Siegler, R. S., & Thompson, C. A. (2014). Numerical landmarks are useful – Except when they're not. *Journal of Experimental Child Psychology, 120*, 39-58. doi: 10.1016/ j.jecp.2013.11.014

Numerical Landmarks Are Useful – Except When They're Not

Robert S. Siegler[a, b], Clarissa A. Thompson[c]

[a]Department of Psychology, Carnegie Mellon University, 331D Baker Hall, 5000 Forbes Ave, Pittsburgh PA, 15213

[b]Siegler Center for Innovative Learning, Beijing Normal University, Beijing 100875, China

[c]Department of Psychology, University of Oklahoma, 455 W. Kindsey St, 727 Dale Hall Tower, Normal, OK 73019, USA

A R T I C L E I N F O

Article history:
Received 8 June 2013
Revised 24 November 2013

Keywords:
Numerical magnitudes
Fractions
Landmarks
Verbal reports
Estimation
Encoding

A B S T R A C T

Placing landmarks on number lines, for example marking each tenth on a 0-1 line with a hatch mark and the corresponding decimal, has been recommended as a useful tool for improving children's number sense. Four experiments indicated that some landmarks do have beneficial effects, but others have harmful effects and yet others no effects on representations of common fractions (N/M). The effects of the landmarks were seen not only on the number line task where they appeared but also on a subsequent magnitude comparison task and on correlations with mathematics achievement tests. Landmarks appeared to exert their effects through the encodings and strategies that they promoted. Theoretical and educational implications are discussed.

Introduction

Mental representations of number and space are complexly intertwined. One source of evidence for this claim is the SNARC (spatial-numerical associations of response codes) Effect, the tendency to respond faster to smaller numbers when they are on the left and to larger numbers when they are on the right (Berch, Foley, Hill, & Ryan, 1999; Dehaene, Bossini, & Giraux, 1993; Hubbard, Piazza, Pinel, & Dehaene, 2005). The SNARC Effect is generally interpreted as reflecting a horizontally oriented mental number line, proceeding from smaller numbers on the left to larger ones on the right (though Santens and Gevers [2008] argue that the SNARC effect does not imply a mental number line). Converging evidence for interrelations between numerical and spatial representations comes from a variety of other behavioral paradigms, including cross-modal transfer designs (e.g., Lourenco & Longo, 2010), comparisons of schooled and unschooled populations (Dehaene, Izard, Spelke, & Pica, 2008), and examinations of people with and without brain damage (Zorzi, Priftis, & Umiltà, 2002). Neural data from both imaging and single cell recording paradigms has provided additional converging evidence for the relation between mental representations of number and space (Ansari, 2008; Hubbard et al., 2005; Nieder & Miller, 2004; Tudusciuc & Nieder, 2007).

*Corresponding author at: Department of Psychology, Carnegie Mellon University, Pittsburgh PA, 15213, USA.
E-mail address: rs7k@andrew.cmu.edu (R. S. Siegler)

Given these connections between numerical and spatial representations, it seems likely that concepts and findings from each domain can be applied to improving understanding of the other. The present study applies this general idea in the context of examining the effects of landmarks on numerical representations.

The role of landmarks in spatial representations is well established for both children and adults (see Lew, 2011, for a review). Even infants in their first year who have more than 6-weeks of crawling experience locate hidden toys more effectively when the hidden toys are near landmarks (Clearfield, 2004). Landmarks can also interfere with spatial cognition, for example by biasing searches for objects to be closer to the landmarks than they actually are (Hubbard & Ruppel, 2000). Both helpful and harmful effects are seen with subjective as well as physical landmarks; by 9-months, infants can use the presence of two physical landmarks to facilitate search for an object midway between them (Lew, Bremner, & Lefkovitch, 2000), and by 20-months, estimates of locations near the midpoint are biased toward the midpoint even when it is unmarked (Huttenlocher, Newcombe, & Sandberg, 1994).

Might landmarks play a similar role in numerical representations? The clear interconnections between spatial and numerical cognition (see de Hevia & Spelke, 2010, for a review), and the few relevant studies that have been conducted with whole numbers (e.g., Ashcraft & Moore, 2012) suggest that they might. Previous research on children's number line estimation with whole numbers provides evidence of a spatial-numerical interconnection (cf. Siegler & Opfer, 2003, Siegler & Booth, 2004), though the exact nature of the connection remains controversial (Barth & Paladino, 2011; Opfer, Young, & Siegler, 2011). Noting that adults' and older children's estimates on a 0-1,000 number line were less variable for numbers near 0, 250, 500, 750, and 1,000 than for other numbers, Siegler and Opfer (2003) hypothesized that accurate estimation on the number line task involves subjectively segmenting the line into quarters. Consistent with this account, Ashcraft and Moore (2012) found that both children's and adults' number line estimates with whole numbers are more accurate for numbers near the midpoint of the scale, and Schneider et al. (2008) reported eye-tracking evidence that children spend a substantial amount of time looking at the midpoint and endpoints of the number line as they attempt to estimate the location of whole numbers. Moreover, two studies of 5th graders' estimation of decimal fractions found that decile landmarks on number lines improve estimation accuracy (Rittle-Johnson, Siegler, & Alibali, 2001; Schneider, Grabner, & Paetsch, 2009). These findings have been among the influences leading to recommendations by mathematics educators (e.g., Cramer & Henry, 2002), textbooks (Bastable et al., 2012), and government panels (IES Fractions Practice Guide, 2010) that numerical landmarks should be used to teach students about whole numbers and fractions.

