Journal of Educational Psychology

When Prior Knowledge Interferes, Inhibitory Control Matters for Learning: The Case of Numerical Magnitude Representations

Elida V. Laski and Alana Dulaney

Online First Publication, May 11, 2015. http://dx.doi.org/10.1037/edu0000034

Journal of Educational Psychology

0022-0663/15/$12.00 http://dx.doi.org/10.1037/edu0000034

When Prior Knowledge Interferes, Inhibitory Control Matters for Learning: The Case of Numerical Magnitude Representations

Elida V. Laski and Alana Dulaney
Boston College

The present study tested the *interference hypothesis*—that learning and using more advanced representations and strategies requires the inhibition of prior, less advanced ones. Specifically, it examined the relation between inhibitory control and number line estimation performance. Experiment 1 compared the accuracy of adults' ($N = 53$) estimates on 2 number line tasks, 1 with standard (power of 10) endpoints (0–1,000) and the other with nonstandard endpoints (364–1,364). Inhibition, as measured by Stroop task performance, predicted the accuracy of estimation on the nonstandard number line task, above and beyond estimation on the standard task. In Experiment 2, changes in kindergartners' ($N = 42$) 0–100 number line estimation were elicited through randomized training conditions, which involved playing a numerical board game. Stroop task performance was related to the rate of improvement in estimation, controlling for pretest number line task performance. The results provide a potential explanation for the relation between inhibitory control and mathematics achievement: Individuals with better inhibitory control may be better able to suppress the activation of prior knowledge and may be less vulnerable to interference from such knowledge. Potential implications for instructional design are discussed.

Keywords: inhibitory control, executive function, number line estimation, mathematics, children's learning

Fewer than half of high school graduates in the United States demonstrate the level of mathematics competence necessary for success in college (ACT, 2006). These inadequate levels of mathematics knowledge negatively impact the national economy and individual citizens' college, career, and economic opportunities (National Mathematics Advisory Panel, 2008; Rivera-Batiz, 1992). To address the problem, it is crucial to understand the processes involved in the development of foundational numerical knowledge.

One aspect of numerical knowledge that has been found to be foundational for later math learning is understanding the magnitude of numbers. Children's ability to accurately estimate the position of numbers on a number line, such that their estimates of numerical magnitude increase linearly with the size of the numbers being estimated, is predictive of rate of growth in mathematics and later mathematics achievement (Booth & Siegler, 2006; Geary, 2011; Holloway & Ansari, 2009). Unfortunately, until about second grade, most children's estimates of numerical magnitude on a 0–100 number line increase logarithmically, in which they generate estimates that overestimate the size of smaller numbers and the differences between them, but underestimate the differences between numbers at the high end. Furthermore, the processes underlying improvement in numerical magnitude representations are not well understood. Motivated by evidence that inhibitory control plays a role in mathematics performance (e.g., Lemaire, 2010), the present study tested the hypothesis that one's ability to inhibit a bias toward logarithmic representations of numerical magnitude contributes to their acquisition and use of linear representations.

This introductory section includes three parts. First, we describe findings regarding the relation between representations of numerical magnitude and performance on mathematics achievement tests as well as changes in these representations over development. Second, we describe current knowledge regarding the role of inhibitory control in mathematics learning. Finally, we present our hypotheses regarding the role of inhibitory control in the use and acquisition of linear and accurate representations of numerical magnitude and describe how these hypotheses were tested in the present experiments.

Representations of Numerical Magnitude

Basic "number sense" involves being able to associate a numeral with the quantity or magnitude it represents. This number sense allows individuals to make comparisons such as "greater-than or less-than" as well as to estimate and judge the plausibility of answers to arithmetic operations (Barth et al., 2006; Berch, 2005; Jordan, Kaplan, Ramineni, & Locuniak, 2009). Performance on measures of numerical magnitude knowledge correlates strongly

Elida V. Laski and Alana Dulaney, Department of Applied Developmental and Educational Psychology, Boston College.

Alana Dulaney is now in the Psychology Department at the University of Chicago.

This research was supported by a Boston College ATIG grant and Research Expense Grant awarded to the first author. We would like to acknowledge Qingyi Yu, Jessica Shapiro, Anne McNamara, Elizabeth Correa, Francisca Ting, and Cote Theriault for their assistance in data collection and coding. Thanks also to Robert Siegler and Marina Vasilyeva, who provided comments on a draft of this article.

Correspondence concerning this article should be addressed to Elida V. Laski, 201 Campion Hall, Lynch School of Education, Department of Applied Developmental and Educational Psychology, Boston College, Chestnut Hill, MA 02467. E-mail: laski@bc.edu

with mathematics achievement test scores at all grade levels from kindergarten through eighth grade (Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Holloway & Ansari, 2009; Schneider, Grabner, & Paetsch, 2009; Siegler & Booth, 2004). Further, children with more advanced numerical magnitude knowledge in first grade show faster growth in math skills over the elementary school years, even after controlling for factors such as intelligence and working memory (Geary, 2011). Causal relations have also been established; experiences that improve the numerical magnitude knowledge of randomly assigned children also improve their subsequent learning of arithmetic and other mathematical skills (Booth & Siegler, 2008; Laski & Siegler, 2007; Siegler & Ramani, 2009; Whyte & Bull, 2008).

Numerical magnitude estimation gradually becomes more precise over several years (Booth & Siegler, 2006; Laski & Siegler, 2007; Laski & Yu, 2014; Siegler & Opfer, 2003). Early numerical magnitude estimates increase logarithmically with the size of the number being estimated; only with age and experience do individuals begin to generate linearly increasing estimates of magnitude, which reflect the formal organization of the number system (Siegler, Thompson, & Opfer, 2009). For example, kindergartners consistently produce estimates on 0–100 number lines that are better fit by a logarithmic function than by a linear one (Ebersbach, Luwel, Frick, Onghena, & Verschaffel, 2008; Siegler & Booth, 2004; Thompson & Opfer, 2008). It is not until second grade that most children generate linearly increasing estimates for the 0–100 range (Geary, Hoard, Nugent, & Byrd-Craven, 2008; Laski & Siegler, 2007; Siegler & Booth, 2004) and not until fourth grade or later for the 0–1000 range (Booth & Siegler, 2006; Opfer & Thompson, 2008).

While children increasingly generate linear estimates rather than logarithmic ones, the logarithmic representation continues to exist and be used in some contexts. For example, the same child will often produce linearly increasing estimates on smaller numerical scales (e.g., 0–100) and logarithmically increasing ones on larger scales (e.g., 0–1,000; Laski & Yu, 2014; Siegler & Opfer, 2003). Even adults have been found to use a logarithmic representation of number in some contexts, such as when their attention is divided (Anobile, Cicchini, & Burr, 2012; Chesney & Matthews, 2013). Thus, just as individuals have "backup" strategies in arithmetic that they use on difficult problems, such as adults' reliance on decomposition rather than retrieval on complex addition problems (e.g., solving 4 + 9 by thinking, "4 + 10 is 14, 9 is 1 less than 10, so 4 + 9 = 13"; LeFevre, Sadesky, & Bisanz, 1996), the logarithmic representation seems to serve as a backup representation.

The numerical magnitude representation used at different ages on different tasks is well documented; however, the process by which individuals select between representations is less clear. We propose that the process of generating estimates of numerical magnitude may involve competition between representations of numerical magnitude. In any given instance, both the logarithmic and linear representations might be activated simultaneously and exert an influence on estimates of numerical magnitude. Because the logarithmic representation develops earlier and involves less formal knowledge of mathematics, there may be a natural bias toward it. This bias may lead the logarithmic representation to exert a greater influence or be given more weight during estimation tasks, particularly in unfamiliar numerical contexts or when cognitive load is high, which could interfere with the use of a linear representation. In this view, the process involved in number line estimation would be similar to the parallel activation models that have been proposed in other domains of cognition, such as bilingualism (e.g., Bechara, Damasio, Tranel, & Damasio, 1997; Bialystok, Craik, Green, & Gollan, 2009; McClelland & Rogers, 2003).

Inhibitory Control

Inhibitory control is the ability to override or suppress a response or way of thinking in favor of a more relevant one while completing a task (cf. Cohen, Dunbar, & McClelland, 1990; Dempster & Corkill, 1999; Posner & Rothbart, 2000). Theorists have identified it as one of three main executive processes, along with switching and updating, within a Supervisory Attentional System (SAS) that monitors whether highly activated representations and responses are appropriate or not to the current situation (e.g., Friedman & Miyake, 2004; Miyake et al., 2000; Norman & Shallice, 1986). Over the course of development, inhibitory control improves and becomes increasingly differentiated from the other executive control processes (Best & Miller, 2010; Lee, Bull, & Ho, 2013; Shing, Lindenberger, Diamond, Li, & Davidson, 2010). In the preschool years, inhibitory control is highly correlated to both updating and switching (Wiebe, Espy, & Charak, 2008). By early adolescence, the processes become only moderately correlated and are best represented by three distinct factors (Rose, Feldman, & Jankowski, 2011).

A large amount of evidence indicates that inhibitory control plays an important role in general mathematics performance and development (e.g., Blair & Razza, 2007; Bull, Espy, & Wiebe, 2008; Bull & Lee, 2014; Clark, Sheffield, Wiebe, & Espy, 2013; Espy et al., 2004; LeFevre et al., 2013; St Clair-Thompson & Gathercole, 2006; Siegler & Pyke, 2013). For example, Clark, Pritchard, and Woodward (2010) found that individual differences in inhibitory control at age 4 were predictive of individual differences in mathematics achievement at age 6. Similarly, Bull and colleagues have consistently found relations between young children's executive control and their mathematical skills, even after controlling for other predictive factors, such as IQ (Bull et al., 2008; Bull, Espy, Wiebe, Sheffield, & Nelson, 2011; Bull & Scerif, 2001). When the components of executive function are analyzed separately, only inhibitory control is associated with preschoolers' and kindergartners' general mathematical skills, above and beyond other executive functions and IQ (Blair & Razza, 2007; Espy et al., 2004; Kroesbergen, van Luit, van Lieshout, van Loosbroek, & van de Rijt, 2009). The relation between inhibitory control and mathematics achievement persists over development. Siegler and Pyke (2013) found relations between sixth and eighth graders' inhibitory control and general mathematics achievement.

