Journal of Experimental Child Psychology 117 (2014) 29–44

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Number line estimation and mental addition: Examining the potential roles of language and education

Elida V. Laski *, Qingyi Yu

Department of Applied Developmental and Educational Psychology, Lynch School of Education, Boston College, Chestnut Hill, MA 02467, USA

ARTICLE INFO

Article history:
Received 25 January 2013
Revised 13 August 2013

Keywords:
Mathematics
Arithmetic
Number line estimation
Cross-cultural
Numerical

ABSTRACT

This study investigated the relative importance of language and education to the development of numerical knowledge. Consistent with previous research suggesting that counting systems that transparently reflect the base-10 system facilitate an understanding of numerical concepts, Chinese and Chinese American kindergartners' and second graders' number line estimation (0–100 and 0–1000) was 1 to 2 years more advanced than that of American children tested in previous studies. However, Chinese children performed better than their Chinese American peers, who were fluent in Chinese but had been educated in America, at kindergarten on 0–100 number lines, at second grade on 0–1000 number lines, and at both time points on complex addition problems. Overall, the pattern of findings suggests that educational approach may have a greater influence on numerical development than the linguistic structure of the counting system. The findings also demonstrate that, despite generating accurate estimates of numerical magnitude on 0–100 number lines earlier, it still takes Chinese children approximately 2 years to demonstrate accurate estimates on 0–1000 number lines, which raises questions about how to promote the mapping of knowledge across numerical scales.

Introduction

Although a range of species possess an approximate representation of numerical magnitude that allows them to reason about the magnitude of nonsymbolic quantities, only humans are able to

* Corresponding author. Fax: +1 617 552 1981.
E-mail address: laski@bc.edu (E.V. Laski).

0022-0965/$ - see front matter
http://dx.doi.org/10.1016/j.jecp.2013.08.007

precisely represent and operate on the magnitude of symbolic numerals (Cantlon, Platt, & Brannon, 2009; Feigenson, Dehaene, & Spelke, 2004). With development, humans become increasingly accurate at representing the magnitudes of large numbers (Barth & Paladino, 2011; Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008; Holloway & Ansari, 2008; Laski & Siegler, 2007; Siegler & Booth, 2004). This uniquely human ability and developmental trajectory is thought to be influenced by cultural practices, including language, the writing system, and formal education (Göbel, Shaki, & Fischer, 2011; Ho & Cheng, 1997; Ng & Rao, 2010; Towse & Saxton, 1998). This study investigated the relative importance of the structure of language and education to the development of accurate numerical estimates in large number ranges (0–100 and 0–1000) and exact mental calculation of sums in the same numerical ranges.

Understanding whether certain factors have a greater influence than others in the early development of numerical knowledge is important for identifying approaches for improving mathematics achievement. Children with more accurate numerical magnitude representations in the first grade show faster growth in math skills over the elementary school years even after controlling for alternative predictive factors such as intelligence and working memory (Geary, 2011). Furthermore, among both Western and Asian children, more accurate magnitude representations are associated with better performance on general number problems (Muldoon, Simms, Towse, Menzies, & Yue, 2011). If certain factors are found to be particularly important for early development of numerical magnitude knowledge, then they could be leveraged to improve children's mathematics achievement trajectory.

Language and numerical thinking

Language allows humans to count and name even very large quantities; thereby, it serves as a tool in the formation of exact mental representations of large numerical quantities and in mental operations on those representations. Evidence of this function of language has been found by examining the numerical understandings of young children who have not yet fully developed language as well as adults with limited linguistic systems for counting. Children and adults exhibit better understanding of numerical magnitude and greater accuracy on exact calculation tasks when they possess language words for the numerical range involved (Dehaene, Izard, Spelke, & Pica, 2008; Pica, Lemer, Izard, & Dehaene, 2004).

The linguistic structure of the counting system might also facilitate or impede numerical development. Languages differ in the extent to which they communicate features about number such as the base-10 system. In Chinese and other Asian languages, multi-digit numbers are expressed by consistent rules for combining the primary numbers (e.g., 12 is "ten-two") that transparently reflect the base-10 system. In contrast, English and other Western languages use inconsistent rules and arbitrary number words to express teens and other multi-digit numbers (e.g., 12 is "twelve"). This difference in the structure of counting systems has been associated with early numerical development in the areas of counting (Miller & Stigler, 1987; Song & Ginsburg, 1988), single-digit and multi-digit mental arithmetic (Dowker, Bala, & Lloyd, 2008; Geary, Bow-Thomas, Liu, & Siegler, 1996), place value understanding (Fuson, 1992), and the precision of numerical estimation (Helmreich et al., 2011; Siegler & Mu, 2008). Individuals from cultures with transparent counting systems consistently outperform those from cultures with less transparent systems in these aspects of numerical thinking.

The advantage in numerical thinking associated with languages with transparent counting systems seems to be independent of formal education. One indication of this dissociation is that the advantage emerges even before formal education (Miller & Stigler, 1987; Paik, van Gelderen, Gonzales, de Jong, & Hayes, 2011). For example, even before formal education, Chinese kindergartners generate estimates that increase linearly with the magnitude of the numbers presented for numbers between 0 and 100 on a number line estimation task, whereas their American peers generate estimates that increase logarithmically (Siegler & Mu, 2008). Another indication is that this advantage is also present in adults with different language backgrounds but similar levels of education (Campbell & Xue, 2001; Imbo & LeFevre, 2008).

Despite this strong evidence in favor of language's contribution to numerical knowledge, there are also findings that suggest it might not influence children's understanding of numerical magnitude. Consider German children's performance on a number line estimation task. The linguistic counting structure in German is even less transparent than that in English; in German number words, the decades and units do not match the order of digits in Arabic numerals (e.g., the word for 84 is *vierundachtzig,* which translates as "four and eighty"). If greater transparency of count words leads to better numerical understanding, then German children should have a poorer understanding of numerical magnitude than same-age English speakers. German children's performance on a 0–100 number line estimation task (Ebersbach et al., 2008), however, is comparable to that of English-speaking American children (Siegler & Booth, 2004).

Education and numerical thinking

Unsurprisingly, education also contributes to numerical development. Amount of education explains numerical knowledge even after controlling for age and other family and home characteristics (Magnuson, Meyers, Ruhm, & Waldfogel, 2004; Morrison, Griffith, & Alberts, 1997). Amount of education also is related to the accuracy of individuals' representations of magnitude as well as their ability to perform exact calculation (Dehaene et al., 2008; Pica et al., 2004).

All educational experiences, however, do not contribute to numerical development equally. Even within the same cultural context, curricular differences are related to differences in understanding of numerical concepts as well as arithmetic performance (Agodini et al., 2010; Ni, Li, Li, & Zhang, 2011). Thus, differences in educational systems and curricula must be considered when examining the role of language in numerical thinking (Wang & Lin, 2009).