It is unclear, however, that providing landmarks is generally helpful for promoting numerical understanding. Instead, the literature on the effects of landmarks on spatial cognition suggests that the effects of landmarks depend on the fit between physical and subjective organizations, in particular, the types of encoding and strategy use that the physical landmarks elicit (Lew, 2011). The same seemed likely to be true with numerical landmarks.

The present study examined the effects of landmarks on representations of common fractions (N/M). Fractions play a crucial role in numerical development, because understanding them requires recognizing that many salient and invariant properties of whole numbers, including each number being represented by a unique symbol, having a unique successor, increasing with multiplication, and decreasing with division, are not true of numbers in general. Instead, the one invariant feature of real numbers is that they have magnitudes that can be located on a number line. Thus, fractions are central to theories of numerical development, because they require learners to discriminate between properties of whole numbers and properties shared by all real numbers (Siegler, Thompson, & Schneider, 2011).

Fractions are also of educational importance. National commissions and panels charged with improving mathematics education have singled out improved understanding of fractions as essential for improving mathematics learning. For example, the National Mathematics Advisory Panel (2008, p. 18) concluded, "The most important foundational skill not presently developed appears to be proficiency with fractions."

One major source of difficulty in many people's fraction knowledge is understanding of fraction magnitudes, (for a recent review, see Siegler, Fazio, Bailey, & Zhou, 2013). Although fractions are generally introduced in the third or fourth grade mathematics curriculum in the U.S. (NCTM, 2007), many older children and adults represent fraction magnitudes inaccurately (cf. Givvin, Stigler, & Thompson, 2011; Hecht, 1998; Hecht & Vagi, 2010; Mazzocco & Devlin, 2008; Opfer & DeVries, 2008; Schneider & Siegler, 2010; Stigler, Givvin, & Thompson, 2010). To cite one example of the problem, on the National Assessment of Educational Progress (NAEP), which is based on a large, nationally representative sample of U. S. children, 50% of eighth graders failed to correctly order from smallest to largest the fractions 2/7, 5/9, and 1/12 (Martin, Strutchens, & Elliott, 2007). On the same test, only 29% of eleventh graders correctly translated a decimal (.029) into the correct fraction (Kloosterman, 2010). Similar findings have emerged in carefully controlled experimental studies with adults; for example, in Schneider and Siegler (2010), U. S. community college students correctly answered only 70% of fraction magnitude comparison problems, where chance was 50% correct.

On tasks measuring knowledge of fraction magnitudes, both children and adults use a variety of strategies, with the particular strategies influencing the quality of performance. This influence is evident on both of the main tasks that have been used to assess fraction magnitude knowledge: magnitude comparison and number line estimation. On magnitude comparison tasks, both children (Meert, Gregoire, & Noel, 2009) and adults (Bonato, Fabbri, Umiltá, & Zorzi, 2007; Schneider & Siegler, 2010) compare numerators when all problems have equal denominators, compare denominators when all problems have equal numerators, and compare magnitudes of the whole fraction when both numerators and denominators are unequal. Avoidance of inappropriate strategies when numerators and denominators are both unequal is highly correlated with accuracy of fraction magnitude comparisons (Fazio, DeWolfe, & Siegler, under review). Children also use varied strategies on fraction number line estimation tasks and, of particular importance for the present study, frequency of at least two strategies that involve creation of landmarks is positively correlated with estimation accuracy (Siegler, Thompson, & Schneider, 2011). One such strategy involves dividing 0-1 lines into the number of units indicated by the denominator (e.g., creating sevenths landmarks to estimate the location of 3/7); another such strategy involves creating whole number landmarks on lines with an endpoint larger than one (e.g., creating landmarks at each whole number on a 0-5 number line).

In the present study, we tested four hypotheses regarding the impact of landmarks (hatch marks with numerical labels) on estimation of fractions on number lines and on numerical magnitude representations more generally. Our first hypothesis is that physical landmarks improve estimation of numerical magnitudes if they promote encoding of structurally important features that are not encoded spontaneously but that make useful strategies possible. Consistent with this finding, on a task involving estimation of decimal magnitudes on 0-1 number lines, presenting physical landmarks marking each tenth promoted encoding of the first digit to the right of the decimal point, and highlighting in red the leftmost digit in two- and three-place decimals improved estimation even more (Rittle-Johnson et al., 2001; Schneider et al., 2009). The landmarks and highlighting also promoted use of accurate strategies based on the tenths digit's value.

