help and support of Bob Siegler, Xinlin Zhou, Lisa Fazio, and the Siegler Center for Innovation in Learning. We would also like to thank our Ohio State research assistant Yiwan Wang and the research assistants at Beijing Normal University.

REFERENCES

Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition*, *122*(3), 454–459.

Antell, S. E., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695–701.

Bailey, D. H., Zhou, X., Zhang, Y., Cui, J., Fuchs, L. S., Jordan, N. C., et al. (2015). Development of fraction concepts and procedures in US and Chinese children. *Journal of Experimental Child Psychology*, *129*, 68–83.

Barth, H., Beckmann, L., & Spelke, E. S. (2008). Nonsymbolic, approximate arithmetic in children: Abstract addition prior to instruction. *Developmental Psychology*, *44*(5), 1466.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(39), 14116–14121.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*(3), 199–222.

Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: Developmental changes in number processing and mental representation. *Journal of Experimental Child Psychology*, *74*, 286–308.

Birnbaum, M. H. (1980). Comparison of two theories of "ratio" and "difference" judgments. *Journal of Experimental Psychology: General*, *109*(3), 304–319.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, *79*, 1016–1031.

Brannon, E. M., & Terrace, H. S. (2000). Representation of the numerosities 1–9 by rhesus macaques (Macaca mulatta). *Journal of Experimental Psychology: Animal Behavior Processes*, *26*(1), 31.

Buckley, P. B., & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology*, *103*(6), 1131.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry*, *46*(1), 3–18.

Carroll, J. B. (1989). The Carroll Model: A 25-year retrospective and prospective view. *Educational Researcher*, *18*, 26–31.

Case, R. (1975). Social class differences in intellectual development: A Neo-Piagetian investigation. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *7*(3), 244.

Case, R., & Griffin, S. (1990). Child cognitive development: The role of central conceptual structures in the development of scientific and social thought. *Advances in Psychology*, *64*, 193–230.

Case, R., & Sandieson, R. (1987). General developmental constraints on the acquisition of special procedures (and vice versa). In *Paper presented at the annual meetings of the American Educational Research Association, Baltimore, April*.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS ONE*, *7*(4), e33832.

Chen, C., & Uttal, D. H. (1988). Cultural values, parents' beliefs, and children's achievement in the United States and China. *Human Development*, *31*, 351–358.

Chi, M. T., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, *19*(3), 434–439.

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*(6), 803–808.

Dacke, M., & Srinivasan, M. V. (2008). Two odometers in honeybees? *Journal of Experimental Biology*, *211*(20), 3281–3286.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*(4), 469–479.

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Attention, Perception, & Psychophysics*, *45*(6), 557–566.

Dehaene, S. (1997). *The number sense*. New York, NY: Oxford University Press.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, *320*(5880), 1217–1220.

Fantz, R. L., & Fagan, J. F., III. (1975). Visual attention to size and number of pattern details by term and preterm infants during the first six months. *Child Development*, *46*, 3–18.

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, *123*, 53–72.

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist*, *50*(1), 24.

Geary, D. C., Berch, D. B., Ochsendorf, R., & Mann Koepke, K. (Eds.), (2017). *Acquiring complex arithmetic skills and higher-order mathematical concepts*. San Diego, CA: Elsevier/Academic Press.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development*, *67*, 2022–2044.

Geary, D. C., Hamson, C. O., Chen, G. P., Liu, F., Hoard, M. K., & Salthouse, T. A. (1997). Computational and reasoning abilities in arithmetic: Cross-generational change in China and the United States. *Psychonomic Bulletin & Review*, *4*, 425–430.

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, H. D. (2013). Adolescents' functional numeracy is predicted by their school entry number system knowledge. *PLoS ONE*, *8*(1), e54651.

Geary, D. C., & vanMarle, K. (2016). Young children's core symbolic and non-symbolic quantitative knowledge in the prediction of later mathematics achievement. *Developmental Psychology*, *52*, 2130–2144.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*(7144), 589–591.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology*, *76*, 104–122.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of US fourth- and eighth-grade students in international context (NCES 2009–001 revised)*. Washington, DC: National Center for Education Statistics.

Griffin, S., Case, R., & Sandieson, R. (1992). *Synchrony and asynchrony in the acquisition of children's everyday mathematical knowledge*. *The mind's staircase: Exploring the conceptual underpinnings of children's thought and knowledge*. Hillsdale, NY: Lawrence Erlbaum Associates, Inc. pp. 75–97.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25–50). Cambridge, MA: The MIT Press/Bradford Books.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(28), 11116–11120.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668.

Hanushek, E. A., & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, *46*, 607–668.

Hofstadter, D. R. (1982). Number numbness, or why innumeracy may be just as dangerous as illiteracy. *Scientific American*, *246*(5), 20–34.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 17–29.