Although the linguistic structures of the counting systems vary substantially between the United States and China, so do the educational systems and curricula. The amount of time children spend at school is greater in China than in the United States (Stevenson, Lee, & Stigler, 1986). In addition, there are differences in the nations' approaches to mathematics instruction. Compared with American teachers, Chinese teachers are more proactive in promoting student engagement during mathematics lessons, spend more time providing information about numerical concepts rather than having the children practice them, and place greater emphasis on arithmetic strategies that involve decomposing numbers into decades and units (Chan & Ho, 2010; Fuson & Kwon, 1992; Lan et al., 2009; Ma, 1999; Stevenson et al., 1986; Stigler, Chalip, & Miller, 1986).

The current study

There were four purposes of the current study. The first purpose was to examine whether the early advantages in numerical thinking associated with languages with transparent counting systems persist with development. Number line estimation was particularly suited for examining this issue for two reasons. First, prominent and recent theories of numerical development focus on the centrality of knowledge about numerical magnitudes (Case & Okamoto, 1996; Dehaene, 2011; Siegler, Thompson, & Schneider, 2011). The second reason is that consistent developmental patterns have been found on the number line estimation task. Children's estimates of numerical magnitude on number line estimation tasks become increasingly accurate with age. One model that has been used to describe developmental changes in number line estimation is a logarithmic-to-linear shift (cf. Siegler, Thompson, & Opfer, 2009). Research using logarithmicity and linearity of estimates as indicators of American children's number line estimation performance has found that the large majority of children do not typically generate linear estimates on 0–100 number lines until second grade; in kindergarten the large majority of children generate estimates best fit by a logarithmic function, and in first grade the two functions fit about equally well (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Opfer & Siegler, 2006; Siegler & Booth, 2004). On number lines from 0 to 1000, the developmental pattern repeats itself between second and fourth grades. In the range from 0 to 1000, the large majority of children do not typically generate linear estimates until fourth grade; in second grade, the large majority of children generate estimates best fit by a logarithmic function, and in third grade, the two functions fit about equally well (Booth & Siegler, 2006; Opfer & Siegler, 2006; Siegler & Opfer, 2003; Thompson & Opfer, 2010).

Chinese children, whose counting system reflects the base-10 system more transparently than the counting systems in English and other Western languages, seem to exhibit a substantially different timing in the development of linear number line estimates. Siegler and Mu (2008) found that the large majority of Chinese kindergartners they tested generated 0–100 number line estimates more linear than American kindergartners tested and comparable to those of American children 2 years older reported in previous studies. Whether this advantage on number line estimation continues with age and on larger numerical scales was an open question. Thus, in this study, we presented Chinese second graders with 0–1000 number lines.

The second purpose of the current study was to examine the relative importance of language and education to the development of numerical thinking on two tasks: number line estimation in large number ranges (0–100 and 0–1000) and mental addition. More specifically, we aimed to discriminate among three possibilities: (a) the linguistic structure of the counting system contributes more than education; (b) the educational system contributes more than the linguistic structure of the counting system; and (c) both the linguistic structure of the counting system and education contribute a comparable amount of influence. To examine this question, we analyzed differences in performance on number line estimation and mental addition between two groups of children—Chinese and Chinese American—with a shared language but different educational experiences.

An analysis of typical performance on these tasks suggested that each theoretical possibility would lead to a different pattern of performance differences between the Chinese and Chinese American children. If the linguistic structure of the counting system influences numerical development more than education, then same-age Chinese and Chinese American children should demonstrate comparable performance on number line estimation and mental addition because they have similar levels of experience with a transparent counting system. On the other hand, if the educational system and curricula influence numerical development more than the linguistic structure of the counting system, then the difference in Chinese and Chinese American children's performance should be consistent with that found in Siegler and Mu (2008). There should be substantial differences between Chinese and Chinese American children because, despite having a shared language, the groups have different educational experiences. Finally, if both linguistic structure of the counting system and education influence numerical development comparably, then Chinese children's performance on both numerical tasks should be more advanced than their Chinese American peers, but these differences should be tempered by the shared language.

A third purpose of the study was to explore the relation between a transparent counting system and base-10 understanding. Previous studies that have found an Asian advantage in understanding the base-10 structure of multi-digit numbers have typically involved asking children to use base-10 blocks to represent multi-digit numbers; greater use of tens blocks than unit cubes was inferred to reflect a better understanding of the base-10 structure of the number system (Miura, 1987; Miura, Kim, Chang, & Okamoto, 1988). It is not possible to determine from this block task, however, whether the better performance of Asian children is due to greater familiarity and practice with the base-10 structure of particular numbers or a more general understanding of base-10.

Examining children's number line estimation on two numerical scales (0–100 and 0–1000) allowed an evaluation of whether experience with a transparent counting system facilitates a general understanding of base-10. Greater knowledge about base-10 relations among numbers is related to better number line estimation performance in larger unfamiliar numerical ranges (Klein et al., 2010; Moeller, Pixner, Kauffman, & Nuerk, 2009; Nuerk, Moeller, Klein, Willmes, & Fischer, 2011; Opfer & Siegler, 2006; Thompson & Opfer, 2010). For example, Moeller and colleagues (2009) found a breakpoint between single-digit and double-digit numbers in the pattern of first graders' number line estimates. Based on this finding, they concluded that an increasing understanding of the integration of tens and units in accordance with the place value system contributes to improvements in number line estimation. Other findings support this view. For example, highlighting the commonalities between small and large numerical scales (e.g., 15:100 is equivalent to 150:1000) improves second graders' estimates on number lines greater than 0–100 (Thompson & Opfer, 2010). Thus, if a transparent counting system facilitates a general conceptual understanding of the base-10 system, then both groups of Chinese-speaking kindergartners should be able to transfer their knowledge of the numerical magnitudes of numbers between 0–100 and 0–1000 number lines. Children with general base-10 knowledge should extend this knowledge across numerical contexts and appropriately adjust their estimates in a larger numerical context because they better understand the relations between numbers (e.g., 50:100 is equivalent to 500:1000). In this case, children's performance on the two scales should be related and kindergartners who generate linear estimates on 0–100 number lines should also do so on 0–1000 number lines. On the other hand, if there is little benefit of a transparent language system to base-10 understanding, or if its benefit is limited to providing information about particular numbers, then there should be little transfer of knowledge between the two numerical scales as seen in American children. In this case, children's performance on the two scales might not be related and Chinese-speaking kindergartners might generate linear estimates only in their counting range (0–100) (Miller, Kelly, & Zhou, 2005; Miller, Smith, Zhu, & Zhang, 1995).

The fourth purpose of this study was to examine the relation between number line estimation and performance on complex addition problems. Studies examining the relation between number line estimation and arithmetic in American children have found that the quality of children's number line estimates is correlationally and causally related to their accuracy on arithmetic problems (Booth & Siegler, 2008; Geary, Hoard, Nugent, & Byrd-Craven, 2008; Geary, 2011). The evidence suggests that children use their numerical magnitude knowledge, or *number sense*, to estimate answers to arithmetic problems, to constrain their choices of possible answers to arithmetic problems, and to judge the probability that their solutions to arithmetic problems are correct. It seemed possible that the use of number sense in arithmetic might vary based on educational approach. Thus, we were interested in whether the pattern of relations would be different for Chinese and Chinese American children.