Our second hypothesis is that providing physical landmarks decreases numerical estimation accuracy if the landmarks reduce appropriate encoding of the numbers' magnitudes and elicit inappropriate strategies. For example, placing 10 equally spaced landmarks on a 0 to 50 number line would be harmful if it reduced encoding of the magnitudes of the numbers being estimated and instead elicited a strategy of counting landmarks. In such a case, counting nine landmarks when asked to locate "9" would lead to estimating that "9" was located where "45" actually would be.

Our third hypothesis is that physical landmarks are inconsequential if they are redundant with spontaneously formed subjective landmarks. Thus, if children spontaneously encode the midpoints of number lines when no physical landmarks are present, as suggested by Ashcraft and Moore (2012) and Schneider et al. (2008), providing a physical landmark at the midpoint is unlikely to affect number line estimates.

Our fourth hypothesis is that that correlations between mathematics achievement test scores and performance on fraction magnitude estimation tasks increase when landmarks promote encoding of magnitudes on the estimation task and decrease when they interfere with such encoding. Both correlational and causal evidence indicate relations between knowledge of fraction magnitudes and performance on standardized mathematics achievement tests (Fuchs et al., 2013; Jordan et al., 2013; Siegler & Pyke, 2013; Siegler, Thompson, & Schneider, 2011). If encoding of fraction magnitudes varies with landmark arrangements, then landmarks that increase encoding of magnitudes on experimental tasks should increase their correlation with achievement test scores, and landmarks that decrease encoding of magnitudes on the experimental tasks should reduce the correlations. The reason is that increasing children's focus on magnitudes should make performance on the experimental tasks a purer measure of magnitude knowledge, whereas decreasing that focus should make performance on the tasks a weaker measure of magnitude knowledge. This non-intuitive prediction seemed unlikely to be correct if either of the underlying hypotheses were incorrect.

Experiment 1

In Experiment 1, we tested these four hypotheses comparing the effects of decile landmarks, quartile landmarks, a midpoint landmark, and no landmarks on 10- and 11-year-olds' number line estimates with common fractions. From perspectives other than the present one, there were reasons to think that any or all of the three landmarks patterns might increase estimation accuracy. Physical decile landmarks have been shown to improve number line estimation with decimal fractions (Rittle-Johnson, et al., 2001; Schneider, et al., 2009), and subjective quartile and midpoint landmarks have been found to be associated with accurate estimation of whole number magnitudes (Ashcraft & Moore, 2012; Siegler & Opfer, 2003).

Despite these prior findings, the present analysis suggested that decile and quartile landmarks patterns would have negative effects on fractions magnitude representations and that midpoint landmarks would have no effect. Although the mapping between decile landmarks and the location of decimal fractions is straightforward (.NM should be located between the N^{th} and $N+1^{th}$ deciles), the mapping between decile landmarks and common fractions is far less straightforward. For example, decile landmarks do not indicate where 5/7 should be located in any simple way, unless 5/7 is translated to a decimal. To the contrary, decile landmarks will decrease estimation accuracy if they increase use of strategies based on dimensions other than magnitude. They might, for instance, lead children to match the numerator of the fraction to the Nth decile landmark (e.g., match ¾ to the landmark at the third decile). Quartile landmarks seem likely to decrease fraction estimation accuracy for the same reason. Finally, if people spontaneously form a subjective landmark at the midpoint, a physical landmark at that location would be expected to have no effect.

To assess whether landmarks exerted an influence on magnitude representations beyond the task where the landmarks appeared, children were presented fraction magnitude comparison problems after they completed the number line task. If the landmarks exercise their effects on number line estimation by influencing encoding of fraction magnitudes, then the earlier encountered pattern of landmarks might exert the same type of influence (positive or negative) on subsequent fraction magnitude comparisons.

Experiment 1 also tested whether the effects of landmarks on estimation accuracy would parallel their effects on correlations between estimation accuracy and mathematics achievement test performance. The logic was that mathematics achievement test scores in large part reflect numerical magnitude representations, as indicated by the strong relations between numerical magnitude representations and math achievement test scores with both whole numbers and fractions, even after general intellectual variables such as reading comprehension and other mathematical knowledge, such as whole number and fraction arithmetic have been statistically controlled (Booth & Siegler, 2006; Siegler, et al., 2011; Siegler & Pyke, 2013). If numerical magnitude knowledge strongly influences mathematics achievement test scores when no interfering conditions are present, then experimental manipulations that decrease attention to magnitudes should decrease the correlations.