Further, general measures of executive functioning are related specifically to the linearity of children's number line estimates (Fuchs et al., 2010; Geary et al., 2007). In addition, studies that have measured and analyzed inhibitory control separately from other executive functions also have found it is correlated to the linearity of young children's number line estimates (Friso-van den Bos, Kolkman, Kroesbergen, & Leseman, 2014; Kolkman, Kroesbergen, & Leseman, 2013). For example, Friso-van den Bos and colleagues (2014) measured 4- to 8-year-olds' inhibitory control using the Eriksen Flanker task, in which children are asked to press the button on the side of the screen a sheep is facing when other animals are facing the opposite direction. They found that individual differences in accuracy on the Flanker test was positively correlated to individual differences in the linearity of 0–10 number line estimates; better inhibition was associated with more linear estimates.

The studies to date, however, do not propose why inhibitory control is related to numerical estimation. Our hypothesis that the process of generating estimates of numerical magnitude may involve parallel activation of the logarithmic and linear representations and a competition between them serves as a possible theoretical explanation for the relation. According to this view, inhibitory control would be related to the linearity of number line estimates because individuals would need to suppress their bias toward a logarithmic representation to generate a linear one. Similarly, Geary and colleagues (2007, 2008) speculated that inhibitory control is involved in the construction of linear representations of magnitude because children must learn to suppress their bias toward early developing logarithmic representations.

Another issue that has not been specified in previous work is precisely what kind of inhibition might be involved in number line estimation. Studies that have examined the relation between inhibitory control and numerical knowledge have used a range of inhibitory control tasks. Some studies have used domain-specific measures that require inhibition of a numerical response (e.g., the number-quantity Stroop tasks, Bull & Scerif, 2001); whereas, others have used more domain-general measures (e.g., day/night Stroop task, Kolkman et al., 2013; or pressing the button on the side of the screen an animal is facing, Friso-van den Bos et al., 2014). Recent models of cognitive control suggest that both domain-specific and domain-general inhibition processes exist and that they may operate independently of each other, at least in adults (Egner, 2008). For instance, both neuroimaging and behavioral studies have found evidence of domain-specific inhibition in the emotional domain (Egner, Delano, & Hirsch, 2007; Soutschek & Schubert, 2013). Thus, it is possible that only tasks that require the inhibition of a numerical response may be related to inhibitory control. Given that inhibitory control becomes more differentiated from other executive processes with age (Lee, Bull, & Ho, 2013); however, it may be that there are developmental differences, with domain-specific control becoming more important for number line estimation with age and mathematical experience.

The Present Experiments

This study included two experiments, both designed to test the *interference hypothesis*—that to acquire and use linear representations, individuals must suppress interference from logarithmic representations. In Experiment 1, we tested the hypothesis with adults. The interference hypothesis posits that the linear and logarithmic representations of numerical magnitude coexist, are simultaneously activated in estimation tasks, and that there is a natural bias toward the earlier developing logarithmic representation; thus, the process of generating numerical magnitude estimates involves inhibition. If this is the case, then the relation between inhibitory control and number line estimation found in young children should also be present in adults.

Previous studies suggested that this relation was most likely to be observed in adults in cognitively challenging contexts. Adults, who generate linear estimates when asked to position sets of dots on a number line under normal conditions, generate logarithmic estimates when required to complete a number line estimation task while concurrently performing an attentionally demanding task (Anobile, Cicchini, & Burr, 2012). This result suggested that contexts with greater mental load reduce adults' ability to inhibit the logarithmic representation to which they are naturally biased, so that in these contexts, individuals' inhibitory control plays a greater role. Studies of adults' scientific reasoning and when they use earlier developing representations versus later developing ones were consistent with this interpretation. Kelemen and Rosset (2009) found that adults were more likely to rely on purpose-based explanations of scientific phenomena—a naïve bias generally associated with children—when asked to make judgments under timed conditions than under untimed conditions. Importantly for the purposes of the present study, individual differences in inhibitory control predicted the frequency with which adults used naïve explanations in the speeded condition, independent of their scientific knowledge.

To test the interference hypothesis in Experiment 1, we presented adults with two number line estimation tasks: one with standard endpoints involving powers of 10 (0 and 1,000) and one with nonstandard endpoints (364–1,364). A number line with nonstandard endpoints precludes the use of well-known landmarks (e.g., 500 on a 0–1,000 number line) and requires more complex calculation than when endpoints are powers of 10. Participants also completed two measures of inhibitory control: a color-word Stroop task as a measure of domain-general inhibitory control and a number-quantity Stroop task as a measure of domain-specific inhibitory control. Including both measures allowed us to examine whether general inhibitory control was related to the quality of individuals' estimates or, specifically, whether it was the ability to inhibit numerical information.

We had three predictions for Experiment 1. The first prediction was that the extra mental load imposed by number lines with nonstandard endpoints would reduce adults' ability to inhibit the logarithmic representations and thus lead them to produce estimates that conformed more closely to a logarithmic pattern than estimates on number lines with standard endpoints. A second prediction was that individuals with poorer inhibitory control would be more likely to generate estimates that conformed to a logarithmic pattern on the nonstandard number line estimation task than those with better inhibitory control. Our third prediction, based on recent work indicating domain-specific inhibition processes in adults (Egner, 2008), was that adults' performance on the number-quantity Stroop task would be more strongly correlated to their performance on the nonstandard estimation task than their performance on the color-word Stroop task.

In Experiment 2, we tested the interference hypothesis with children in the context of a training study with kindergartners. This approach allowed us to test whether the proposed process underlying estimation—parallel activation and suppression of representations—and whether the same kind of inhibition (domain-specific vs. domain-general) is similar at different developmental stages. Further, Experiment 2 explored whether individual differences in inhibitory control are related to the extent to which children benefit from instruction relevant to number line estimation. In other words, we aimed to test whether children who might be better at suppressing their bias toward the logarithmic representation more readily acquire a linear representation.

Our main predictions for Experiment 2 were that individual differences in children's inhibitory control are related to the rate at which children learn a linear representation during instruction and that this relation depends on the context in which learning occurs. These predictions were based on findings—which are discussed further in the introduction to Experiment 2— that an inability to inhibit misleading or inaccurate prior beliefs often interferes with learning (e.g., Bartolotti, Marian, Schroeder, & Shook, 2011; Guzzetti, Snyder, Glass, & Gamas, 1993; Ni & Zhou, 2005).

Experiment 1

Method

Participants and procedure. The experiment included 53 undergraduate students (24 males, 29 females) in varied academic programs (25% Human Development, 17% Elementary/Secondary Education, 8% Undecided, 6% Biology, 6% Economics, 4% Communications, and 34% Other) at a selective university. Participants were recruited through flyers posted around the university. The majority of students were sophomores (30%) and juniors (36%), and a minority were freshmen (15%) and seniors (19%), with an even distribution of genders within each group, $\chi^2(3) = 3.43$, $p = .330$ Students were asked to rate, on a 5-point scale, their comfort with math ("not at all comfortable" to "extremely comfortable") and their performance in math courses ("poor" to "excellent"). Mean ratings of comfort with math fell between "somewhat comfortable" and "fairly comfortable" ($M = 3.69$, $SD = .98$) and ratings of performance in math fell between "average" and "good" ($M = 3.64$, $SD = .85$). Students also reported on the number of math courses they had taken in college, with 34% having taken 1 course, 34% having taken 2 courses, 12% having taken 3 courses, and 20% having taken 4 or more courses.

Participants met individually with an experimenter for a 15-min session in a laboratory testing room on campus. After providing written consent, participants completed two measures of number line estimation and of inhibitory control. The order in which participants completed the tasks was counterbalanced across participants.

Measures.

Inhibitory control. Participants completed two frequently used measures of inhibitory control, a *color-word Stroop* task and a *number-quantity Stroop* task (Friedman & Miyake, 2004; Stroop, 1935). The order of the two tasks was counterbalanced across participants. Each Stroop task involved three conditions (baseline, congruent, and incongruent), which were counterbalanced across participants, as were the 12 test trials within each condition. The trials were presented on a computer using the experimental software, Eprime. Participants were told to verbally respond as quickly as possible and to simultaneously hit the {enter} button to proceed to the next trial. Response time in milliseconds for each trial (i.e., trial onset to response) was recorded by the experimental software and accuracy was derived by reviewing audio recordings of participants' responses. On both tasks, the outcome measure was an interference score calculated by subtracting the mean reaction time (RT) on correct baseline trials from the mean RT (in milliseconds) on correct incongruent trials.

Color-word Stroop. The *color-word Stroop* task involves reading a word as quickly as possible when the color of the font is either congruent with word meaning (e.g., "YELLOW" displayed in yellow) or incongruent (e.g., YELLOW displayed in green). Words in the baseline condition were simple nouns (e.g., car), which were also displayed in different colors to control for the frequency of each color.

Number-quantity Stroop. The *number-quantity Stroop* task involves indicating the quantity of numerals that appear on a screen as quickly as possible when the quantity and numeral are congruent (e.g., 333) or incongruent (e.g., 33333). The baseline condition used asterisks, rather than numerals (e.g., ***).

Number line estimation. Two number line estimation tasks, *standard endpoint* (0–1,000) and *nonstandard endpoint* (364−1,364), were used to measure number line estimation ability and the cognitive representations that inform the estimates. On each item, participants were asked to estimate the position of a numeral on a 20-cm number line with only the endpoints labeled. Previous estimates were not visible on later trials; each trial was estimated on a separate sheet of paper. The order of the two tasks was counterbalanced across participants. On each task, participants were presented 26 trials in a random order unique to that participant.

Standard endpoint number line estimation. The standard endpoint number lines had "0" on the left end and "1,000" on the right end. The numbers presented were: 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.

Nonstandard endpoint number line estimation. The nonstandard endpoint number lines had "364" on the left end and "1,364" on the right end. The 26 test trials for this task were generated by adding 364 to each number presented on the standard endpoint task. Thus, the correct position of the trials on the number line was the same for both tasks.

Results

First, we describe adults' performance on the standard and nonstandard number lines. Given evidence of gender differences in mathematics performance at the college level (Ganley & Vasilyeva, 2014; Lindberg, Hyde, Petersen, & Linn, 2010), we examined the potential role of gender in individual differences in estimation. Second, we report the relation between inhibitory control and the fit of the logarithmic function on nonstandard number lines.