Method

Participants

Participants were two groups of kindergarten and second-grade children: Chinese and Chinese American. The Chinese children were born and reared in China, whereas the Chinese American children either were born in the United States or moved to the United States at a young age but had parents who were first-generation immigrants and spoke Chinese in the home. The Chinese group consisted of 41 kindergartners (20 boys and 21 girls, mean age = 6.07 years, *SD* = 0.23) and 31 second graders (16 boys and 15 girls, mean age = 7.89 years, *SD* = 0.42) recruited from a public school in a predominantly middle-income district in Shanghai, China. The Chinese American group consisted of 42 kindergartners (19 boys and 23 girls, mean age = 5.98 years, *SD* = 0.43) and 20 second graders (13 boys and 7 girls, mean age = 8.05 years, *SD* = 0.50). These participants were recruited from a middle-income bilingual Chinese public school and bilingual Chinese after-school program that used an American mathematics curriculum. At least half of the instruction during the school day and in the after-school program occurred in Chinese.

The Chinese participants' only language was Chinese, whereas the Chinese American participants were fluent in Chinese as well as English. All of the Chinese participants had attended school only in China, and 95% of the Chinese American participants had attended school only in the United States. The Chinese children were tested in Chinese, and the Chinese American children were tested in English. Both groups were tested during the spring of their school year; an analysis of the school calendars indicated that at this time the groups had spent an equivalent amount of time in school during the year. Although the children included in the study are not representative of an entire educational system, the instruction they received is reflective of the common approach in each nation. In China, the teachers followed the national curriculum materials. In the United States, the teachers used *Investigations in Number, Data, and Space*, which is a widely used curriculum across the United States (Agodini et al., 2010).

Materials and procedure

Children met one-on-one with an experimenter for a single session and completed two numerical tasks: number line estimation and mental addition. The order of task presentation was counterbalanced between participants.

Number line estimation

Both kindergartners and second graders completed two number line estimation tasks (0–100 and 0–1000) as a measure of their representations of numerical magnitude. The order of the tasks was counterbalanced between participants. The number line was presented on a cardboard strip 20 cm long with a removable pointer. The numbers to be estimated were presented one at a time on cards. On each trial, the pointer was handed to the children and they were instructed to place the pointer at the position on the line where the number presented belonged. Following 2 practice trials (0 and 100 or 0 and 1000), children completed the test trials with no feedback. The 0–100 number line estimation task consisted of 20 trials: 2, 5, 14, 18, 21, 26, 34, 39, 45, 48, 56, 59, 61, 67, 72, 78, 83, 89, 92, and 97. The 0–1000 number line estimation task consisted of 26 trials: 2, 5, 18, 21, 34, 45, 56, 67, 78, 89, 97, 122, 179, 246, 350, 366, 486, 517, 523, 606, 725, 754, 818, 881, 938, and 992. On both tasks, the order of the numbers to be estimated was randomized across participants.

Mental addition

Children's ability to retrieve or mentally calculate the answers to addition problems was tested using timed addition tasks. Kindergartners solved simple addition problems with two one-digit addends and solved complex addition problems with two-digit addends. The simple and complex addition problems were randomized in blocks, and the order of the presentation of each block was counterbalanced across participants. Second graders solved only complex addition problems with either two-digit addends or three-digit addends. Simple and complex addition problems were randomized in blocks, and the order of the presentation of each block was counterbalanced across participants.

Simple addition. Kindergartners were asked to solve as many single-digit addition problems as possible (out of 70) within 1 min. The problems were displayed one at a time on a computer screen and read aloud to the children. As soon as a child provided his or her answer verbally, the next problem was displayed. The problems were randomly selected among all possible problems with sums between 2 and 10, with the constraint that in half of the problems the largest addend was first and in the other half the largest addend was second. The problems were presented in the same random order for every participant.

Complex addition. Kindergartners solved six complex addition problems: one two-digit + one-digit problem (6 + 78) and five two-digit + two-digit problems (51 + 19, 35 + 38, 79 + 13, 32 + 49, and 66 + 23). The problems were presented one at a time on a computer screen and read aloud to the children. Children were instructed to solve them mentally and to respond verbally. For these two-digit problems, a 4 s time limit was imposed to encourage retrieval. Second graders solved the same problems as the kindergartners under the same time limit plus an additional six problems that consisted of three-digit numbers: three three-digit + two-digit problems (432 + 37, 215 + 16, and 38 + 290) and three three-digit + three-digit problems (127 + 214, 660 + 145, and 389 + 127). For these two-digit problems, an 8 s time limit was imposed. All children were encouraged to "give me your best guess" if they did not respond during the time limit. The problems were presented in the same random order for all participants.

Results

Number line estimation

Estimates for 0–100 number line

The first analysis examined the effects of grade and group on children's 0–100 number line estimates. The dependent measure was a composite score calculated by summing individuals' *z* scores for percentage absolute error [PAE = (|estimate − estimated quantity|/scale of estimates) × 100], linearity (R^2 of the best fitting linear function), and slope of the best fitting linear function. Each measure provides a somewhat different type of information. The ideal function relating actual and estimated magnitudes is a linear function with a slope of 1.00. However, estimates can increase linearly with a slope far less than 1.00, and estimates can increase with a slope of 1.00 but not fit a linear function

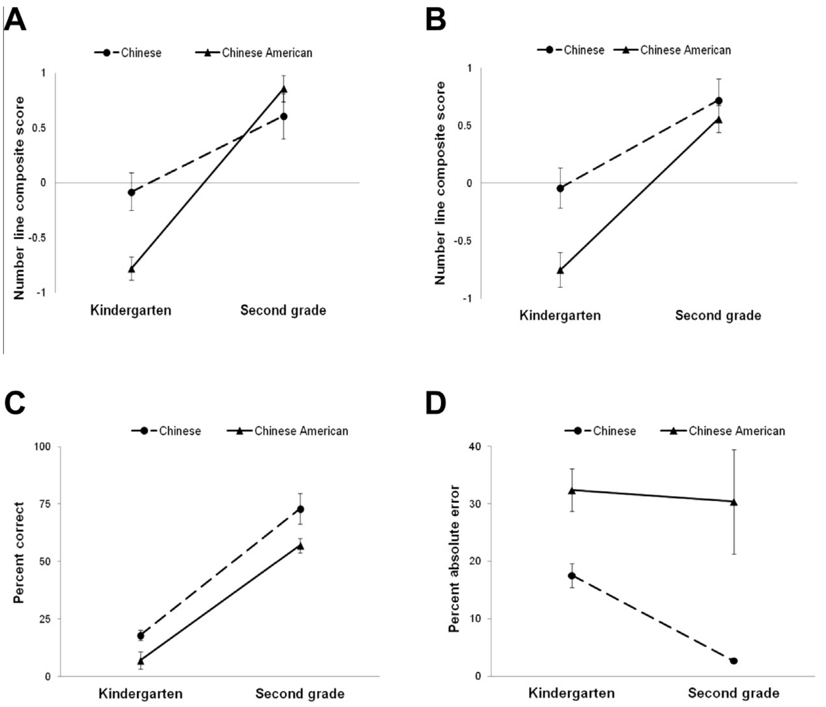

Fig. 1. Performance of children in each grade (kindergarten and second grade) and each group (Chinese and Chinese American) for each measure: (A) 0–100 number line estimation; (B) 0–1000 number line estimation; (C) complex addition percentage correct; (D) complex addition percentage absolute error.