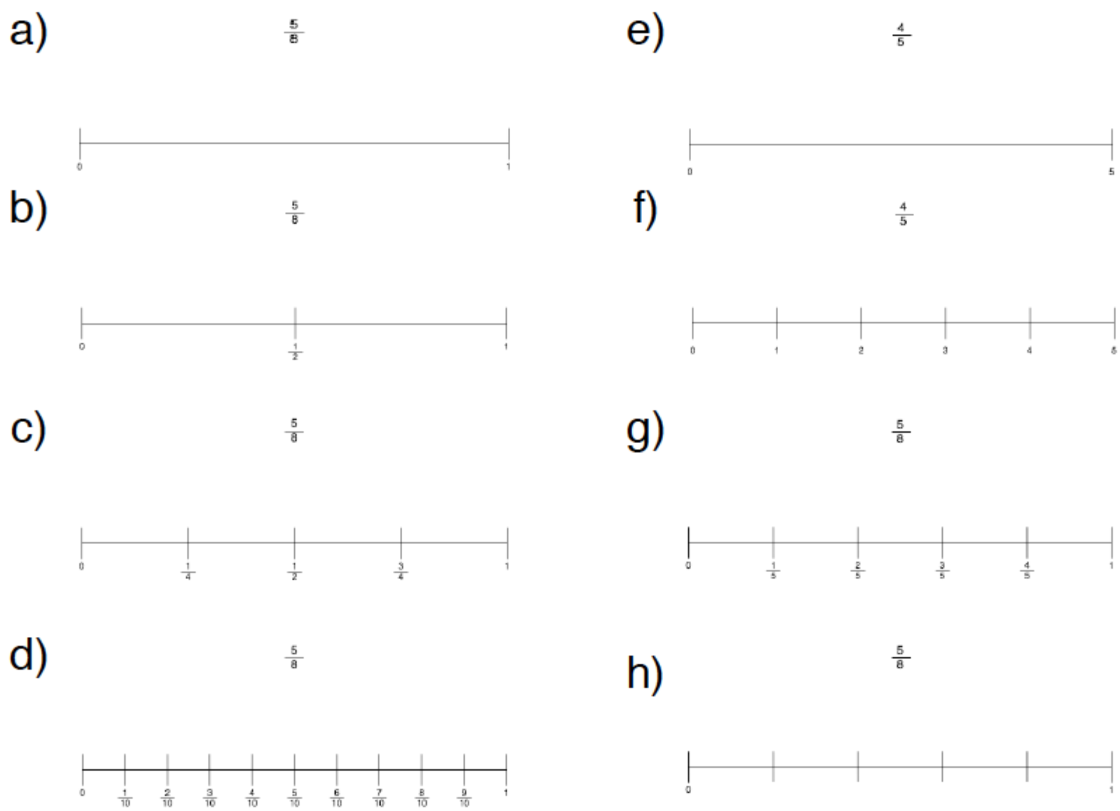

Figure 1. Experimental stimuli from the no landmarks 0-1 condition (Panel a: Experiments 1 and 3), the midpoint condition (Panel b: Experiment 1), the quartiles condition (Panel c: Experiment 1), the deciles condition (Panel d: Experiment 1), the no landmarks 0-5 condition (Panel e: Experiment 2), the whole numbers landmarks condition (Panel f: Experiment 2), the quintiles landmarks condition (Panel g: Experiment 3), and the hatch-marks-alone condition (Panel h: Experiment 4).

Method

Participants

The children were 60 fifth grade students (*Mean age* = 10.99 years, *SD* = 0.37 years; 52% females; 93% Caucasian, 3.3% Biracial, 1.7% African-American, 1.7% Asian) from three suburban public elementary schools near Pittsburgh, PA. Percent of children eligible for free lunches approximated Pennsylvania's state average (26% versus 33%). A female research assistant presented the procedure.

Tasks

Here and in all experiments, children first performed number line estimation and then magnitude comparison, so that effects of the former on the latter task could be assessed.

Number line estimation. Children were sequentially presented 20 number lines. Each line was 20 cm. long, included a left endpoint labeled 0 and a right endpoint labeled 1, and had above its midpoint a fraction whose position children needed to estimate. Thus, these were bounded rather than unbounded number lines, a variable that is important in interpreting the findings (Cohen & Blanc-Goldhammer, 2011). Fractions were drawn from each tenth of the number line (e.g., two fractions with decimal equivalents between .2 and .29). The fractions in all experiments are listed in Appendix A.

Children were randomly assigned to one of four experimental groups that differed only in the landmarks on the number lines they were presented. In the no landmarks condition (Figure 1a), no internal positions were labeled; in the midpoint condition, ½ was labeled (Fig. 1b); in the quartile landmarks condition, ¼, ½, and ¾ were labeled (Fig. 1c); and in the decile landmarks condition, all tenths were labeled (Fig. 1d). After children marked the location of a fraction on the number line, they were asked to explain why they chose that location. No feedback was provided on this or other tasks in any of the current experiments.

Magnitude comparison. Children were asked to compare the sizes of 16 pairs of fractions with magnitudes between 0 and 1 and with single digit numerators and denominators (Appendix A). Children were to press the "a" key if the fraction on the left side of the computer screen was larger and the "l" key if the fraction on the right side was larger. The side of the computer screen on which each fraction was presented was counterbalanced across children.

Achievement test. Mathematics scores on the Pennsylvania System of School Assessment (PSSA) were obtained from the children's schools. The test included sections on number and arithmetic operations, measurement, geometry, algebra, and data analysis/probability. Sample items can be found here: http://www.montroseareasd.k12.pa.us/pssa/pssa_samples/Gr5Math06.pdf

The PSSA was administered in the spring of the academic year before the number line and magnitude comparison data were obtained. Scores on the test did not differ across the four experimental conditions, $F(3, 33) = .57$, $p > .05$ (no landmarks $M = 1587$, $SD = 229$; midpoint landmark $M = 1612$, $SD = 240$; quartile landmarks $M = 1672$, $SD = 209$; decile landmarks $M = 1543$, $SD = 193$).