Number line estimation. Descriptive statistics for adults' estimates on standard and nonstandard number lines are displayed in Table 1. To test our first prediction that estimates would tend to be more logarithmic on nonstandard number lines than standard ones, we first examined the fit of the linear and logarithmic functions to the median estimates on each kind of number line. The linear function accounted for a greater amount of variance in adults' median estimates than the logarithmic function on both number lines, as indicated by dependent means t tests: on nonstandard number lines, $R^2_{\text{lin}} = .99$ and $R^2_{\text{log}} = .98$, $t(34) = -9.52$, $p < .001$, $d = 1.05$, and on standard number lines, $R^2_{\text{lin}} = 1.0$ and $R^2_{\text{log}} = .65$, $t(34) = -4.77$, $p < .001$, $d = 2.23$. However, as expected, a dependent means t test showed that adults' median estimates on

Table 1
Descriptive Statistics for Experiment 1

	M	*SD*
Standard number line		
Linear R^2	0.77	0.40
Logarithmic R^2	0.51	0.27
PAE	0.04	0.01
Nonstandard number line		
Linear R^2	0.93	0.07
Logarithmic R^2	0.91	0.07
PAE	0.06	0.03
Inhibitory control[a]		
Color-word Stroop	58.77	256.98
Number-quantity Stroop	60.41	203.42

Note. PAE = percent absolute error.
[a] Interference scores (milliseconds).

nonstandard number lines fit the logarithmic function much better than did their median estimates on standard number lines, $R^2_{\text{log nonstandard}} = .98$ versus $R^2_{\text{log standard}} = .65$, $t(34) = 7.46$, $p < .001$, $d = 1.79$. Further, adults' median estimates were somewhat less accurate on nonstandard number lines. The average percent absolute error [PAE = (|estimate-estimated quantity|/scale of estimates) × 100] of estimates on nonstandard number lines was 3% versus 2% on standard lines, $t(34) = 2.31$, $p = .027$, $d = 0.59$.

Analyses of individuals' estimates on each kind of number line revealed the same pattern: adults' estimates on nonstandard number lines conformed more closely to a logarithmic pattern and were less accurate than their estimates on standard number lines. A 2 (estimation task: nonstandard vs. standard) × 2 (gender: male vs. female) mixed-design multivariate analysis of variance (MANOVA), with logarithmic fit of estimates and PAE as the two dependent variables, found a main effect of estimation task, $F(2, 50) = 140.05$, $p < .001$, $\eta^2_p = 0.85$, a main effect of gender, $F(2, 50) = 12.44$, $p < .001$, $\eta^2_p = 0.33$, and an interaction between estimation task and gender, $F(2, 50) = 10.01$, $p < .001$, $\eta^2_p = 0.29$. The main effect of estimation task indicated that the best-fitting logarithmic function accounted for greater variance on the nonstandard task than the standard task: nonstandard number lines, $R^2_{\log} = .91$ ($SD = .07$), and standard number lines, $R^2_{\log} = .51$ ($SD = .27$). Adults' accuracy was also lower on nonstandard number lines, PAE = 6% (SD = 2%), than on standard number lines, PAE = 4% (SD = 1%). The main effect of gender indicated that, across estimation tasks, the best-fitting logarithmic function accounted for greater variance in females' estimates ($R^2_{\log} = .77$, $SD = .04$) than in males' estimates ($R^2_{\log} = .64$, $SD = .16$); similarly, females were less accurate (PAE = 6%, SD = 2%) than males (PAE = 4%, SD = 2%).

To interpret the interaction effect, tests of simple effects were conducted in which we compared differences in performance between estimation tasks separately for males and females. Both males and females' estimates were more logarithmic and less accurate on nonstandard number lines than standard ones, ps < .001, but the magnitude of the decrement in performance differed for males and females. In relation to the variance accounted for by the best fitting logarithmic function, the difference between the nonstandard and standard estimation tasks was greater for males (nonstandard $R^2_{\log} = .95$, $SD = .05$ vs. standard $R^2_{\log} = .36$, $SD = .34$) than for females (nonstandard $R^2_{\log} = .90$, $SD = .08$ vs. standard $R^2_{\log} = .64$, $SD = .03$). For PAE, the difference between estimation tasks was greater for females (nonstandard PAE = 7%, SD = 3% vs. standard PAE = 4%, SD = 1%) than for males (nonstandard PAE = 5%, SD = 2% vs. standard PAE = 4%, SD = 2%). Because of the unusual pattern of results we suspected that these gender differences were not necessarily meaningful. Thus, to explore this finding further, we compared the magnitude of decrement in each measure of performance between males and females. We first converted the $R^2_{\log}$ and PAE to z scores to allow for comparison, then we computed a score for each measure of the difference between the standard and nonstandard number lines (i.e., scores on the standard estimation task were subtracted from scores on the nonstandard estimation task). An independent samples t test indicated that the magnitude of the decrement in $R^2_{\log}$ among males did not differ from the decrement in PAE among females, $t(37.71) = 1.34$, $p = .188$.

In summary, the results were consistent with first our prediction: adults' estimates were more likely to conform to a logarithmic pattern and be less accurate on nonstandard number lines compared with standard number lines.

Inhibitory control. An interference score was calculated for each Stroop task by subtracting the mean RT on correct baseline trials from the mean RT on correct incongruent trials. Larger differences in times under the two conditions indicate lower levels of inhibitory control. As can be seen in Table 1, adults' average levels of performance on the Stroop tasks and the spread of the scores on the two tasks was similar. The mean number-quantity Stroop interference score was 60.41 ms ($SD = 202.42$) and the mean color-word Stroop interference score was 58.77 ms ($SD = 256.98$). There were no gender differences in the interference scores on either measure: number-quantity Stroop task ($M_{\text{males}} = 73.96$, $SD = 204.30$ vs. $M_{\text{females}} = 49.19$, $SD = 205.62$, $t(51) = 0.44$, $p = .663$) and color-word Stroop task ($M_{\text{males}} = 53.55$, $SD = 242.04$ vs. $M_{\text{females}} = 63.08$, $SD = 272.91$, $t(51) = -0.13$, $p = .895$).

Inhibitory control as a predictor of the fit of logarithmic functions. Next, we tested our predictions that (a) individuals' inhibitory control would predict the extent to which they rely on a logarithmic representation of numerical magnitude on cognitively demanding estimation tasks and (b) adults' performance on the number-quantity Stroop task would be more strongly correlated to their performance on the nonstandard estimation task than their performance on the color-word Stroop task.

We conducted separate regression analyses using color-word Stroop and number-quantity Stroop scores as predictors of the fit of the logarithmic function and the percent absolute error of adults' estimates on nonstandard number lines. Separate analyses were conducted because collinearity between the measures would lead to unreliable estimates of the coefficient (Belsey, Kuh, & Welsch, 2004; Pedhazur, 1997). Because individuals may differ in use of a linear or logarithmic function regardless of inhibitory control and task difficulty, we controlled for the fit of the logarithmic function to individuals' estimates on standard number lines. Similarly, regression analyses using percent absolute error controlled for individuals' accuracy on standard number lines. Gender was not included in the regressions because the analyses above indicated no meaningful differences on the nonstandard estimation task and no differences on the measures of inhibitory control.

The results were consistent with both predictions—adults with poorer inhibitory control were more likely to generate estimates that conformed to a logarithmic function, but the presence of the relation depended on which inhibitory control measure was included as a predictor. The color-word Stroop task interference score did not predict either the fit of the logarithmic function on nonstandard number lines, $b = 0.0004$, $t(50) = 1.16$, $p = .251$, $r = .16$, or adults' percent absolute error on them, $b = 0.002$, $t(50) = 1.27$, $p = .210$, $r = .18$.

In contrast, as illustrated in Table 2, the number-quantity interference score significantly predicted the amount of variance accounted for by the best-fitting logarithmic function in adults' estimates on nonstandard number lines. This measure of inhibitory control explained 21% of the variance in the fit of the logarithmic function, over and above individuals' natural propensity to use a logarithmic function (i.e., the fit of the logarithmic function to estimates on the standard task). Specifically, an increase of 10 points in interference scores is associated with an increase of 2% in the fit (R^2) of the logarithmic function, $b = 0.0002$, $t(50) = 3.74$, $p < .001$, $r = .47$. Similarly, a greater number-quantity interference score was related to less accurate estimates. The number-quantity interference score accounted for nearly 6% of the variance in the percent absolute error of adults' estimates on the nonstandard number lines above and beyond their percent absolute error on standard number lines, $b = 0.003$, $t(50) = -1.96$, $p = .055$, $r = .27$.

To test for differences in the predictive power of the number-quantity interference scores versus the color-word interference scores on the fit of a logarithmic function to adults' estimates and their percent absolute error, we conducted Fisher's z transformations for correlated coefficients (Meng, Rosenthal, & Rubin, 1992) using the effect size rs obtained from the two regression coefficients, respectively. For the fit of adults' estimates to a logarithmic function, number-quantity Stroop scores were a significantly stronger predictor than color-word Stroop scores, $z = 1.91$, $p = .028$. For PAE, there was no significant difference in the effects of the two Stroop tasks, $z = 0.53$, $p = .298$.

Discussion

The results of Experiment 1 were consistent with the hypotheses that motivated the experiment. First, adults generated different patterns of estimates depending on the difficulty of the number line estimation task. Although a linear function better fit participants' estimates on both standard and nonstandard tasks, the fit of a logarithmic function to participants' estimates was much greater on the nonstandard number line task. This finding supports the view that individuals possess multiple representations of numerical magnitude that may be simultaneously activated in estimation tasks and they have a tendency to increase their weighting of the logarithmic representation when confronted with difficult numerical tasks.

Second, adults' ability to inhibit prepotent numerical responses, as measured by the number-quantity Stroop task, was related to the degree to which their estimates fit the logarithmic function on number lines with nonstandard endpoints, even after controlling for the tendency to use a logarithmic representation on number lines with standard endpoints. On the other hand, performance on the color-word Stroop task was not related to the fit of the logarithmic function to adults' nonstandard number line estimates, after controlling for its fit on number lines with standard endpoints. These findings suggest that, at least for adults, the ability to inhibit prepotent numerical responses is more important than general inhibitory control. They also add to the evidence suggesting individuals are more likely to use sophisticated formal reasoning when they possess better inhibitory control. This interpretation suggested that individuals with better inhibitory control would more easily acquire a linear representation of magnitude. This hypothesis was tested in Experiment 2.