very closely. Recent research also demonstrates that an increase in accuracy (PAE) does not necessarily mean a change in the underlying representation of numbers (see Opfer & Martens, 2012). Thus, creating a composite with the three measures allows for a more comprehensive and global assessment of number line estimation quality in addition to simplicity in reporting of the results.[1]

As shown in Fig. 1A, a 2 (Grade: kindergarten or second grade) × 2 (Group: Chinese or Chinese American) analysis of variance (ANOVA) indicated that children's 0–100 number line estimates varied by grade, $F(1,130) = 55.28$, $p < .001$, $\eta_p^2 = .30$, and the grade by group interaction, $F(1,130) = 9.18$, $p = .011$, $\eta_p^2 = .07$. At second grade, there was no difference in Chinese and Chinese American children's estimates (mean composite score = 0.61, *SD* = 0.67 and mean composite score = 0.86, *SD* = 0.92, respectively). However, at kindergarten, Chinese children's estimates were better than those of Chinese American children, (mean composite score = −0.08, SD = 0.67 and mean composite score = −0.78, *SD* = 1.11, respectively), $t(81) = 3.46$, $p = .001$, $d = 0.76$.

Next, we examined the fit of a linear and logarithmic function to each group's median estimates in order to categorize each group's performance on the number line task. Although there are other models to describe children's number line estimates (e.g., Barth & Paladino, 2011; Ebersbach et al., 2008), we used this approach because it is the model that has been most commonly used in the literature; thus, it most easily allowed comparison of the performance of our sample with that of previous studies (e.g., Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Mu, 2008; Siegler & Opfer, 2003). For each grade within each group, paired-samples *t* tests were used to compare the absolute value of the difference between the children's median estimate for each number and the predicted estimate for that number generated by the best fitting linear and logarithmic functions.

[1] The overall pattern of the results was the same when the measures were examined separately.

Both Chinese and Chinese American second graders' median estimates were better fit by the linear function than by the logarithmic function: Chinese second graders ($R^2_{lin} = .99$ vs. $R^2_{log} = .80$), $t(19) = 5.72$ $p < .001$, $d = 2.12$, and Chinese American second graders ($R^2_{lin} = .90$ vs. $R^2_{log} = .78$), $t(19) = 6.06$, $p < .001$, $d = 2.78$. On the other hand, Chinese kindergartners' median estimates were better fit by the linear function than by the logarithmic function ($R^2_{lin} = .95$ vs. $R^2_{log} = .87$), $t(19) = 3.14$, $p = .005$, $d = 1.44$, and the fit of the two functions did not differ for Chinese American kindergartners ($R^2_{lin} = .90$ vs. $R^2_{log} = .91$), $t(19) = 1.28$, $p = .216$.

To examine whether the findings regarding group medians also fit individual children's performance, we examined the fit of the linear functions to each child's estimates (see Fig. 2A). The model that fit the most individual children varied by grade and group. There was no difference in the percentage of Chinese and Chinese American second graders best fit by a linear function (100% and 95%), but a greater percentage of Chinese kindergartners' estimates were best fit by a linear function than those of Chinese American kindergartners (71% vs. 45%), $\chi^2(1, N = 81) = 5.53$, $p = .019$.

Thus, the results indicated that on 0–100 number lines, Chinese kindergartners' estimates were more accurate and more linear than those of their Chinese American peers. In addition, as expected, there were no differences between Chinese and Chinese American second graders' estimates.

Estimates for 0–1000 number line

As shown in Fig. 1B, a 2 (Grade) × 2 (Group) ANOVA on children's composite scores on 0–1000 number lines indicated that their estimate composite scores varied by grade, $F(1,130) = 36.17$, $p < .001$, $\eta^2_p = .22$, and by group, $F(1,130) = 6.31$, $p = .001$, $\eta^2_p = .05$. Second graders' estimates were better than those of kindergartners (mean composite score = 0.65, $SD = 0.72$ and mean composite score = -0.40, $SD = 1.11$, respectively), and Chinese children's estimates were better than those of Chinese American children (mean composite score = 0.28, $SD = 0.93$ and mean composite score = -0.33, $SD = 1.21$, respectively).

The analysis of the median number lines estimates of each group indicated that Chinese second graders' median estimates were better fit by the linear function than by the logarithmic function ($R^2_{lin} = .96$ vs. $R^2_{log} = .83$), $t(25) = 3.56$, $p = .006$, $d = 1.42$, and the fit of the two functions did not differ for Chinese American second graders ($R^2_{lin} = .94$ vs. $R^2_{log} = .85$), $t(25) = -0.25$, $p = .807$. Both groups of kindergartners, however, were better fit by the logarithmic function than by the linear function: Chinese kindergartners ($R^2_{lin} = .69$ vs. $R^2_{log} = .96$), $t(25) = 8.98$, $p < .001$, $d = 3.59$, and Chinese American kindergartners ($R^2_{lin} = .65$ vs. $R^2_{log} = .93$), $t(25) = 8.39$, $p < .001$, $d = 3.36$.

Again, the same pattern was found in individuals' estimates (see Fig. 2B). There was no difference in the percentage of Chinese and Chinese American second graders whose 0–1000 estimates were best fit by a linear function (71% and 55%) or in the percentage of Chinese and Chinese American kindergartners whose 0–1000 estimates were best fit by a linear function (19% and 19%). Thus, the results indicated that on 0–1000 number lines, Chinese second graders' estimates were more linear than

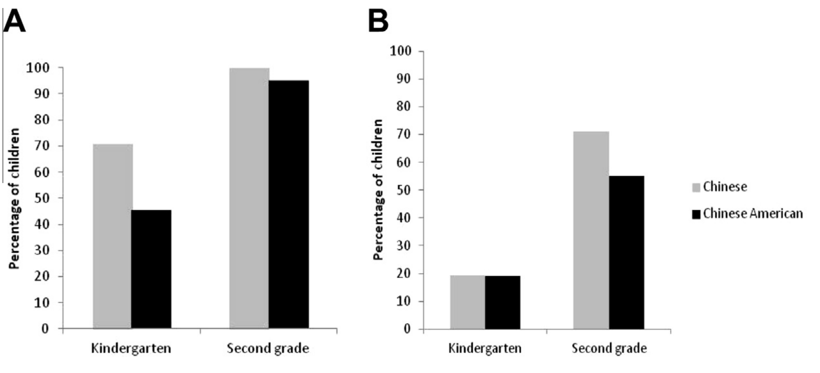

Fig. 2. Percentages of children whose 0–100 and 0–1000 number line estimates were best fit by a linear function in each grade and group: (A) 0–100 number line estimation; (B) 0–1000 number line estimation.

those of their Chinese American peers, whereas Chinese and Chinese American kindergartners performed similarly.