Procedure

Number line and magnitude comparisons tasks were presented to children individually in a quiet room in their school during a 20-minute session. Number line estimation was done with paper and pencil; magnitude comparison was done on a computer. Problems on both tasks were ordered randomly. There was no time limit on either task; the instructions on both indicated that accuracy was most important, and that speed was also somewhat important. Sessions were videotaped so that verbal reports could be coded later.

Results

Number line task

Accuracy. Accuracy of number line estimation was indexed by percent absolute error, defined as:

$$PAE = (|\text{Child's Answer} - \text{Correct Answer}|)/\text{Numerical Range}.$$

Thus, if a child was asked to locate 3/5 on a 0-1 number line and marked the location corresponding to .67, PAE would be 7%, ([|0.67-0.60|])/1. PAE varies inversely with accuracy: The more accurate the estimate, the lower the PAE.

To examine effects of landmarks on this and other tasks, performance in the no landmarks and midpoint landmarks conditions was collapsed, as was performance in the quartile and decile landmarks conditions. The reason was that these pairs of conditions were expected to produce similar performance, and no differences within either pair were found. PAEs in the midpoint and no landmarks conditions were 8% ($SD = 6\%$) and 9% ($SD = 6\%$), $t(28) = .64$, $p > .05$; those in the quartiles and deciles condition were 12% (SD = 9%) and 17% ((SD = 11%, $t(28) = 1.59$, $p > .05$.

As predicted, and contrary to the view that quartile and decile landmarks would be helpful, children estimated less accurately in the quartile/decile landmarks condition than in the no landmarks/midpoint condition, PAE = 14% (*SD* = 10%) versus 9% (*SD* = 6%), $t(58) = 2.67$, $p < .01$, $d = .61$. Estimates of children in the quartile/decile landmarks condition also were less linear, $R^2_{lin} = .67$ (*SD* = .30) versus .85 (*SD* = .21), $t(58) = 2.64$, $p = .01$, $d = .70$, and had slopes further from the ideal 1.00, *M* = .85 (*SD* = .33) versus 1.01 (*SD* = .21), $t(58) = 2.14$, $p = .04$, $d = .58$.

Encoding. To measure the main dimension of each child's encoding of the fractions, we computed gamma correlations between the child's estimate for each fraction and a) the fraction's magnitude b) its numerator, and c) its denominator. Whichever correlation was strongest was interpreted as indicating the variable that the child predominantly encoded. The one exception was that when all three correlations were less than .30, the child's main encoding was classified as unknown.

Gamma correlations are non-parametric statistics based entirely on rank order data. If the task was to estimate the locations of 3/7, 1/5, and 2/13 on a 0-1 number line, those children who estimated 3/7 as largest and 2/13 smallest would have the highest rank order correlation for fraction magnitude and therefore would be said to be encoding it; those who estimated that 3/7 was largest and 1/5 smallest would have the highest rank order correlation for numerator size and would be said to be encoding it; and those who estimated that 1/5 was largest and 2/13 smallest or 2/13 smallest and 1/5 largest would have the highest rank order for denominator size and therefore be classified as encoding it (some children thought a small denominator size indicated the larger fraction, but others thought the opposite). We computed these gamma correlations because children who predominantly encode numerator or denominator size might not generate interval scale number line estimates (Cohen & Goldhammer, 2011).

The gamma correlations indicated that experimental condition and encoding strategy were associated, χ^2 (df = 1) = 4.19, $p < .05$. Consistent with the hypothesis that landmarks that reduced estimation accuracy (the decile and quartile landmarks) did so through reducing encoding of fraction magnitudes, a smaller percentage of children in the quartile/decile landmarks condition were classified as encoding fraction magnitudes than in the no/midpoint landmarks condition, 57% versus 87%. Conversely, a higher percentage of children in the quartile/decile landmarks condition were classified as encoding numerator size, 33% versus 9%.

To validate the measure of encoding, we tested whether the standardized test scores of the 98 children in the four experiments who were classified as encoding fraction magnitude were higher than those of the 49 children who were not. The reason for combining data from the four experiments was that too few children in each experiment relied on predictors other than fraction magnitude to provide meaningful statistical comparisons.

As hypothesized, achievement test scores of children classified as encoding fraction magnitudes were higher than scores of children classified as using other strategies, mean *z* scores = .38 and -.77, $F(1, 145) = 62.05$, $p < .001$, $d = 2.9$. The pattern held true in all four experiments: Experiment 1, mean achievement test score = 1686 (*SD* = 172) versus 1436 (*SD* = 189); Experiment 2, mean achievement test score = 861 (*SD* = 81) versus 756 (*SD* = 208); Experiment 3, mean achievement test score = 847 (*SD* = 82) versus 685 (*SD* = 92); Experiment 4, mean achievement test score = 868 (SD = 77) versus 714 (SD = 112).