Husén, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vols. 1 & 2). New York, NY: Wiley.

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin & Review*, *18*(6), 1222–1229.

Jordan, N. C., Huttenlocher, J., & Levine, S. C. (1992). Differential calculation abilities in young children from middle-and low-income families. *Developmental Psychology*, *28*(4), 644.

Kim, D., & Opfer, J. (2017). A unified framework for bounded and unbounded numerical estimation. *Developmental Psychology, 53* (6), 1088.

Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, *78*(6), 1723–1743.

LeFevre, J. A., Lira, C. J., Sowinski, C., Cankaya, O., Kamawar, D., & Skwarchuk, S. L. (2013). Charting the role of the number line in mathematical development. *Frontiers in Psychology*, *4*, 641.

Lezak, M. D., Howieson, D. B., Loring, D. W., Hannay, J., & Fischer, J. (2004). *Neuropsychological assessment*. New York, NY: Oxford University Press.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*, 1292–1300.

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Numerical approximation abilities correlate with and predict informal but not formal mathematics abilities. *Journal of Experimental Child Psychology*, *116*(4), 829–838.

Lipton, J. S., & Spelke, E. S. (2003). Origins of number sense: Large-number discrimination in human infants. *Psychological Science*, *14*(5), 396–401.

Lo, J. J., & Luo, F. (2012). Prospective elementary teachers' knowledge of fraction division. *Journal of Mathematics Teacher Education*, *15*(6), 481–500.

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(46), 18737–18742.

Lynn, R. (1982). IQ in Japan and the United States shows a growing disparity. *Nature*, *297*(5863), 222–223.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*, 256–261.

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, *17*(5), 714–726.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mankiw, N. G., & Weinzierl, M. (2010). The optimal taxation of height: A case study of utilitarian income redistribution. *American Economic Journal: Economic Policy*, *2*(1), 155–176.
Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS ONE*, *6*(9), e23749.
McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776–781.
McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology*, *103*(4), 400–408.
McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: Operational momentum in nonsymbolic arithmetic. *Attention, Perception, & Psychophysics*, *69*(8), 1324–1333.
Mechner, F. (1958). Probability relations within response sequences under ratio reinforcement. *Journal of the Experimental Analysis of Behavior*, *1*(2), 109–121.
Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520.
Muldoon, K., Towse, J., Simms, V., Perra, O., & Menzies, V. (2013). A longitudinal analysis of estimation, counting skills, and mathematical ability across the first school year. *Developmental Psychology*, *49*(2), 250.
Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*(4), 490–502.
Murnane, R. J., Willett, J. B., & Levy, F. (1995). *The growing importance of cognitive skills in wage determination (No. w5076)*. Cambridge, MA: National Bureau of Economic Research.
Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*(10), 966–974.
National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, D.C.: US Department of Education.
National Research Council (NRC). (2005). How students learn: History, mathematics, and science in the classroom. Committee on How People Learn. A targeted report for teachers. In M. S. Donovan & J. D. Bransford (Eds.), Washington, DC: National Academics Press.
OECD. (2010). *PISA 2009 results: Executive summary*. Paris: OECD.
Opfer, J. E., Kim, D., Siegler, R. S., Fazio, L. K., & Zhou, X. (submitted for publication). Cognitive mechanisms of nation effects on mathematics proficiency.
Opfer, J. E., & Martens, M. A. (2012). Learning without representational change: Development of numerical estimation in individuals with Williams syndrome. *Developmental Science*, *15*(6), 863–875.
Opfer, J. E., Qin, J., & Kim, D. (submitted for publication). Varieties of numerical estimation: A unified developmental framework.
Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*(3), 169–195.
Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science*, *14*(5), 1194–1204.
Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development*, *79*(3), 788–804.
Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition*, *149*, 11–17.

Osborne, R. T., & Lindsey, J. M. (1967). A longitudinal investigation of change in the factorial composition of intelligence with age in young school children. *The Journal of Genetic Psychology*, *110*(1), 49–58.

Park, J., Bermudez, V., Roberts, R. C., & Brannon, E. M. (2016). Non-symbolic approximate arithmetic training improves math performance in preschoolers. *Journal of Experimental Child Psychology*, *152*, 278–293.

Piazza, M., & Izard, V. (2009). How humans count: Numerosity and the parietal cortex. *The Neuroscientist*, *15*(3), 261–273.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–555.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503.

Piffer, L., Petrazzini, M. E. M., & Agrillo, C. (2013). Large number discrimination in newborn fish. *PLoS ONE*, *8*(4), e62466.

Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica*, *140*(1), 50–57.

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*(2), 375–394.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614.

Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, *47*(3), 189–203.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science*, *316*, 1128–1129.