Relation of estimates on each scale

To examine the relation between children's performance on the two numerical scales, we computed correlations between children's composite scores on the 0–100 number line estimation task and their composite scores on the 0–1000 number line task. Overall, across grades and groups, individual differences on the two scales proved to be moderately correlated, $r(132) = .52$, $p < .001$. Chinese American children's estimates on the two scales were not correlated at kindergarten, $r(40) = .18$, $p = .25$, but were correlated at second grade, $r(18) = .53$, $p = .018$. In contrast, Chinese children's estimates were correlated at both kindergarten, $r(39) = .36$, $p = .02$, and second grade, $r(29) = .58$, $p = .001$.

To better understand the relation between Chinese children's performance on each scale, we examined their estimates on the subset of 11 numbers that were presented in both the 0–100 and 0–1000 estimation tasks. The median estimates of Chinese kindergartners and second graders on the 0–100 numerical context were best fit by a linear function. This performance could be because of knowledge of that particular scale of numbers or a more general understanding of base-10 or both. If a general understanding of base-10 was involved in children's estimates, then the function best fitting estimates of numbers below 100 on the 0–1000 number line task should also be linear. Children with base-10 knowledge should extend this knowledge across numerical contexts and appropriately adjust their estimates in a larger numerical context because they better understand the relations between numbers (e.g., 15:100 is equivalent to 150:1000).

The results did not support the idea that children's performance on the number line estimation tasks was related through base-10 knowledge. As shown in Fig. 3, for the 11 numbers presented on both tasks, the linear function fit Chinese kindergartners' and second graders' median estimates much better in the 0–100 numerical context than in the 0–1000 context: kindergarten ($R^2_{lin} = .96$ vs. .81), $t(10) = 5.84$, $p < .001$, $d = 2.27$, and second grade ($R^2_{lin} = .99$ vs. 90), $t(10) = 6.16$, $p < .001$, $d = 2.66$. In the 0–1000 numerical context, the logarithmic function fit kindergartners' and second graders' median estimates for these same 11 numbers better than the linear function: kindergarten ($R^2_{log} = .94$ vs. $R^2_{lin} = .83$), $t(10) = 4.05$, $p = .002$, $d = 2.56$, and second grade ($R^2_{log} = .92$ vs. $R^2_{lin} = .90$), $t(10) = 2.81$, $p = .018$, $d = 1.78$.

Mental addition

Simple addition

Chinese kindergartners' ability to quickly and accurately mentally solve simple addition problems far exceeded that of Chinese American kindergartners; Chinese kindergartners provided more than

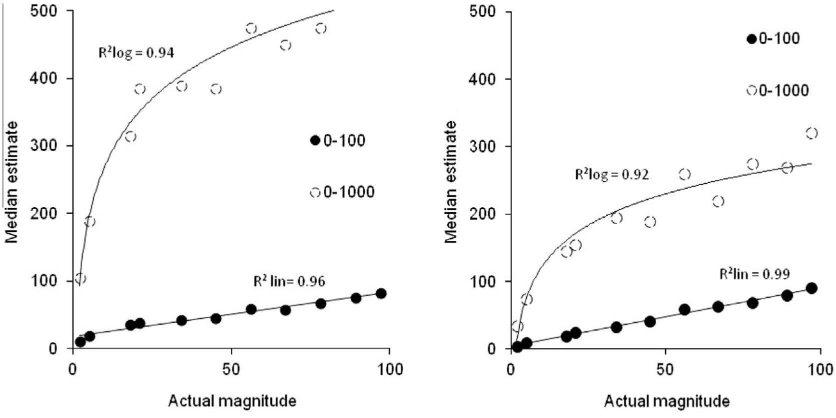

Fig. 3. Chinese kindergartners' (left panel) and second graders' (right panel) median estimates on the 0–100 and 0–1000 number lines for numbers presented on both tasks. Also shown are the best fitting functions (linear or logarithmic) and the variances accounted for by those functions.

twice as many correct answers to the simple addition problems within the 1 min time limit than Chinese American kindergartners (*M* = 16.98, *SD* = 7.81 and *M* = 8.60, *SD* = 6.12, respectively), $t(81) = 5.45$, $p < .001$, $d = 1.21$. In addition, Chinese kindergartners also attempted to answer more questions than Chinese American kindergartners (*M* = 17.46, *SD* = 7.54 and *M* = 10.57, *SD* = 5.68, respectively), $t(81) = 4.71$, $p < .001$, $d = 1.05$, and achieved a higher percentage correct among the attempted answers than Chinese American kindergartners (*M* = 96%, *SD* = 7.40% and *M* = 76%, *SD* = 26.92%, respectively), $t(81) = 4.67$, $p < .001$, $d = 1.04$.

Complex addition

As shown in Fig. 1C, Chinese children were also more accurate than Chinese American children on the complex addition problems regardless of grade. An ANOVA found main effects for grade, $F(1,130) = 196.72$, $p < .001$, $\eta_p^2 = .60$, and group, $F(1,130) = 12.81$, $p < .001$, $\eta_p^2 = .09$, but no grade by group interaction. Chinese kindergartners answered more problems correctly than Chinese American kindergartners (18%, *SD* = 24.11% vs. 7%, SD = 13.89%), $t(81) = 2.50$, $p = .015$, $d = 0.56$, and Chinese second graders answered more problems correctly than Chinese American second graders (73%, *SD* = 17.40% vs. 57%, *SD* = 29.65%), $t(49) = 2.48$, $p = .017$, $d = 0.71$.

To better understand children's performance on the complex arithmetic problems, we also examined percentage absolute error, calculated as PAE = (|child's answer − correct answer|/scale of answers) × 100. Problems involving two-digit addends had answers in the 0–100 numerical range, and problems involving three-digit numbers had answers in the 0–1000 numerical range. We calculated children's mean PAE across all of the complex addition problems with which they were presented. PAE provides a continuous measure of the degree to which children's answers are of the right magnitude, which is likely to be more closely related to understanding of numerical magnitudes than percentage correct (Booth & Siegler, 2008). Linear representations of numerical magnitude are characterized by greater differentiation between numbers than logarithmic representations, which in turn probably constrains the plausible range of sums to problems. Thus, although children with linear representations might not be able to estimate the exact answer to an addition problem, their estimates should be closer misses than those of children who possess logarithmic representations of numerical magnitude.

The pattern of results was the same. An ANOVA found main effects for grade, $F(1,130) = 4.59$, $p = .008$, $\eta_p^2 = .03$, and group, $F(1,130) = 29.00$, $p < .001$, $\eta_p^2 = .18$. The magnitudes of Chinese kindergartners' answers were about twice as close to the correct magnitudes than those of Chinese American kindergartners (PAE = 18%, *SD* = 13.35% vs. PAE = 32%, *SD* = 24.06%), $t(81) = 3.46$, $p = .001$, $d = 0.76$. The difference between groups was even greater between second graders. The magnitudes of Chinese second graders' answers were more than four times as close to the correct magnitudes as those of Chinese American second graders (PAE = 2%, *SD* = 2.37% vs. PAE = 30%, *SD* = 40.44%), $t(49) = 3.82$, $p < .001$, $d = 0.97$.

Relation between number line estimation and mental addition

To determine the degree to which individual differences in performance on the numerical tasks were related, we first computed correlations between each child's composite scores for the 0–100 and 0–1000 number line estimation tasks and each child's percentage absolute error on the complex addition task.