To summarize, children in the quartile and decile landmarks conditions estimated the magnitudes of fractions on 0-1 number lines less accurately than did children in the no landmark and midpoint landmark conditions. The less accurate estimates were accompanied by less frequent encoding of fraction magnitudes and more frequent encoding of numerators and denominators in isolation from each other. Encoding fraction magnitudes was associated not only with accurate fraction number line estimation but also with high mathematics achievement test scores.

Strategy use. Children's explanations of their number line estimates indicated use of four strategies. *Numerical transformations* involved changing the fraction to a more tractable numerical form, through rounding or translating the fraction to a decimal or percentage. For example, one student explained her estimate for 4/9 by saying, "I divided both of these by two (pointing to the numerator and denominator). I got 2/4.5, so it's a little bit less than 2/5." *Number line segmentation* involved imposing subjective landmarks on the number line; thus, a child in the no landmarks condition explained her estimate of 1/3 by saying, "Because they're thirds, and you just count off the thirds." *Magnitude* strategies involved relying on the size of the fraction relative to the numerical range; for example, a child in the quartile landmarks condition estimated 5/6 by saying, "It's almost 1 whole, so I put it by the 1." *Independent components* referred to estimates based solely on the numerator and/or solely on the denominator; for example, a child in the midpoint landmarks condition explained an estimate of 5/12 by saying, "I counted up from 0 to 5." The strategies were non-exclusive; children often used more than one strategy on a given trial.

In both conditions, frequencies of three strategies were associated with accurate number line estimation (low PAE): numerical transformations (no/midpoint landmarks, r (28) = -.64, p < .0001; quartile/decile landmarks, r (28) = -.56, p < .01), number line segmentation no/midpoint landmarks, r (28) = -.46, p < .05; quartile/decile landmarks, r (28) = -.37, p < .05), and fraction magnitudes (no/midpoint landmarks, r (28) = -.74, p < .0001; quartile/decile landmarks, r (28) = -.62, p < .0001). Use of independent components (numerator or denominators) in the no/midpoint landmarks condition was associated with inaccurate estimates (high PAE, r (28) = .70, p < .0001).

Differences in accuracy between experimental conditions appeared to derive from the conditions influencing the frequency of use of these strategies. A repeated-measures ANOVA on use of the four strategies showed a main effect of strategy F(3, 174) = 45.09, p < .0001, $\boldsymbol{\eta}^2$ = .39, a main effect of condition, F(1, 58) = 18.25, p < .001, $\boldsymbol{\eta}^2$ = .24, and a significant strategy x condition interaction, F(3, 174) = 12.60, p < .0001, $\boldsymbol{\eta}^2$ = .11. All three strategies that were associated with accurate estimation were more frequent in the no/midpoint landmarks condition than in the quartile/decile landmarks condition: numerical transformations, 53% (*SD* = 28%) versus 25% (*SD* = 24%) of trials, t (58) = 4.13, p < .001, d = 1.07; number line segmentation, 31% (*SD* = 30%) versus 12% (*SD* = 25%) of trials, t (58) = 2.65, p = 01, d = .69; and fraction magnitude, 69% (*SD* = 32%) versus 29% (*SD* = 31%) of trials, t (58) = 4.98, p < .001, d = 1.27. Thus, estimation accuracy was lower in the quartile and decile landmark conditions not only because those conditions led to less accurate encoding of fraction magnitudes but also because they led to less use of helpful estimation strategies.

Magnitude comparison task. The type of landmarks that children encountered on the number line task also influenced their subsequent magnitude comparisons. Encountering quartile or decile landmarks on number lines decreased subsequent magnitude comparison accuracy (percent correct) relative to encountering no landmarks or only a midpoint landmark, 78% (*SD* = 25%) versus 92% (*SD* = 11%) correct, t(58) = 2.74, p < .01, d = .72.

If numerical magnitude comparison and number line estimation both reflect understanding of magnitudes, individual differences on these tasks should be related. Consistent with this reasoning, magnitude comparison accuracy was correlated with all three measures of number line performance in both the no/midpoint landmarks condition and in the quartile/decile landmarks condition: for PAE, *r's* (28) = -.51 and -.67, *p's* < .01; for linearity, r (28) = .44, p < .05, and r (28) = .58, p < .01; for slope, r (28) = .41, p < .05, and r (28) = .58, p < .01.

Relations of fraction magnitude and achievement test performance

In the no/midpoint landmarks condition, mathematics achievement scores were correlated with number line PAE, r (16) = -.84, p < .01; linearity, r (16) = .80, p < .01; and slope, r (16) = .53, p < .05, as well as with percent correct on the magnitude comparison task, r (16) = .49, p < .05 (the relatively small degrees of freedom reflects some parents not agreeing to let us access their children's achievement test scores). In the quartile/decile condition, which reduced the frequency of encoding of magnitudes and strategies based on magnitudes, only one of these four correlations was significant, that between math achievement test score and number line PAE, r (17) = -.46, p < .05. These relations could not be explained by mean differences between test scores, because the mean achievement test scores of children in the two conditions were almost identical (*M* = 1600, *SD* = 227.8, and *M* = 1597, *SD* = 204.8).