Experiment 2

This experiment was designed to test the part of the interference hypothesis that predicts that learning more advanced representations and strategies requires the inhibition of prior, less advanced ones. In the present context, this hypothesis implied that learning a linear representation of numerical magnitudes requires suppression of the logarithmic representation. Kindergartners played a 0–100 board game designed to promote a linear representation of numerical magnitudes in one of two conditions: *counting-on* from the number in the square of their current position on the game board (e.g., child on 5 who spun a 2 said, "6, 7") or *counting-from-one* the number of spaces moved (e.g., child on 5 who spun a 2 said, "1, 2"). Their inhibitory control was assessed at pretest using two Stroop tasks. To examine changes in the understanding of numerical magnitudes, a microgenetic design was used. Microgenetic designs involve frequent assessment of knowledge, making it possible to determine the rate and path of change (Siegler, 2006). Kindergartners played the numerical board game described in Laski and Siegler (2014) eight times and completed a 0–100 number line estimation task at pretest, after playing the game four

Table 2
Predictive Relations Between Adults' Number-Quantity Stroop Task Performance and R^2_{log} and PAE on Nonstandard Number Lines, Experiment 1

	b	SE	t	$R^2\Delta$	R^2
Fit of the logarithmic function					
1. R^2_{log} standard task	−0.032	.032	−1.01	0.046	0.046
2. Number-quantity Stroop	0.0002***	<.001	3.74	0.208	0.254
Percent absolute error (PAE)					
1. PAE standard task	0.751	.21	3.52	.210	.210
2. Number-quantity Stroop	−0.003†	<.001	1.963	.056	.266

† $p < .10$. *** $p < .001$.

times, after playing the game eight times, and at posttest 2 weeks later.

The experiment had three purposes. The first was to test the hypothesis that inhibitory control would be related to the rate of improvement in the linearity of children's number line estimates. The logic underlying this prediction was that children must suppress their bias toward a logarithmic representation of magnitudes to benefit from the linear cues on the game board. This idea is consistent with the results of studies across many domains that have found that children often bring prior approaches and representations to learning tasks which, in some cases, interfere with learning (e.g., Bartolotti et al., 2011; Guzzetti et al., 1993; Ni & Zhou, 2005). More specifically, it is consistent with recent evidence that activation of children's logarithmic representations at the onset of instruction decreased learning about numerical magnitudes. Children who generate logarithmic estimates before receiving instruction benefit less than those who generate either linear ones or those who are not asked to generate estimates before receiving instruction (Booth & Siegler, 2008; Opfer & Thompson, 2008).

The second purpose was to test whether the extent to which inhibitory control is related to learning depends on the context in which learning occurs. The two board game conditions (count-from-one and count-on from the larger addend) used in the present study differed in ways that allowed us to test this hypothesis. In both conditions, the game boards had cues to the linearly increasing magnitude of numbers (i.e., equal size spaces between numbers). When children use a count-on procedure, however, they are more likely to attend to these cues (Laski & Siegler, 2014); thus, there is more information available to them that contradicts a logarithmic representation. In contrast, the count-from-one procedure affords fewer opportunities for children to attend to the linear cues; the repetition of smaller numbers on each turn may actually accentuate a logarithmic representation. This analysis led to the specific prediction that inhibitory control would be more important for learning in the count-from-one condition, where instruction provided less contradictory information.

The third purpose was to examine whether there are developmental differences in the kind of inhibition involved in number line estimation. To parallel Experiment 1, kindergartners' inhibitory control was assessed using both a domain-general (i.e., color-shape) Stroop task and domain-specific (i.e., number-quantity) Stroop task. There was no clear prediction. On one hand, it was possible that domain-specific inhibition would be more strongly related to the rate of improvement in the linearity of children's number line estimates like adults. Bull and Scerif (2001) found that 7-year-olds' interference score on a number-quantity Stroop task was more strongly correlated to their mathematics achievement than their interference score on a color-word Stroop task. On the other hand, given that inhibitory control becomes more differentiated from other executive processes with age (Lee, Bull, & Ho, 2013), it was possible that domain-general inhibition would be more strongly related to improvements in linearity or that there would be no differences between the two kinds of inhibition.

Method

Participants and procedure. Participants included 42 kindergartners (mean age = 5.8 years, *SD* = 3.9 months) recruited from two charter schools serving low- to lower-middle income families. The percentages of children eligible for the free or reduced lunch program in the two schools were 93% and 55%, respectively. Twenty-four of the children were male and 18 were female; the mean age was equivalent across genders.

Children within each school were randomly assigned to one of two conditions: *count-from-one* or *count-on*. The count-from-one condition included 21 children (mean age = 5.80, 33% female, 33% Black, 52% White, and 15% Other). The count-on condition also included 21 children (mean age = 5.80, 52% female, 48% Black, 52% White, and 0% Other). Children met individually with an experimenter during the spring of the academic year in a quiet area of their school. Based on the NCTM and Common Core standards for mathematics, children at this point in the year had been exposed to counting to 100 by both ones and tens as well as counting and estimating sets including up to 20 objects (National Council of Teachers of Mathematics, 2006; National Governors Association, 2010).

Each child met with the experimenter for two sessions per week for 3 weeks. In Session 1, children completed the inhibitory control measures and a pretest number line estimation task. During Sessions 2–5, children played the board game eight times, twice during each of the four training sessions. At the end of Sessions 3 and 5, children completed the number line estimation task after playing the game. These assessments of number line estimation allowed us to examine rate of learning over the training sessions. In Session 6, children completed a posttest measure of number line estimation. All sessions were videotaped.

Board game training conditions. All children played a board game called *Race to Space*. The game board had the numbers 1–100 arranged in a 10 × 10 matrix. The blue background color of the board became darker every two rows, as the numbers on the game board increased, providing an added cue to numerical magnitude. The spinner, which determined how far participants would move their tokens on each turn, had five sections labeled 1–5.

Children participated in one of two experimental conditions: count-from-one and count-on. The conditions differed in what children were instructed to say as they moved their token across spaces on the game board. In the count-from-one condition, children counted aloud from 1 as they moved their token, until they reached the number indicated on the spinner (e.g., children on 17 who spun a 2 said "1" as they put their token on the square labeled 18 and "2" as they put their token on the square labeled 19). In the count-on condition, children counted on from the number labeled in the square where they began the turn (e.g., children who began a turn on 17 and spun a 2 said "18, 19"). The same experimenter ran all participants, and did so within a relatively short time period (January to June), following a scripted protocol for each condition.

In both conditions, if the child could not perform the requested activities, the experimenter provided assistance. The videos of the training sessions were coded to determine the amount of assistance provided by the experimenter. Raters coded instances in which the experimenter demonstrated how to keep track of how many spaces a child moved the token, provided the name of numerals on the board, reminded the child how to play the game, or asked questions or prompted to promote children's thinking about the game.

A subset (10%) of games was coded by two raters: $K = .88$, $p < .01$.

Measures.

Inhibitory control. Participants completed two measures of inhibitory control that paralleled the tasks in Experiment 1: shape-color Stroop task and number-quantity Stroop task. The order of the two tasks was counterbalanced across participants.

Shape-color Stroop. In the shape-color Stroop task (adapted from Espy, Bull, Martin, & Stroup, 2006), three conditions were presented in a fixed order: a control condition (baseline) and two conditions requiring inhibitory control (response suppression and rule-switching). Each condition had 12 trials arranged in three lines of four on a page and presented in the same random order for every child. In the baseline condition, the child was told to name the stimulus figures (colored squares and circles with cartoon faces, legs, and arms) by saying their colors as quickly as possible. The shapes were comprised of an equal number of four colors: red, blue, yellow, and green. In the response suppression condition, children were instructed only to name the colors of the figures with happy faces; this condition required children to say the colors of six figures and suppress their reaction to say the colors of the other six. In the rule-switching condition, children were shown stimulus figures that had hats or no hats. They were told that the names of figures with hats were the figure shapes, while the names of figures without hats were their colors, and were asked to name the shape or color of all the figures as quickly as possible. In all conditions, children received practice with feedback before proceeding to the test trials. Accuracy of naming the shapes in accordance with the given rule was recorded.

Number-quantity Stroop. The number-quantity Stroop task was adopted from Bull and Scerif (2001). Children were instructed to name the quantity of items (one, two, three, or four) as quickly as possible, and allowed to self-correct their responses. Three conditions, each with 12 trials, were presented in a counterbalanced order across individuals: baseline, congruent, and incongruent. Baseline trials consisted of triangles (e.g., ▲▲▲). On the congruent trials, the quantity and printed numeral corresponded (e.g., 3 3 3) and on the incongruent trials they did not (e.g., 2 2 2). As in Experiment 1, trials were presented on a computer using Eprime and response time was calculated from the time a trial appeared on the screen until the student provided a correct response.

Number line estimation. The number line estimation task was identical to the one administered to adults in Experiment 1, except that the endpoints were 0 and 100 and there were only 22 trials. The numbers presented were 2, 3, 5, 8, 12, 17, 21, 26, 34, 39, 42, 46, 54, 58, 61, 67, 73, 78, 82, 89, 92, and 97. A different random order of the numbers was generated for each child.

Results

First, we describe performance on the inhibitory control measures; then, we examine the relation between inhibitory control and rate of learning across conditions; finally, we compare that relation between conditions to test whether the extent to which inhibitory control is related to learning depends on the context in which learning occurs. Preliminary analyses indicated no gender differences on any of the outcome measures; thus, all analyses were collapsed across genders.

Inhibitory control. Performance on the shape-color Stroop task was measured using an inhibitory cost score: Inhibitory Cost = |[percent correct in inhibition conditions—percent correct in control condition]|. As shown in Table 3, the mean inhibitory cost score across conditions was 8%. There was no difference in the mean score among children in the count-from-one and count-on conditions ($M = 12\%$, $SD = 16.87\%$, and $M = 6\%$, $SD = 10.91\%$), $t(40) = 1.18$, $p = .246$.