As shown in Table 1, in general, better number line estimation was associated with lower percentage absolute error on the addition task, but the relation between the tasks changed with age. Kindergartners' performance on the 0–100 number line task was related to their percentage absolute error on complex addition performance, but there was no relation between the tasks at second grade. On the other hand, second graders' performance on the 0–1000 number line task was related to their percentage absolute error on complex addition, but there was no relation between the tasks at kindergarten. This finding probably reflects the magnitude of the addition problems that were presented at each grade level. Kindergartners were asked to solve addition problems only with sums up to 100, whereas second graders were asked to solve six additional problems involving three-digit numbers with sums up to 1000.

Table 1
Correlations between kindergartners' and second graders' 0–100 and 0–1000 number line estimates and their percentage absolute error on the complex addition task.

Number line estimation	Complex addition[a]		
	All children Percentage absolute error (PAE)	Chinese Percentage absolute error (PAE)	Chinese American Percentage absolute error (PAE)
0–100 estimate composite score			
Kindergarten	−.56**	−.56**	−.47**
Second grade	.12	−.06	.08
0–1000 estimate composite score			
Kindergarten	−.15	−.22	.04
Second Grade	−.41**	−.22	−.56**

** $p < .01$.
[a] Kindergartners solved complex addition problems with sums up to 100, and second graders solved problems with sums up to 1000.

Among second graders, Chinese American children's performance on the 0–1000 number line task was related to their complex addition performance, but there was no relation between the tasks for Chinese second graders. There was little variance in Chinese second graders' performance on the complex addition task (mean PAE of only 2%, $SD = 2.37\%$), which may have led to the lack of a correlation.

Discussion

Influence of language and education on numerical thinking

The current study examined differences in performance between two groups of children—Chinese and Chinese American—with a shared language but different educational experiences to gain insight into the extent to which language and educational experiences might contribute to children's developing understanding of the magnitude of larger numbers. Previous studies have suggested that counting systems that transparently reflect the base-10 system, such as in the Chinese language, facilitate an understanding of the magnitudes of numbers and an ability to operate on larger numbers. The pattern of findings in the current study provides some support for this view but provides stronger evidence for the importance of educational approach to the development of numerical knowledge.

In support of the idea that languages with a transparent counting structure facilitate the development of numerical knowledge, both groups of Chinese-speaking children in this study performed better on number line estimation than what is typical of samples of American children. Chinese kindergartners' median estimates on 0–100 number lines and Chinese second graders' median estimates on 0–1000 number lines were comparable to the performance of American children approximately 2 years older in previous studies (cf. Booth & Siegler, 2006; Laski & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003). The median estimates of Chinese American children, who were fluent in Chinese but had been educated in America on 0–100 number lines in kindergarten and 0–1000 number lines in second grade, were comparable to the performance of American children approximately 1 year older in previous studies (cf. Booth & Siegler, 2006; Laski & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003). Thus, the current findings suggest that a transparent base-10 counting system may facilitate an earlier understanding of the magnitude of large numbers that persists with age even across different educational contexts.

The differences found between Chinese and Chinese American children, however, point to the importance of educational approach in the development of numerical knowledge. Chinese children performed better than same-age Chinese American children who were fluent in Chinese but had been educated in America. This finding was true at both kindergarten and second grade on every task except for 0–100 number line estimation at second grade, where both groups were expected to perform well.

Furthermore, the differences between the groups on addition and number line estimation at kindergarten and second grade mirrored expected curricular and instructional differences between China and the United States. Curriculum analyses and classroom observations indicate that Chinese teachers spend a greater amount of time in the primary grades teaching and practicing arithmetic strategies than American teachers (e.g., Geary et al., 1996; Stevenson et al., 1986). Consistent with the differences in instructional time spent on arithmetic between kindergarten and second grade, differences in Chinese and Chinese American children's performance on complex addition, in terms of both percentage correct and percentage absolute error, increased between kindergarten and second grade. At kindergarten, Chinese children's answers to the complex addition problems were about twice as close to the correct magnitudes as those of Chinese American kindergartners but more than four times as close at second grade.

Differences in Chinese and Chinese American children's performance on 0–100 and 0–1000 number line estimation decreased between kindergarten and second grade. This pattern is also consistent with curricular differences in China and the United States. Although it is unlikely that children in either context receive instruction on this task, instruction that provides greater exposure to numbers in these number ranges would benefit number line estimation. One point of agreement among the varied theoretical accounts of number line estimation is that familiarity with the particular numbers being estimated improves the accuracy of estimates (Barth & Paladino, 2011; Ebersbach et al., 2008; Moeller et al., 2009; Siegler & Booth, 2004).

Prior research has found that even by preschool the learning environments are richer in mathematics in China than in the United States. For instance, Starkey and Klein (2006) found that Chinese teachers provide, on average, 36 min of math support to 4-year-old children per day, whereas American teachers provide only 21 min through mostly whole group activities. The current finding that Chinese kindergartners provided more than twice as many correct answers to the simple addition problems within the 1 min time limit than Chinese American kindergartners also suggests that Chinese children had had more mathematics experience by kindergarten than Chinese American children. Furthermore, the finding that Chinese kindergartners' performance, but not Chinese American kindergartners' performance, on the 0–100 and 0–1000 number line estimation tasks was related, and this relation did not seem to be due to mapping base-10 knowledge across the scales, suggests that Chinese children had greater exposure to larger numbers by kindergarten. It is possible that by second grade, differences in children's mathematics experience with the numbers up to 1000 in China and the United States were less stark, which attenuated the difference in their performance on number line estimation.

Other possible group differences might have influenced the current results. The current study did not control for proficiency in Chinese or aspects of Chinese language use such as frequency, context (school vs. home), and content (academic vs. informal). Thus, it is possible that Chinese children performed better than Chinese American children on the numerical tasks because they were more proficient in or had greater experience with the linguistic structure of the Chinese counting system rather than because of differences in educational approaches. We believe, however, that this is unlikely. The Chinese American children included in the current study attended a bilingual Chinese school in which at least half of the mathematics instruction was conducted in Chinese. Furthermore, the children in the after-school program received additional instruction in Chinese.

Another potential reason why Chinese children may have performed better than Chinese American children on the numerical tasks could be because the Chinese American children were bilingual. Previous research suggests, however, that although bilingualism may lead to poorer performance in mathematics in preschool, by 5 years of age bilingual children perform comparable to or better than monolingual speakers, at least in terms of their counting ability (Song & Ginsburg, 1988). Furthermore, if the Chinese American kindergartners and second graders were at a disadvantage because of their bilingualism, then one might expect the Chinese American children's performance to have been worse than that of monolingual English speakers who have been tested in other studies. Yet, the Chinese American children's number line estimates were more accurate and linear than those of same-age English speakers reported in previous studies (e.g., Booth & Siegler, 2006; Siegler & Booth, 2004). For instance, in this study, Chinese American kindergartners' median estimates on 0–100 number lines were equally well fit by the linear and logarithmic functions, whereas Siegler and Booth (2004) found that the median estimates of English-speaking American kindergartners were better fit by a logarithmic function than by a linear function. Similarly, other studies that have examined the mathematics performance of bilingual children of the same age as those in this study have found that bilingual children's mathematics is no worse, and often better, than that of monolingual speakers. For example, Dowker and colleagues (2008) found that Welsh–English bilingual children performed better on a numerical magnitude comparison task than monolingual English speakers.