Discussion

Results from Experiment 1 indicated that quartile and decile landmarks interfered with fraction number line performance and hindered subsequent fraction magnitude comparison. The interfering effect of quartile and decile landmarks also was apparent in the lower correlations with mathematics achievement test scores when such landmarks were present.

As anticipated, the deleterious effects of quartile and decile landmarks appeared to result from their eliciting encoding and strategy use based on components of fractions, in particular their numerators, rather than on the fractions' magnitudes. An alternative possibility, however, was that the results arose from the presence of multiple landmarks on a number line confusing the children, rather than from the landmarks' relation to encoding and strategy use.

Experiment 2 was designed to distinguish between these interpretations, as well as to test whether landmarks can have positive effects on fraction magnitude estimation. We presented 0-5 number lines with either no landmarks or landmarks at each whole number and asked children to locate fractions on the line. In Siegler, et al. (2011), generation of subjective landmarks at these whole number positions on 0-5 number lines was associated with accurate estimation. Therefore, we predicted that providing physical landmarks at these locations to randomly chosen children would lead to more accurate estimates and higher correlations of magnitude task performance with mathematics achievement test scores.

Experiment 2

Method

Participants

The children were 60 fifth graders (*Mean age* = 11.05 years, *SD* = 0.42 years; 45% females; 83% Caucasian, 7% Hispanic, 5% Native American, 3% Asian, and 2% Biracial), who were recruited from four public schools in Norman, Oklahoma. Eligibility for the free or reduced-price lunch program was lower in this district than the state average (25% versus 56% of students). One male and two female research assistants conducted the experiment.

Tasks and procedure

Number line. The number line task was the same as in Experiment 1 except that in both conditions, the right endpoint of the number line was labeled 5 rather than 1 (Fig. 1e), and in the whole number landmarks condition, marks and numerical labels were present at the points corresponding to 1, 2, 3, and 4 (Fig. 1f). Children were asked to estimate the positions of two fractions from each tenth of the 0-5 range (e.g. two fractions between 2 and 2 ½, see Appendix A for a list of all fractions).

Magnitude comparison. On the magnitude comparison task, children compared the reference fraction of 5/2 to 19 fractions drawn evenly from each tenth of the 0-5 range, except with only one fraction between 2 ½ and 3.

Achievement test. The mathematics part of the Oklahoma Core Curriculum Tests (OCCT) was used to measure math achievement. It included sections on number and arithmetic operations, measurement, geometry, algebra, and data analysis/probability. Sample problems from it can be found at: http://www.glencoe.com/sites/common_assets/workbooks/math/MAC3OK/m3okccw2.pdf

The OCCT was administered in the spring of the year before the number line estimation and magnitude comparison data were obtained. Scores on the test did not differ between experimental conditions: *M*'s = 809 (*SD* = 117) and 826 (*SD* = 98), $t(53) = .61$, $p > .05$. A few (5 of 60) parents denied permission to access their children's test scores, so the degrees of freedom are lower in analyses involving test scores.

Results

Number line estimation

Accuracy. As predicted, the whole number landmarks elicited more accurate estimates than no landmarks did, PAE = 15% (SD = 14%) versus 23% (SD = 11%), t(58) = 2.36, p < .05, d = .64. The whole number landmarks also elicited more linear estimates, R^2_{lin} = .59 (SD = .41) versus .39 (SD = .37), t(58) = 2.0, p < .05, d = .51, and slopes that tended to be closer to 1.00, M = .59 (SD = .45) versus .38 (SD = .44), t(58) = 1.84, p = .07, d = .47.

Encoding. Gamma correlations indicated that in the whole number landmarks condition, 63% of children encoded fraction magnitudes; in the no landmarks condition, 50% of children did so. The numerator strategy was the next most common approach; it was used by 10% of children in the whole number landmarks condition and by 20% in the no landmarks condition. There was no significant association between condition and encoding strategy, Fisher Exact Probability Test, p > .05.

Strategy use. The categorization of strategies that was used in Experiment 1 yielded similar findings to the ones from that experiment. In the whole number landmarks condition, accurate estimation (low PAE) was strongly correlated with frequency of explanations that cited numerical transformations, r (28) = -.95, p < .001; number line segmentation, r (28) = -.53, p < .01; and fraction magnitude r (28) = -.95, p < .001. In contrast, inaccurate estimation was strongly associated with reliance on the numerator or denominator, r (28) = .91, p < .01. Similarly, in the no landmarks condition, accurate estimation was associated with frequency of numerical transformations, r (28) = -.71, p < .001; number line segmentation, r (28) = -.58, p < .01; and reliance on magnitudes, r (28) = -.81, p < .001. Inaccurate performance was associated with frequency of citing numerator or denominator size alone, r (28) = .80, p < .001.

Analysis of explanations in Experiment 2 revealed a specific type of numerical transformation strategy that had a major impact on the accuracy of children who used it: the mixed number strategy. This strategy involves encoding whether a fraction is greater than one, and if it is, translating the fraction into a mixed number (a whole number and a fraction). For example, one child explained his accurate estimate for 7/3 by saying, "3 into 7 goes 2 times, 1 number leftover in 3rds."