Table 3
Descriptive Statistics for Experiment 2

	Across conditions		Count-from-one		Count-on	
	M	*SD*	*M*	*SD*	*M*	*SD*
Log R^2						
Pretest	0.61	0.25	0.62	0.20	0.60	0.30
Training 2	0.59	0.22	0.55	0.23	0.64	0.21
Training 4	0.63	0.21	0.56	0.22	0.69	0.19
Posttest	0.63	0.22	0.55	0.26	0.71	0.15
Linear R^2						
Pretest	0.48	0.26	0.46	0.23	0.49	0.29
Training 2	0.55	0.29	0.46	0.28	0.64	0.28
Training 4	0.58	0.30	0.47	0.30	0.69	0.25
Posttest	0.58	0.28	0.46	0.29	0.69	0.22
PAE						
Pretest	0.20	0.07	0.21	0.07	0.20	0.08
Training 2	0.18	0.08	0.20	0.08	0.15	0.07
Training 4	0.17	0.07	0.19	0.08	0.14	0.06
Posttest	0.17	0.07	0.19	0.07	0.14	0.07
Inhibitory control						
Shape color Stroop[a]	8.93	14.27	11.51	16.87	6.35	10.91
Number quantity Stroop[b]	12.09	9.72	12.38	9.06	11.79	10.55

Note. PAE = percent absolute error.
[a] Inhibitory cost score based on percent correct. [b] Interference score (seconds)

Performance on the number-quantity Stroop task was measured with an interference score based on response time on correct trials only: Interference = incongruent condition mean RT—baseline condition mean RT. The mean interference score was 1.01 (*SD* = .81) seconds (see Table 3); there was no difference in the mean score between children in the count-from-one and count-on conditions ($M = 1.03$, $SD = 0.76$, and $M = 0.99$, $SD = 0.87$, respectively), $t(40) = 0.21$, $p = .836$. Scores on the two inhibition tasks were moderately correlated, $r(40) = .34$, $p < .05$.

Relations of inhibitory control and rate of learning across conditions. We used multilevel growth modeling to test the predictions that (a) inhibitory control would be related to the rate of improvement of children's number line estimates and (b) that developmental differences in the kind of inhibition involved in number line estimation may exist. Multilevel growth modeling was used to calculate the rate of improvement in the linearity of estimates and PAE over the four time points at which children completed the number line estimation task: pretest, after Session 2, after Session 4, and posttest (see Table 3). The growth modeling procedure began by determining the functional form of growth across time. This step entailed running two unconditional growth models, one linear and one quadratic. The linear model included no predictors other than a variable representing *time* at Level 1; in parallel, the quadratic model included only *time* and $time^2$. In both models, intercepts and slopes were allowed to vary randomly. The results showed that the fixed and random effects for the two models were similar, but the fit of the linear model, as estimated by −2 Restricted Log Likelihood, was significantly better than the fit of the quadratic model, $\chi^2(1) = 9.80$, $p < .01$. Therefore, we adopted the linear model for the main analyses.

To examine which kind of inhibition was involved in children's number line estimation, the linearity of estimates and PAE were analyzed separately in two parallel sets of linear growth models. As with Experiment 1, because collinearity between the two measures of inhibitory control would lead to unreliable parameter estimates, the measures were used as predictors in separate models. Thus, in one set of models color-shape Stroop inhibitory cost scores (Equation 1) were used to predict linearity and PAE and in the other set number-quantity Stroop interference scores (Equation 2) were used to predict linearity and PAE, resulting in four models.

$$\begin{aligned} Y_{ti} &= \pi_{0i} + \pi_{1i}(time_{ti}) + e_{ti} \\ \pi_{0i} &= \beta_{00} + \beta_{01}(age_i) + \beta_{02}(colorstroop_i) + r_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11}(age_i) + \beta_{12}(colorstroop_i) + r_{1i} \end{aligned} \quad (1)$$

$$\begin{aligned} Y_{ti} &= \pi_{0i} + \pi_{1i}(time_{ti}) + e_{ti} \\ \pi_{0i} &= \beta_{00} + \beta_{01}(age_i) + \beta_{02}(numberstroop_i) + r_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11}(age_i) + \beta_{12}(numberstroop_i) + r_{1i} \end{aligned} \quad (2)$$

In Equations 1 and 2, β_{00} refers to the average level of linearity of estimates or PAE at pretest, and β_{01} represents the association between the predictor (shape-color Stroop in Equation 1 and number-quantity Stroop in Equation 2) and linearity of estimates or PAE at pretest. In both models, a variable representing time was included as a predictor at Level 1, which produced an estimate of the average slope of children's improvement across the four time points. The slope estimate is denoted by β_{10}, and the coefficient β_{11} represents variation in slopes as a function of inhibitory control. The time variable was coded so that the *y*-intercept represented children's average performance at pretest (i.e., pretest/Time 1 = 0, Time 2 = 1, Time 3 = 2, posttest/Time 4 = 3). Thus, the estimate of the slope (i.e., estimate of average rate of improvement) controlled for pretest linearity or accuracy. By controlling for pretest linearity, the aim was to control for any unmeasured variables related to improvements in children's performance across training, such as IQ or general mathematical ability. We also included the child's age as a covariate to capture any individual differences in age-related cognitive abilities that might not have been captured by the child's pretest performance and inhibitory control. The intercept and slope in each model were allowed to vary randomly, which yielded estimates of the proportion of variance in individual children's intercepts (i.e., pretest number line estimation linearity and PAE), and slopes (i.e., average rate of improvement), that could be explained by inhibitory control. Estimates of variance explained in intercepts and slopes are denoted by r_{0i} and r_{1i}, respectively. The term e_{ti} represents variation in individual children's scores.

The two models that included the number-quantity interference scores found that children's performance on this Stroop task was not related to the rate of their improvement on number line estimation (for linearity of estimates, $b = 0.0001$, $t(39) = 0.11$, $p = .910$, $r = .02$ and for PAE, $b = -0.0004$, $t(39) = -1.33$, $p = .192$, $r = .21$). One potential explanation is that, among young children, the number-quantity Stroop task measures numerical knowledge in addition to inhibitory control. Thus, individual differences on this task would already be captured by pretest number line estimation performance. Indeed, performance on this task positively correlates with mathematics performance among children between the ages of 6 and 8 years (Bull & Scerif, 2001). Similarly, in the present study, interference scores on the number-quantity Stroop task were related to both pretest linearity ($b = -0.01$, $t(39) = -2.42$, $p = .020$, $r = .36$) and percent absolute error ($b = 0.004$, $t(39) = 3.39$, $p = .002$, $r = .48$).

In contrast, the results of the two models that included color-shape Stroop performance as a predictor were consistent with our prediction: children's inhibitory control was related to the rate at which children's understanding of numerical magnitude improved from playing the board game. The analysis revealed that playing the game led to improvement in linearity, $b = .03$, $t(39) = 3.24$, $p = .002$, $r = .46$, but the effect depended on children's inhibitory cost scores. The rate of improvement (i.e., the estimated slope of the improvement over the four time points) in the linearity of estimates among children with poorer inhibitory control (those with an inhibitory cost *z*-score that was at least 1 *SD* above the mean) was 0.06 points less than the rate observed among children with better inhibitory control (those with an inhibitory cost *z*-score that was at least 1 *SD* below the mean), $b = -0.03$, $t(39) = -2.81$, $p = .008$, $r = .41$. Random effects estimates revealed that children tended to vary reliably in their rate of growth across time, $\hat{\lambda} = .422$, $\chi^2 = 70.90$, $p = 003$, and inhibitory cost scores explained 31% of this variance. At pretest (i.e., the model intercept), the linearity in children's estimates varied reliably, $\hat{\lambda} = .93$, $\chi^2 = 558.79$, $p < .001$, but inhibitory cost scores only explained 10% of this variance, and were not significantly predictive of pretest linearity, $b = .05$, $t(39) = -1.25$, $p = .220$, $r = .20$. Children's age was related to the linearity of estimates at pretest, $b = 0.08$, $t(39) = 2.05$, $p = .047$, $r = .31$, but was not related to their rate of improvement over time, $b = -0.004$, $t(39) = -0.411$, $p = .683$, $r = .07$.

The model examining the rate of improvement in number line estimation accuracy provided congruent evidence that inhibitory control was predictive of improvement in number line estimation. Children's PAE decreased significantly over time, $b = -0.01$, $t(39) = -5.15$, $p < .001$, $r = .64$, but the rate (i.e., the estimated slope) of improvement varied as a function of inhibitory cost scores. The PAE of children with an inhibitory cost score that was 2 *SD* above the mean increased about 1% points more than children with an inhibitory cost score that was 2 *SD* below the mean, $b = -0.006$, $t(39) = 2.63$, $p = .029$, $r = .39$. Random effects estimates showed that there was reliable variability in the rate at which PAE decreased over time, $\hat{\lambda} = .37$, $\chi^2 = 64.64$, $p = .011$, although inhibitory cost scores explained 30% of this variance. At pretest, there was significant variability in the PAE of children's number line estimates, $\hat{\lambda} = .87$, $\chi^2 = 341.32$, $p = <.001$, but inhibitory cost scores explained none of this variance, and were not significantly predictive of pretest performance, $b = 0.02$, $t(39) = 1.35$, $p = .184$, $r = .21$. Age was marginally related to pretest differences in accuracy, $b = -0.02$, $t(39) = -1.88$, $p = .068$, $r = .29$, but was unrelated to improvements in accuracy over time, $b = -0.0007$, $t(39) = 0.288$, $p = .775$, $r = .05$.

To test for differences in the predictive power of the shape-color Stroop scores versus the number-quantity Stroop scores, we conducted Fisher's *z* transformations (adjusting for the correlation between the two Stroop tasks) on the effect size *r*s obtained from the two regression coefficients, respectively. For linearity of children's estimates, shape-color Stroop scores were a significantly stronger predictor of the rate of improvement over time than number-quantity Stroop scores, $z = 1.84$, $p = .039$. For PAE, there was no significant difference in the effects of the two Stroop tasks, $z = 0.88$, $p = .189$.

In summary, children's inhibitory control, but only when measured by the shape-color Stroop task, was related to the rate at which children's understanding of numerical magnitude improved from playing the board game. Figure 1 illustrates the patterns of learning of children with better versus poorer inhibitory control (shape-color cost scores below the median vs. above it). Children with below average inhibitory control (relative to the sample) demonstrated little to no improvement across the training sessions in the linearity and PAE of their number line estimates. In contrast, children with above-average inhibitory control demonstrated rapid and substantial improvement.