The current study also did not control for family values, including parental expectations and support, related to mathematics. This factor might have contributed to the better performance of the Chinese-speaking children in this study relative to the performance of American children in previous studies. However, it is less likely to have contributed to the differences between Chinese and Chinese American children. Starting in preschool, both Chinese and Chinese American parents set higher expectations for their children's mathematics achievement, engage their children in working more on mathematics at home, and use more formal and systematic instructional approaches at home than Euro-American parents (Huntsinger, Jose, Larson, Balsink Krieg, & Shaligram, 2000; Huntsinger, Jose, Liaw, & Ching, 1997; Starkey & Klein, 2008).

Relation between understanding and operating on large numbers

Previous research has found that the quality of children's number line estimates is correlationally and causally related to their accuracy on arithmetic problems (Booth & Siegler, 2008; Geary, 2011; Geary et al., 2008). The current study expands these findings by suggesting that the relation depends on a match between the magnitude of the numbers being estimated and the magnitude of the sums of the addition problems. Kindergartners solved problems with sums only up to 100; their performance on the 0–100 number line task, but not on the 0–1000 number line task, was related to their percentage absolute error on complex addition performance. On the other hand, second graders solved problems with sums up to 1000; their performance on the 0–1000 number line task was related to their percentage absolute error on complex addition. This finding is consistent with the view that children use their knowledge of numerical magnitude to constrain their choices of possible answers to arithmetic problems.

The current study also provides evidence that number line estimation predicts performance on addition problems among children educated outside the United States, at least at kindergarten. At second grade, Chinese American children's estimates, but not Chinese children's estimates, on 0–1000 number lines were found to be associated with their percentage absolute error on complex addition. The lack of finding of a relation between the tasks for Chinese second graders could be because there was little variance in Chinese second graders' performance on the complex addition task.

Alternatively, it may be that Chinese second graders did not rely on their numerical magnitude knowledge for generating the solutions to the complex addition problems. It seems possible that, because Chinese instruction spends a greater amount of time on practicing arithmetic strategies, Chinese second graders quickly used procedural knowledge to generate the sums and relied less on their knowledge of numerical magnitude to retrieve the correct answers or to estimate plausible answers. This hypothesis is consistent with Chinese second graders' high accuracy and low percentage absolute error on the complex addition problems, which seems unlikely to have been from retrieving answers to problems involving two three-digit numbers. This hypothesis is also consistent with evidence that the number processing areas of the brain show different levels of activation depending on problem complexity and training type (drill based vs. strategy based) (Zamarian, Ischebeck, & Delazer, 2009). It would be interesting to test whether the relation between number line estimation and arithmetic emerges when Chinese second graders are faced with even more complex problems or with the same problems using a shorter time limit.

Developmental progression of numerical magnitude understanding

Children progressively come to understand the magnitude of larger numbers and to operate on them. Consistent with other studies, the current findings indicate that it takes approximately 2 years after children are able to accurately estimate the magnitude of the numbers between 0 and 100 for them to do so for the numbers between 0 and 1000 (e.g., Booth & Siegler, 2006; Siegler & Booth, 2004).

Although earlier acquisition of a linear representation on one scale may lead to an earlier representation on another larger scale, it does not seem to affect the time course of learning. Just as in Siegler and Mu (2008), the Chinese kindergartners in this study generated linear estimates of numerical magnitude for the numbers between 0 and 100. It was not until second grade, however, that Chinese children generated linear estimates on 0–1000 number lines. Thus, the length of time between when the majority of Chinese children generate linear estimates on 0–100 number lines and when the majority do so on 0–1000 number lines was similar to that of American children who demonstrate the transition between second and fourth grades. This finding suggests that the acquisition of a linear representation of numbers is not "developmentally" constrained but that there may be a minimum level of experience necessary.

What kind of experience helps children to map their knowledge of numbers in one scale to a larger less familiar scale? We had hypothesized that a transparent base-10 counting system might facilitate this mapping; however, the current results suggest that this is not the case. Rather, the results seem to suggest that experience with the particular numbers in a given scale is necessary. In other words, the current results suggest that children will demonstrate linear estimates in a number range at about the time when their instruction includes practice with the numbers in that range.

This interpretation raises the question of what kind of practice with numbers in a given scale are necessary to support linear estimates. Is counting experience in a range sufficient? Studies that have looked at the association between counting and estimation ability indicate that it is not. Children are able to count correctly in a numerical range at least a year before they show much knowledge of the numerical magnitudes in the same range (Le Corre, Van de Walle, Brannon, & Carey, 2006; Lipton & Spelke, 2005). Also suggesting that counting is not sufficient, the Common Core Mathematics Standards in the United States expect second graders to count to 1000, but U.S. second graders do not generate linear estimates on 0–1000 number lines (National Governors Association Center for Best Practices, 2010; Siegler & Opfer, 2003). This suggests that other kinds of practice must be important to understanding numerical magnitudes.

One possibility is that practice incrementing and decrementing numbers in a given number range, such as when solving arithmetic problems, may be an important contributor to developing an accurate understanding of the magnitude of numbers within that range. A cursory examination of when Chinese instruction and American instruction present addition problems with sums up to 1000 supports this idea. According to the Common Core Mathematics Standards, U.S. students do not practice adding and subtracting numbers within 1000 until third grade (National Governors Association Center for Best Practices, 2010). On the other hand, a review of the mathematics workbooks used as part of the national curriculum in China found that nearly a third of the arithmetic problems practiced in second grade involved adding and subtracting numbers within 1000. It remains up to future studies to test this possibility more systematically.

Acknowledgments

Thanks are due to the children, teachers, and principals of Huanjiang Kindergarten, Kwong Kow Chinese School, Josiah Quincy Elementary School, and Kongjiang Secondary School; to Zi Chen, Xiaoli He, Fei Lou, Cynthia Soo Hoo, Wei Tang, Zhao Qin, and Wei Wu for assistance in collecting the data; and to Boston College's Academic Technology Office (ATIG grant), which supported this research.

References

Agodini, R., Harris, B., Thomas, M., Murphy, R., Gallagher, L., & Pendleton, A. (2010). *Achievement effects of four early elementary school math curricula: Findings for first and second graders*. Executive summary, National Center for Education Evaluation and Regional Assistance, No. 2011-4002.

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*, 125–135.

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*, 545–551.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*, 189–201.

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

Campbell, J. I. D., & Xue, Q. L. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*, 299–315.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13*, 83–91.

Case, R., & Okamoto, Y. (1996). The role of central conceptual structures in the development of children's thought. *Monographs for the Society for Research in Child Development*, *61*(1–2, Serial No. 246).

National Governors Association for Best Practices, Council of Chief State School Officers (2010). *Common Core Standards for Mathematics*. Washington, DC: Author.

Chan, B. M. Y., & Ho, C. S. H. (2010). The cognitive profile of Chinese children with mathematics difficulties. *Journal of Experimental Child Psychology, 107*, 260–279.