Rodic, M., Zhou, X., Tikhomirova, T., Wei, W., Malykh, S., Ismatulina, V., et al. (2015). Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science*, *18*, 165–174.

Rose, H., & Betts, J. R. (2004). The effect of high school courses on earnings. *Review of Economics and Statistics*, *86*(2), 497–513.

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology*, *81*, 74–92.

Rushton, J. P. (1992). Cranial capacity related to sex, rank, and race in a stratified random sample of 6,325 U.S. military personnel. *Intelligence*, *16*, 401–413.

Ryan, J. J., & Lopez, S. J. (2001). Wechsler adult intelligence scale-III. *Understanding psychological assessment*. New York City, NY: Springer. pp. 19–42.

Sadler, P. M., & Tai, R. H. (2007). The two high-school pillars supporting college science. *Science*, *317*, 457–458.

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, *30*(2), 344–357.

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*(3), 418–431.

Saxe, G. B., Guberman, S. R., & Gearhart, M. (1989). *Social processes in early number development. Monographs of the society for research in child development:* Vol. 52 (serial no. 216).

Scarf, D., Hayne, H., & Colombo, M. (2011). Pigeons on par with primates in numerical competence. *Science*, *334*(6063), 1664.

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., et al. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), 1–16.

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology*, *101*(2), 359.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequality. *Child Development*, *48*(2), 630–633.

Sella, F., Berteletti, I., Lucangeli, D., & Zorzi, M. (2015). Varieties of quantity estimation in children. *Developmental Psychology*, *51*(6), 758–770.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., et al. (2012). Early predictors of high school mathematics achievement. *Psychological Science*, *23*, 691–697.

Siegler, R. S., & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science*, *19*, 759–763.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*(3), 237–250.

Siegler, R. S., & Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Developmental Science*, *11*(5), 655–661.

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*(3), 545.

Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic-to-linear shift: One learning sequence, many tasks, many time scales. *Mind, Brain, and Education*, *3*(3), 143–150.

Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology*, *62*(4), 273–296.

Starkey, P., & Cooper, R. G. (1980). Perception of numbers by human infants. *Science*, *210*(4473), 1033–1035.

Stevenson, H. W., Chen, C., & Lee, S. Y. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, *231*, 693–699.

Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science*, *231*, 693–699.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap*. New York, NY: Summit.

Stevenson, H. W., Stigler, J. W., Lee, S.-Y., Lucker, G. W., Kitamura, S., & Hsu, C.-C. (1985). Cognitive performance and academic achievement of Japanese, Chinese, and American Children. *Child Development*, *56*, 718–734.

Stigler, J. W., & Hiebert, J. (1999). *Understanding and improving. Comparing Standards Internationally: research and practice in mathematics and beyond*. Oxford, UK: Symposium Books Ltd. 119.

Stigler, J. W., Gonzales, P., Kwanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, *52*(4), 1146–1152.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(13), 7836–7841.

Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effects of context on age and sex differences in symbolic magnitude estimation. *Journal of Experimental Child Psychology*, *101*(1), 20–51.

Tyler, J. H. (2004). Basic skills and the earnings of dropouts. *Economics of Education Review*, *23*(3), 221–235. https://doi.org/10.1016/j.econedurev.2003.04.001.

Vanbinst, K., Ansari, D., Ghesquière, P., & De Smedt, B. (2016). Symbolic numerical magnitude processing is as important to arithmetic as phonological awareness is to reading. *PloS ONE*, *11*(3), e0151045.

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2012). Numerical magnitude representations and individual differences in children's arithmetic strategy use. *Mind, Brain, and Education*, *6*(3), 129–136.

Wechsler, D. (1991). *Manual for the Wechsler intelligence scale for children-(WISC-III).* San Antonio, TX: Psychological Corporation.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, *358*(6389), 749–750.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*(1), 88–101.

Xu, X., Chen, C., Pan, M., & Li, N. (2013). Development of numerical estimation in Chinese preschool children. *Journal of Experimental Child Psychology*, *116*, 351–366.

Zhou, X., Chen, Y., Chen, C., Jiang, T., Zhang, H., & Dong, Q. (2007). Chinese kindergartners' automatic processing of numerical magnitude in Stroop-like tasks. *Memory & Cognition*, *35*, 464–470.

FURTHER READING

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*, 394–406.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., et al. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10593–10597.

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological Science*, *24*, 2013–2019.

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

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry*, *47*(2), 190–204.

Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Research*, *1106*(1), 177–188.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293–305.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.

Rousselle, L., & Noël, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, *102*, 361–395.

Stevenson, H. W., Lee, S.-Y., Chen, C., Lummis, M., Stigler, J., Fan, L., et al. (1990). Mathematics achievement of children in China and the United States. *Child Development*, *61*, 1053–1066.