Relations of inhibitory control and rate of learning within conditions. To test whether the extent to which inhibitory control is related to learning depends on the context in which learning occurs and, more specifically, the prediction that inhibitory control would be more important for learning in the count-from-one condition than in the count-on condition, we examined the relation between inhibitory control and rate of learning in each condition. Because number-quantity interference scores were not related to the rate of learning across the sample, only shape-color inhibitory cost scores were used in these analyses. Equation 3 shows the model that was tested. In parallel to the previous analysis, we ran this model twice—once with the linearity of children's estimates as the outcome and once with PAE as the outcome.

$$\begin{aligned} Y_{ti} &= \pi_{0i} + \pi_{1i}(time_{ti}) + e_{ti} \\ \pi_{0i} &= \beta_{00} + \beta_{01}(age_i) + \beta_{02}(condition_i) + \beta_{03}(colorstroop_i) \\ &\quad + \beta_{04}(conditionXcolorstroop_i) + r_{0i} \\ \pi_{1i} &= \beta_{10} + \beta_{11}(age_i) + \beta_{12}(condition_i) + \beta_{13}(colorstroop_i) \\ &\quad + \beta_{14}(conditionXcolorstroop_i) \end{aligned} \quad (3)$$

In building this model, we found that random variance in the rate of children's learning was nonsignificant ($p = .123$ for linearity of estimates; $p = .113$ for PAE) and less reliable once we added the condition variable to the model. Therefore, we fixed the model slopes (i.e., we removed the error term r_{1i}) when we added the three-way interaction term between time, condition, and inhibitory control.

We first report results pertaining to improvement in the linearity of children's estimates over time. In contrast to the previous analysis, there was no overall improvement in the linearity of children's estimates across the four time points once the condition variable was included in the model, $b = 0.003$, $t(121) = 0.302$, $p = .763$, $r = .03$. Instead, the rate of improvement over time varied by condition, with a significantly higher rate of growth in the count-on condition that in the count-from-one condition, $b = 0.059$, $t(121) = 3.364$, $p = .001$, $r = .29$. Specifically, the rate of growth for a child in the count-on condition was 6% higher, on average, than the rate of a child in the count-from-one condition. The results pertaining to inhibitory control from the previous analysis were replicated in this analysis: inhibitory control was not significantly related to pretest scores, $b = -0.002$, $t(37) = -0.667$, $p = .509$, $r = .11$, but was associated with more rapid improvement across conditions, $b = -0.002$, $t(121) = -3.08$, $p = .003$, $r = .27$, even after accounting for the effect of condition. Critically, however, this analysis revealed a marginally significant interaction effect between condition and inhibitory control, which suggested that the negative relation between inhibitory cost scores and the rate of learning was diminished in the count-on condition, compared with the count-from-one condition, $b = 0.03$, $t(121) = 1.683$, $p = .095$, $r = .15$. Specifically, a child in the count-on condition with an inhibitory cost score that was 1 *SD* above the mean (reflecting poor inhibitory control) grew at a rate that was 3% greater than a child in the count-from-one condition with the same inhibitory cost score. Though child age was marginally related to linearity of estimates at pretest, $b = 0.23$, $t(37) = 1.949$, $p = .059$, $r = .31$, it was not related to improvement over time, $b = -1.007$, $t(121) = -0.261$, $p = .794$, $r = .02$.

A similar pattern was found when examining improvement in PAE. Although there was overall improvement in PAE across the four time points (indicated by a reduction in PAE), $b = -0.006$, $t(121) = -2.086$, $p = .039$, $r = .19$, the rate of improvement varied by condition, with a significantly greater rate of improvement in the count-on condition than in the count-from-one condition, $b = -0.011$, $t(121) = -2.680$, $p = .008$, $r = .24$. Specifically, the rate of growth for a child in the count-on condition was 6% higher, on average, than the rate of a child in the count-from-one condition. Inhibitory control was not significantly related to pretest scores, $b =$

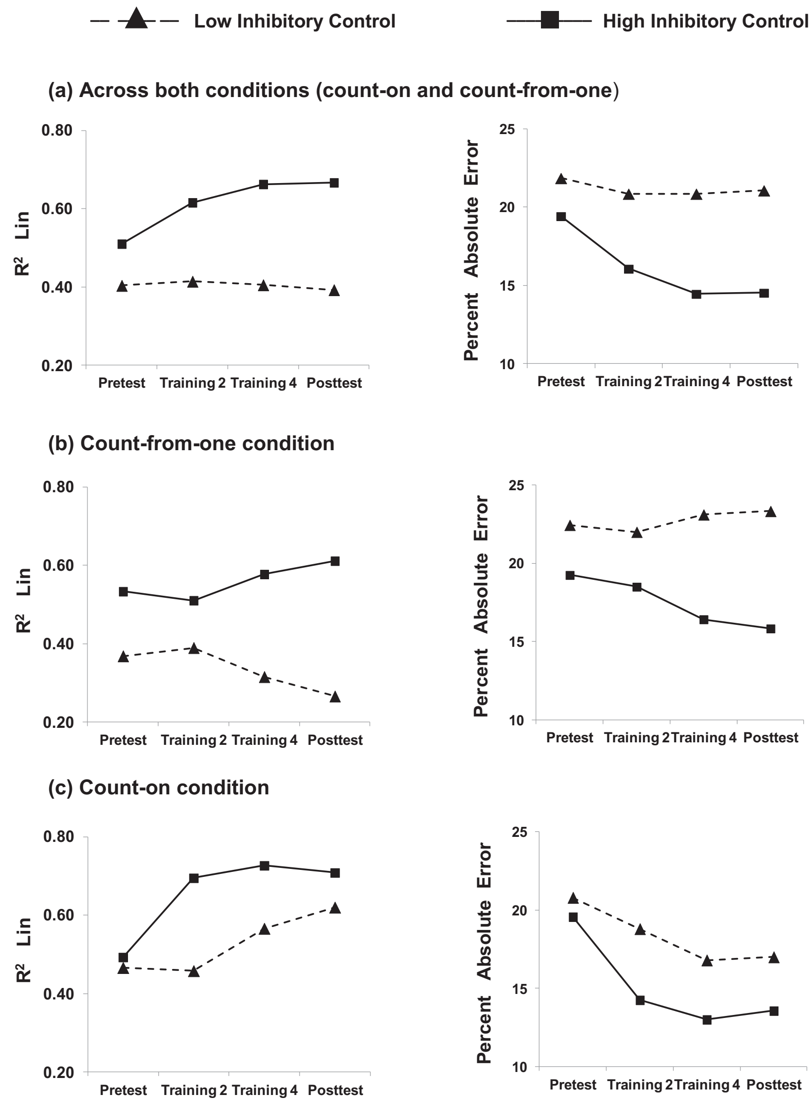

Figure 1. Experiment 2: Rate of improvement of number line estimation over training across and between board game training conditions for children with low versus high inhibitory control. Greater inhibitory control was associated with more rapid improvement in the count-on condition, but not in the count-from-one condition.

0.001, $t(37) = 1.216$, $p = .232$, $r = .20$, but was associated with more rapid improvement across conditions, $b = 0.001$, $t(121) = 2.619$, $p = .010$, $r = .23$, even after accounting for the effect of condition. The interaction effect between condition and inhibitory control did not reach significance, although the overall pattern was similar to that observed when predicting the linearity of children's estimates, $b = -0.005$, $t(121) = -1.291$, $p = .199$, $r = .12$. Age was not related to linearity of estimates at pretest, $b = -.052$, $t(37) = -1.609$, $p = .116$, $r = .26$, or to improvement over time, $b = 0.001$, $t(121) = 0.169$, $p = .866$, $r = .02$.

In summary, children in the count-on condition showed a faster rate of improvement than children in the count-from-one condition, both in the linearity of their estimates and their percent absolute error. Inhibitory control was related to rate of learning, even after accounting for the effect of condition. There was also evidence to suggest that although poor inhibitory control was associated with a lower rate of learning, the count-on condition buffered against this effect. This pattern of results is illustrated in Figure 1.

It was possible that inhibitory control was less related to learning in the count-on condition because the difficulty of the counting procedure elicited more assistance from the experimenter, offsetting poorer inhibitory control. To explore this possibility, the videos of the eight times that children played the game were coded for any instance of assistance. The frequency of instances of experimenter assistance was calculated by dividing the total instances of instruction each individual received by the number of games that were coded for each individual. Because of technical difficulties, a complete set of videos was not available for five participants.

The average amount of assistance received by participants in the count-on condition ($M = 27.54$, $SD = 22.77$) was higher than that received by participants in the count-from-one condition ($M = 3.45$, $SD = 2.90$), $t(19.59) = 4.70$, $p < .001$, $d = 0.73$. This difference, however, did not explain differences in the influence of inhibitory control between the conditions. Following the Barron and Kenny (1986) approach for testing moderation, a hierarchical multiple regression analysis, controlling for pretest linearity, indicated no moderation effect: the interaction of average amount of instructional support and shape-color Stroop interference did not account for any unique variance in children's posttest linearity above and beyond that accounted for by each measure alone, $b = 0.0002$, $t(36) = 1.58$, $p = .12$, $r = .26$. The same result was found using PAE as the outcome, $b = -0.00004$, $t(36) = -1.46$, $p = .15$, $r = .25$.

General Discussion

This study indicates that inhibitory control is an important process in number line estimation under some conditions for children and adults. The results were consistent with the hypothesis that ability to inhibit early developing logarithmic representations of numerical magnitude in favor of later developing more formal linear ones is important for mathematics performance and learning. Individual differences in inhibitory control predicted the quality of adults' and children's number line estimates, as well as the rate at which children's number line estimates improved from playing a numerical board game. In this concluding section, the potential implications of these findings for mathematics learning and instruction are discussed.

The Role of Inhibitory Control in Number Line Estimation

Individual differences in inhibitory control predicted differences in number line estimation. Adults' domain-specific inhibitory control, as measured by a number-quantity Stroop task, predicted the quality of their number line estimates on a number line with nonstandard endpoints (364–1,364) in Experiment 1. Kindergartners' domain-general inhibitory control, as measured by a color-shape Stroop task, predicted their estimates on 0–100 number lines in Experiment 2. Further, individual differences in children's domain-general inhibitory control predicted the rate of improvement in their number line estimates from playing a numerical board game, with a medium effect size across conditions ($r = .39$). This pattern of results, indicating the importance of inhibitory control in number line estimation, suggests that to use and acquire linear representations of magnitude, interference from the natural bias toward a logarithmic representation must be suppressed.