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (rev. ed.). New York: 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*, 1217–1220.

Dowker, A., Bala, S., & Lloyd, D. (2008). Linguistic influences on mathematical development: How important is the transparency of the counting system? *Philosophical Psychology, 21*, 523–538.

Ebersbach, M., Luwel, K., Frick, A., Onghena, P., & Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year-old children: Evidence for a segmented linear model. *Journal of Experimental Child Psychology, 99*, 1–17.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.

Fuson, K. C. (1992). Issues in place-value and multidigit addition and subtraction learning and teaching. *Journal for Research in Mathematics Education, 21*, 273–280.

Fuson, K., & Kwon, Y. (1992). Korean children's understanding of multidigit addition and subtraction. *Child Development, 63*, 491–506.

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

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., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development, 78*, 1343–1359.

Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33*, 277–299.

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology, 42*, 543–565.

Helmreich, I., Zuber, J., Pixner, S., Kaufmann, L., Nuerk, H. C., & Moeller, K. (2011). Language effects on children's nonverbal number line estimations. *Journal of Cross-Cultural Psychology, 42*, 598–613.

Ho, C. S.-H., & Cheng, F. S.-F. (1997). Training in place-value concepts improves children's addition skills. *Contemporary Educational Psychology, 22*, 495–506.

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science, 11*, 644–649.

Huntsinger, C. S., Jose, P. E., Larson, S. L., Balsink Krieg, D., & Shaligram, C. (2000). Mathematics, vocabulary, and reading development in Chinese American and European American children over the primary school years. *Journal of Educational Psychology, 92*, 745–760.

Huntsinger, C. S., Jose, P. E., Liaw, F. R., & Ching, W. D. (1997). Cultural differences in early mathematics learning: A comparison of Euro-American, Chinese-American, and Taiwan-Chinese families. *International Journal of Behavioral Development, 21*, 371–388.

Imbo, I., & LeFevre, J. A. (2008). Complex-arithmetic problem solving: Differences among Belgian, Canadian, and Chinese participants. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings from the 30th annual meeting of the Cognitive Science Society* (pp. 2144–2149). Austin, TX: Cognitive Science Society.

Klein, E., Moeller, K., Dressel, K., Domahs, F., Wood, G., Willmes, K., et al (2010). To carry or not to carry—is this the question? Disentangling the carry effect in multi-digit addition. *Acta Psychologica, 135*, 67–76.

Lan, X., Ponitz, C. C., Miller, K. F., Li, S., Cortina, K., Perry, M., et al (2009). Keeping their attention: Classroom practices associated with behavioral engagement in first grade mathematics classes in China and the United States. *Early Childhood Research Quarterly, 24*, 198–211.

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*, 1723–1743.

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52*, 130–169.

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development, 76*, 978–988.

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

Magnuson, K. A., Meyers, M. K., Ruhm, C. J., & Waldfogel, J. (2004). Inequality in preschool education and school readiness. *American Educational Research Journal, 41*, 115–157.

Miller, K. F., Kelly, M., & Zhou, X. (2005). Learning mathematics in China and the United States. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 163–178). New York: Psychology Press.

Miller, K. F., Smith, C., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming system. *Psychological Science, 6*, 56–60.

Miller, K. F., & Stigler, J. W. (1987). Counting in Chinese: Cultural variation in a basic cognitive skill. *Cognitive Development, 2*, 279–305.
Miura, I. T. (1987). Mathematics achievement as a function of language. *Journal of Educational Psychology, 79*, 79–82.
Miura, I. T., Kim, C. C., Chang, C., & Okamoto, Y. (1988). Effects of language characteristics on children's cognitive representation of number: Cross-national comparisons. *Child Development, 59*, 1445–1450.
Moeller, K., Pixner, S., Kauffman, L., & Nuerk, H. C. (2009). Children's early mental number line: Logarithmic or decomposed linear? *Journal of Experimental Child Psychology, 103*, 503–515.
Morrison, F. J., Griffith, E. M., & Alberts, D. M. (1997). Nature–nurture in the classroom: Entrance age, school readiness, and learning in children. *Developmental Psychology, 33*, 254–262.
Muldoon, K., Simms, V., Towse, J., Menzies, V., & Yue, G. (2011). Cross-cultural comparisons of 5-year-olds' estimating and mathematical ability. *Journal of Cross-Cultural Psychology, 42*, 669–681.
Ng, S. S. N., & Rao, N. (2010). Chinese number words, culture, and mathematics learning. *Review of Educational Research, 80*, 180–206.
Ni, Y., Li, Q., Li, X., & Zhang, Z. (2011). Influence of curriculum reform: An analysis of student mathematics achievement in Mainland China. *International Journal of Educational Research, 50*, 100–116.
Nuerk, H. C., Moeller, K., Klein, E., Willmes, K., & Fischer, M. H. (2011). Extending the mental number line: A review of multi-digit number processing. *Journal of Psychology, 219*, 3–22.
Opfer, J. E., & Martens, M. A. (2012). Learning without representational change: Development of numerical estimation in individuals with Williams syndrome. *Developmental Science, 15*, 863–875.
Opfer, J. E., & Siegler, R. S. (2006). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169–195.
Paik, J. H., van Gelderen, L., Gonzales, M., de Jong, P. F., & Hayes, M. (2011). Cultural differences in early math skills among U.S., Taiwanese, Dutch, and Peruvian preschoolers. *International Journal of Early Years Education, 19*, 133–143.
Pica, P., Lemer, C., Izard, W., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499–503.
Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428–444.
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*, 237–243.
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*, 143–150.
Siegler, R. S., Thompson, C. A., & Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cognitive Psychology, 62*, 273–296.
Song, M. J., & Ginsburg, H. P. (1988). The effect of the Korean number system on young children's counting: A natural experiment in numerical bilingualism. *International Journal of Psychology, 23*, 319–332.
Starkey, P., & Klein, A. (2006). *The early development of mathematical cognition in socioeconomic and cultural contexts*. Paper presented at the Institute for Education Sciences Research Conference, Washington, DC.
Starkey, P., & Klein, A. (2008). Sociocultural influences on young children's mathematical knowledge. In O. Saracho & B. Spodek (Eds.), *Contemporary perspectives on mathematics in early childhood education* (pp. 253–276). Charlotte, NC: IAP/Information Age.
Stevenson, H. W., Lee, S. Y., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 231*, 693–699.
Stigler, J. W., Chalip, L., & Miller, K. F. (1986). Consequences of skill: The case of abacus training in Taiwan. *American Journal of Education, 94*, 447–479.
Thompson, C., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81*, 1768–1786.
Towse, J., & Saxton, M. (1998). Mathematics across national boundaries. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 129–150). Hove, UK: Psychology Press.
Wang, J., & Lin, E. (2009). A meta-analysis of comparative studies on Chinese and U.S. students' mathematics performance: Implications for mathematics education reform and research. *Educational Research Review, 4*, 177–195.
Zamarian, L., Ischebeck, A., & Delazer, M. (2009). Neuroscience of learning arithmetic: Evidence from brain imaging studies. *Neuroscience & Biobehavioral Reviews, 33*, 909–925.