An interference effect in number line estimation would not be possible if linear representations of numerical magnitude replace logarithmic ones with age and experience. This study and previous ones, however, suggest that logarithmic representations remain, despite the acquisition of linear ones (Anobile et al., 2012; Laski & Yu, 2014; Siegler & Opfer, 2003). In this study, adults generated estimates that conformed more closely to a logarithmic pattern when the estimation task was difficult and less familiar. Similarly, most second graders who generate linear estimates on 0–100 number lines generate logarithmic ones on 0–1,000 number lines (Laski & Yu, 2014; Siegler & Opfer, 2003). We propose that the representations may not only coexist, but that the process of generating estimates of numerical magnitude in any given situation may involve parallel activation of both representations, competition between them, and a natural bias toward the logarithmic one, which leads it to exert greater influence and interference on difficult tasks.

While it is possible that even if the logarithmic representation remains, it does not create any interference, this seems unlikely. Interference from prior representations to later mathematics performance is common. For example, when children represent fractions, they often apply their understanding of whole numbers to the novel context of fractions (Ni & Zhou, 2005), such as judging fractions with larger denominators to be greater than fractions with smaller ones (e.g., claiming that [1/4] is greater than [1/2]). Similarly, children often solve mathematical equivalence problems incorrectly (e.g., $7 + 2 = 10 - \_\_$), because they rely on the more familiar operational patterns of basic arithmetic problems (e.g., $4 + 3 = 7$; McNeil, 2008). Even adults find it difficult to answer addition facts after having just completed a series of multiplication problems (Campbell & Timm, 2000). In all these instances, inhibitory control would be necessary to suppress related prior knowledge to produce more accurate performance. Thus, a plausible explanation for the relation between inhibitory control and the accuracy of number line estimates in the current study is that suppressing interference from the logarithmic representation is crucial to generating linear patterns of estimates.

An alternative explanation for the relations between inhibitory control and number line estimation is that individuals with poor inhibitory control have behavioral issues, which undermine engagement and attention and, thus, learning more generally (cf. Dempster & Corkill, 1999). Arguing against this interpretation, however, Experiment 2 provided children a brief one-to-one learning situation with minimal distractions and immediate response to inappropriate behavior, which left little possibility for individual differences in uncontrolled behavior or attention to affect learning. Further, the relation between inhibitory control and learning existed after controlling for pretest estimation performance; poor control of behavior or attention would presumably influence performance on both occasions.

It is also possible that the relations between inhibitory control and number line estimation were because of a factor not measured in the present study, such as IQ. In both experiments, however, we controlled for general mathematics ability. In Experiment 1, the relation between domain-specific inhibitory control and adults' tendency to use a logarithmic representation existed after controlling for their tendency to generate logarithmic patterns of estimates on standard number lines. In Experiment 2, domain-general inhibitory control explained variance in the linearity of children's estimates at posttest, above and beyond the linearity of their pretest estimates.

Implications for Mathematics Learning and Instruction

The present findings help explain individual differences in children's number line estimation. Kindergartners who have better domain-general inhibitory control generate more accurate and more linear number line estimates than do those with poorer inhibitory control. They also are more likely to benefit from experiences that promote linear estimates, such as a numerical board game. These results add to recent evidence that inhibitory control is involved in young children's mathematics learning (e.g., Blair & Razza, 2007; Bull et al., 2008; Clark et al., 2013; Espy et al., 2004; St Clair-Thompson & Gathercole, 2006).

The results also provide a potential explanation for the relation between inhibitory control and mathematics achievement: individuals with better inhibitory control may be better able to suppress the activation of prior knowledge and thus may be less vulnerable to interference from such knowledge. In the context of number line estimation, the data support the view that children need to suppress their bias toward a logarithmic representation to acquire and generate a linear representation.

If this kind of causal relation exists, then improving children's inhibitory control could be one mechanism for improving children's mathematics achievement. A future study should test whether children who receive inhibition training demonstrate greater improvement on number line estimation tasks than their peers who do not receive training. Recent studies indicate that young children's inhibitory control can be improved through repeated practice (Diamond, 2012; Kray & Ferdinand, 2013). The transfer effects of inhibition training, however, are quite narrow—sometimes not even transferring to other untrained inhibition tasks (Kray & Ferdinand, 2013). Thus, improvement in inhibitory control may be a necessary, but not sufficient factor, for improving number line estimation. Children may need simultaneous practice with relevant mathematics concepts, such as counting and arithmetic operations. Indeed, there is some evidence to suggest that increased experience with a domain improves individuals' ability to overcome domain-specific interference. For example, Bialystok and DePape (2009) found that adults with extended musical experience outperformed adults without musical experience on auditory Stroop tasks involving a word and pitch conflict.

Of interest to the authors, there were developmental differences in which inhibitory control task was most predictive of the accuracy of number line estimates, suggesting that domain-specific control becomes more important for number line estimation with age and mathematical experience. In Experiment 1, adults' performance on the number-quantity Stroop task predicted the quality of their number line estimates, but their performance on the color-word Stroop task did not. In Experiment 2, children's performance was predicted by the color-shape Stroop task in the count-from-one condition, but not the number-quantity one. One potential explanation is that the number-quantity Stroop task is unreliable for young children. Children under 8 years often do not automatically activate numerical magnitude representations when seeing Arabic numerals (Berch, Foley, Hill, & Ryan, 1999; Girelli, Lucangeli, & Butterworth, 2000; van Galen & Reitsma, 2008). Thus, the number-quantity Stroop task may not require inhibitory control for kindergarten children. The failure of the color-word Stroop task to predict adults' number line performance might indicate that with increased expertise in a subject, domain-general inhibitory control becomes less important, whereas the importance of domain-specific inhibitory control increases. This explanation would imply that young children's learning could benefit from generalized tasks designed to increase executive functioning, whereas improving older children's and adults' knowledge would require more specialized training. Future studies exploring this idea may be worthwhile.

The present findings also suggest that design of instruction might be able to alleviate differences related to inhibitory control and improve all children's learning and performance. In the present study, inhibitory control was more strongly related to the rate of children's improvement on number line estimation when they played a number board game by counting-from-one as they moved their token than when they played by counting-on from their current position. This effect was not explained by differences in the amount of assistance children in the two conditions received from the experimenter.

The key difference between the board game conditions might have been the extent to which the counting procedure provided information that contradicted the logarithmic representation. Studies of instructional approaches based on conceptual change theory have demonstrated the value of activating prior naïve knowledge in the learning process. More important, however, these studies have also demonstrated that without explicit instructional strategies that challenge the naïve knowledge (e.g., providing counter examples; requesting self-explanation) that knowledge, whether intentionally activated or not, may fail to promote and may even impede learning (e.g., Chi, 2000; Guzzetti et al., 1993; Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). For example, Groβe & Renkle (2007) found that college students asked to explain only incorrect examples demonstrated less improvement in understanding probability than those who explained only correct examples. In contrast, Durkin and Rittle-Johnson (2012) found that children benefitted from incorrect examples if the instruction directed them to why the example was incorrect. In their study, they required children to compare incorrect and correct examples rather than just presenting incorrect examples alone. They found that children who compared incorrect and correct examples showed greater improvements in placing decimals on a number line than those who compared two correct examples. Our results add to this previous work by suggesting it is children with better inhibitory control who are positioned to resist interference from naïve prior knowledge when the instruction does not offer supports to do so.

In the current context, the logarithmic representation would have been activated because of the numerical nature of the game for both groups. Experience playing the game using the count-on procedure provided more information that contradicted the logarithmic representation; whereas, using the count-from-one procedure provided less contradictory information and, in fact, may have reinforced the logarithmic representation. The count-on procedure facilitates attention to the linearly increasing magnitude of numbers (i.e., equal size spaces between numbers) present on the game board (Laski & Siegler, 2014). Thus, this context provides more information that highlights the differences between the linear and logarithmic representation and helps children to discount it, lessoning the need for an individual's own inhibitory control. On the other hand, the count-from-one procedure emphasizes the numbers 1–5, because children repeatedly use these numbers to count as they move their token. This emphasis on smaller numbers may have reinforced the logarithmic representation, which exaggerates differences between smaller values. This reinforcement of the logarithmic function might have interfered with encoding of the linear features and thereby increased the extent to which an individual's inhibitory control mattered for learning. Further investigations into the role of inhibitory control in different learning contexts may provide insights into how to design instruction that minimizes interference effects and helps alleviate inhibitory control deficits.

Summary and Limitations

While we controlled for adults' performance on standard number lines and kindergartners' pretest number line estimates, we did not have a direct measure of intelligence. The use of a single domain-general and domain-specific inhibitory control task in each experiment is also a limitation. While other studies have similarly used a single inhibition task to examine relations between inhibitory control and mathematics, the use of multiple measures of both domain-general and domain-specific inhibition would have reduced the possibility of measurement error and strengthened the conclusions. In choosing measures for future studies it is important to consider the extent to which cognitive control is involved. Some studies that have examined the relation between inhibitory control and mathematics have used inhibitory control tasks that required inhibition of a motor or behavioral response (e.g., pressing button on the side of the screen an animal is facing, Espy et al., 2004; Friso-van den Bos et al., 2014). On the other hand, the Stroop tasks used in the present study involved inhibition of a more cognitive component because we hypothesized cognitive interference was involved in estimation. Use of multiple measures may only be better if the control processes (i.e., behavioral vs. cognitive) are similar across the tasks. A final limitation is the correlational nature of the study. Further work is necessary to test whether there are indeed causal relations between inhibition and number line estimation, perhaps by providing training in inhibitory control and examining whether subsequent improvements in inhibitory control have any effect on number line estimation. Despite these limitations, this study is among the first to posit and investigate a potential explanation—the interference hypothesis—for the relation between inhibition and mathematics performance, and number line estimation specifically.

References

ACT. (2006). *Developing the STEM education pipeline*. Iowa City, IA: ACT.

Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition, 122,* 454–459. http://dx.doi.org/10.1016/j.cognition.2011.11.006

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173.

Barth, H. L. A. Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98,* 199–222. http://dx.doi.org/10.1016/j.cognition.2004.09.011

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology, 2,* 324–337. http://dx.doi.org/10.3389/fpsyg.2011.00324

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275,* 1293–1295. http://dx.doi.org/10.1126/science.275.5304.1293

Belsey, D. A., Kuh, E., & Welsch, R. E. (2004). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York, NY: Wiley.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38,* 333–339. http://dx.doi.org/10.1177/00222194050380040901

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. http://dx.doi.org/10.1006/jecp.1999.2518

Best, J. R., & Miller, P. H. (2010). A developmental perspective on executive function. *Child Development, 81,* 1641–1660. http://dx.doi.org/10.1111/j.1467-8624.2010.01499.x

Bialystok, E., Craik, F. I., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest, 10,* 89–129. http://dx.doi.org/10.1177/1529100610387084

Bialystok, E., & Depape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 565–574. http://dx.doi.org/10.1037/a0012735

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78,* 647–663. http://dx.doi.org/10.1111/j.1467-8624.2007.01019.x

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

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

Bull, R., Espy, K., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Cognitive Neuroscience Laboratory–Faculty and Staff Publications*. Paper 40.

Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematic achievement. *Developmental Science, 14,* 679–692. http://dx.doi.org/10.1111/j.1467-7687.2010.01012.x

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives, 8,* 36–41. http://dx.doi.org/10.1111/cdep.12059

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology, 19,* 273–293. http://dx.doi.org/10.1207/S15326942DN1903_3

Campbell, J. I., & Timm, J. C. (2000). Adults' strategy choices for simple addition: Effects of retrieval interference. *Psychonomic Bulletin & Review, 7,* 692–699. http://dx.doi.org/10.3758/BF03213008

Chesney, D. L., & Matthews, P. G. (2013). Knowledge on the line: Manipulating beliefs about the magnitudes of symbolic numbers affects the linearity of line estimation tasks. *Psychonomic Bulletin & Review, 20,* 1146–1153. http://dx.doi.org/10.3758/s13423-013-0446-8

Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Education design and cognitive science* (Vol. 5, pp. 161–238). Mahwah, NJ: Erlbaum.

Clark, C. A. C., Pritchard, V. E., & Woodward, L. J. (2010). Preschool executive functioning abilities predict early mathematics achievement. *Developmental Psychology, 46,* 1176–1191. http://dx.doi.org/10.1037/a0019672

Clark, C. A., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development, 84,* 662–677. http://dx.doi.org/10.1111/j.1467-8624.2012.01854.x

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361. http://dx.doi.org/10.1037/0033-295X.97.3.332

Dempster, F. N., & Corkill, A. J. (1999). Interference and inhibition in cognition and behavior: Unifying themes for educational psychology. *Educational Psychology Review, 11,* 1–88. http://dx.doi.org/10.1023/A:1021992632168

Diamond, A. (2012). Activities and programs that improve children's executive functions. *Current Directions in Psychological Science, 21,* 335–341. http://dx.doi.org/10.1177/0963721412453722

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction, 22,* 206–214. http://dx.doi.org/10.1016/j.learninstruc.2011.11.001

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. http://dx.doi.org/10.1016/j.jecp.2007.08.006

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in Cognitive Sciences, 12,* 374–380. http://dx.doi.org/10.1016/j.tics.2008.07.001

Egner, T., Delano, M., & Hirsch, J. (2007). Separate conflict-specific cognitive control mechanisms in the human brain. *NeuroImage, 35,* 940–948. http://dx.doi.org/10.1016/j.neuroimage.2006.11.061

Espy, K. A., Bull, R., Martin, J., & Stroup, W. (2006). Measuring the development of executive control with the shape school. *Psychological Assessment, 18,* 373–381. http://dx.doi.org/10.1037/1040-3590.18.4.373

Espy, K. A., McDiarmid, M. M., Cwik, M. F., Stalets, M. M., Hamby, A., & Senn, T. E. (2004). The contribution of executive functions to emergent mathematic skills in preschool children. *Developmental Neuropsychology, 26,* 465–486. http://dx.doi.org/10.1207/s15326942dn2601_6

Friedman, N. P., & Miyake, A. (2004). The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology, 133,* 101–135. http://dx.doi.org/10.1037/0096-3445.133.1.101

Friso-van de Bos, I., Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. (2014). Explaining variability: Numerical representations in 4-to 8-year-old children. *Journal of Cognition and Development, 15,* 325–344. http://dx.doi.org/10.1080/15248372.2012.742900

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81,* 1520–1533. http://dx.doi.org/10.1111/j.1467-8624.2010.01489.x

Ganley, C. M., & Vasilyeva, M. (2014). The role of anxiety and working memory in gender differences in mathematics. *Journal of Educational Psychology, 106,* 105–120. http://dx.doi.org/10.1037/a0034099

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

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

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. http://dx.doi.org/10.1080/87565640801982361

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76,* 104–122. http://dx.doi.org/10.1006/jecp.2000.2564

Groβe, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17,* 612–634. http://dx.doi.org/10.1016/j.learninstruc.2007.09.008

Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly, 28,* 116–159. http://dx.doi.org/10.2307/747886

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. http://dx.doi.org/10.1016/j.jecp.2008.04.001

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45,* 850–867. http://dx.doi.org/10.1037/a0014939

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition, 111,* 138–143. http://dx.doi.org/10.1016/j.cognition.2009.01.001

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25,* 95–103. http://dx.doi.org/10.1016/j.learninstruc.2012.12.001

Kray, J., & Ferdinand, N. K. (2013). How to improve cognitive control in development during childhood: Potentials and limits of cognitive interventions. *Child Development Perspectives, 7,* 121–125. http://dx.doi.org/10.1111/cdep.12027

Kroesbergen, E. H., van Luit, J. E. H., van Lieshout, E. C. D. M., van Loosbroek, E., & van de Rijt, B. A. M. (2009). Individual differences in early numeracy: The role of executive functions and subitizing. *Journal of Psychoeducational Assessment, 27,* 226–236. http://dx.doi.org/10.1177/0734282908330586

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. http://dx.doi.org/10.1111/j.1467-8624.2007.01087.x

Laski, E. V., & Siegler, R. S. (2014). Learning from number board games: You learn what you encode. *Developmental Psychology, 50,* 853–864. http://dx.doi.org/10.1037/a0034321

Laski, E. V., & Yu, Q. (2014). Number line estimation and mental addition: Examining the potential roles of language and education. *Journal of Experimental Child Psychology, 117,* 29–44. http://dx.doi.org/10.1016/j.jecp.2013.08.007

Lee, K., Bull, R., & Ho, R. M. (2013). Developmental changes in executive functioning. *Child Development, 84,* 1933–1953. http://dx.doi.org/10.1111/cdev.12096

LeFevre, J. A., Berrigan, L., Vendetti, C., Kamawar, D., Bisanz, J., Skwarchuk, S. L., & Smith-Chant, B. L. (2013). The role of executive attention in the acquisition of mathematical skills for children in Grades 2 through 4. *Journal of Experimental Child Psychology, 114,* 243–261. http://dx.doi.org/10.1016/j.jecp.2012.10.005

LeFevre, J-A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 216–230. http://dx.doi.org/10.1037/0278-7393.22.1.216

Lemaire, P. (2010). Executive functions and strategic aspects of arithmetic performance: The case of adults' and children's arithmetic. *Psychologica Belgica, 50,* 335–352. http://dx.doi.org/10.5334/pb-50-3-4-335

Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin, 136,* 1123–1135. http://dx.doi.org/10.1037/a0021276

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience, 4,* 310–322. http://dx.doi.org/10.1038/nrn1076

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: Typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development, 79,* 1524–1537. http://dx.doi.org/10.1111/j.1467-8624.2008.01203.x

Meng, X. L., Rosenthal, R., & Rubin, D. (1992). Comparing correlated correlation coefficients. *Psychological Bulletin, 111,* 172–175. http://dx.doi.org/10.1037/0033-2909.111.1.172

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49–100. http://dx.doi.org/10.1006/cogp.1999.0734

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for PreK-Grade 8 Mathematics: A quest for coherence*. Reston, VA: National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers: Washington, DC.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U. S. Department of Education.

Ni, Y., & Zhou, Y.-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist, 40,* 27–52. http://dx.doi.org/10.1207/s15326985ep4001_3

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York, NY: Plenum Press. http://dx.doi.org/10.1007/978-1-4757-0629-1_1

Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development, 79,* 788–804. http://dx.doi.org/10.1111/j.1467-8624.2008.01158.x

Pedhazur, E. J. (1997). *Multiple regression in behavioral research*. South Melbourne: Wadsworth.

Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and Psychopathology, 12,* 427–441. http://dx.doi.org/10.1017/S0954579400003096

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *The Journal of Human Resources, 27,* 313–328. http://dx.doi.org/10.2307/145737

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2011). Modeling a cascade of effects: The role of speed and executive functioning in preterm/full-term differences in academic achievement. *Developmental Science, 14,* 1161–1175. http://dx.doi.org/10.1111/j.1467-7687.2011.01068.x

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,* 359–372. http://dx.doi.org/10.1037/a0013840

Shing, Y. L., Lindenberger, U., Diamond, A., Li, S. C., & Davidson, M. C. (2010). Memory maintenance and inhibitory control differentiate from early childhood to adolescence. *Developmental Neuropsychology, 35,* 679–697. http://dx.doi.org/10.1080/87565641.2010.508546

Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2: Cognition, perception, and language* (6th ed., pp. 464–510). Hoboken, NJ: Wiley.

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

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237–250. http://dx.doi.org/10.1111/1467-9280.02438

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding of fractions. *Developmental Psychology, 49,* 1994–2004. http://dx.doi.org/10.1037/a0031200

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

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. http://dx.doi.org/10.1111/j.1751-228X.2009.01064.x

Soutschek, A., & Schubert, T. (2013). Domain-specific control mechanisms for emotional and nonemotional conflict processing. *Cognition, 126,* 234–245. http://dx.doi.org/10.1016/j.cognition.2012.10.004

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 59,* 745–759. http://dx.doi.org/10.1080/17470210500162854

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. http://dx.doi.org/10.1037/h0054651

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,* 20–51. http://dx.doi.org/10.1016/j.jecp.2008.02.003

van Galen, M. S., & Reitsma, P. (2008). Developing access to number magnitude: A study of the SNARC effect in 7- to 9-year-olds. *Journal of Experimental Child Psychology, 101,* 99–113. http://dx.doi.org/10.1016/j.jecp.2008.05.001

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11,* 381–419. http://dx.doi.org/10.1016/S0959-4752(00)00038-4

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

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology, 44,* 575–587. http://dx.doi.org/10.1037/0012-1649.44.2.575

Received April 20, 2014
Revision received January 27, 2015
Accepted January 30, 2015 ■