Roughly half of children (14 of 30) in the whole number landmarks condition consistently used the mixed number strategy on fractions greater than one. These children cited this strategy to explain 98% (SD = 5%) of their estimates of fractions greater than one. In contrast, the other half of children in the whole number landmarks condition cited use of the mixed number strategy on only 25% (SD = 39%) of trials with fractions greater than one, and children in the no landmarks condition cited it on 47% (SD = 45%) of trials.

Frequency of citation of the mixed number strategy was closely related to estimation accuracy (PAE). The two variables correlated r (28) = -.94, p < .001 in the whole number landmarks condition and r (28) = -.74, p < .001, in the no landmarks condition. Children in the whole number landmarks condition who consistently explained their estimates of fractions greater than one in terms of the mixed number strategy were very accurate in absolute terms, far more accurate than peers in the no landmarks condition, PAEs = 3% (SD = 2%) versus 23% (SD = 11%), t(42) = 6.36, p < .001, d = 2.53. In contrast, the other children in the whole number landmarks condition estimated no more accurately than children in the no landmarks condition, PAE's = 25% (SD = 11%) versus 23% (SD = 11%), t(44) = .76, p > .05.

Correlations between frequency of the mixed number strategy and PAE were strong in both experimental conditions, but the relation was stronger in the whole number landmarks condition than in the no landmarks condition, r (28) = .94 versus r (28) = .60, Fisher's r-to-z transformation, p < .01 (Preacher, 2002). The difference likely reflected the greater ease of placing estimates in the correct interval when landmarks indicated which interval that was. Thus, the whole number landmarks appeared to exercise their effect by promoting consistent use of the mixed number strategy, but it had this effect on only about half of the children who were presented the landmarks.

To summarize, children who received 0-5 number lines with landmarks at each whole number were more accurate and used the highly effective strategy of translating fractions into mixed numbers more often than children who were presented the same task without the landmarks.

Magnitude comparison

Number of correct magnitude comparisons did not differ between the whole number landmarks and the no landmarks conditions, 63% (SD = 22%) versus 65% (SD = 21%) correct comparisons, $p > .05$. However, differences were present in the magnitude comparison accuracy of a) children in the whole number landmarks condition who consistently reported using the mixed number strategy, b) children in that condition who did not consistently report using that strategy, and c) children in the no landmarks condition, $F(2, 57) = 6.76$, $p < .01$, $\eta^2 = .19$. Post-hoc tests using a Bonferroni correction for family wise error rate showed that the 14 children in the whole number landmarks condition who consistently used the mixed number strategy on the number line task were more accurate on the magnitude comparison task than the 16 children who did not, 77% (SD = 16%) versus 51% (SD = 20%), $p < .01$, $d = 1.44$.

As in Experiment 1, number line and magnitude comparison performance were related. In the whole number landmarks condition, correct magnitude comparisons was related to number line PAE, $r (28) = -.48$, $p < .01$; linearity, $(28) = .54$, $p < .01$; and slope, $r (28) = .44$, $p < .05$. In the no landmarks condition, magnitude comparison accuracy also was related to number line PAE, $r (28) = -.56$, $p < .01$, linearity, $r (28) = .50$, $p < .01$; and slope, $r (28) = .63$, $p < .01$.

Relations of fraction magnitude knowledge to achievement test performance

Mathematics achievement test scores of children in the whole number landmarks condition who consistently reported using the mixed number strategy were considerably higher than those of the other children in the condition, mean OCCT scores = 887 (SD = 78) versus 761 (SD = 71), $t(25) = 4.37$, $p < .001$, $d = 1.69$. Viewed as a continuous measure, percent estimates in the correct whole number interval also was strongly related to OCCT scores, $r (25) = .69$, $p < .0001$. In contrast, in the no landmarks condition, percent estimates in the appropriate whole number interval was unrelated to overall mathematics achievement, $r (26) = .28$, ns. This difference could not be attributed to differences in average scores on the mathematics part of the OCCT, because mean scores on the test were virtually identical for children in the two experimental conditions.

In the whole number landmarks condition, which was hypothesized to promote attention to fraction magnitudes, mathematics achievement test scores were strongly correlated with all measures of magnitude knowledge: number line PAE, $r (25) = -.66$, $p < .01$; linearity; $r (25) = .67$, $p < .01$; and slope, $r (25) = .62$, $p < .01$; as well as magnitude comparison accuracy, $r (25) = .48$, $p < .05$. In contrast, in the no landmarks condition, achievement test scores were not significantly correlated with any of the measures of fraction magnitude knowledge.

Discussion

Consistent with the hypothesis that the presence of whole number landmarks would improve processing of magnitudes of improper fractions, all measures of number line estimation were superior in this condition to those in the no landmarks condition. The effect stemmed in large part from the whole number landmarks promoting use of the strategy of translating fractions into mixed numbers. The difference in estimation accuracy between children in the two conditions derived totally from the children in the whole number landmarks condition who consistently used the mixed number strategy. Consistent with the hypothesis that the whole number landmarks promoted encoding of magnitudes, correlations between the measures of magnitude knowledge and mathematics achievement scores were again stronger in the condition that promoted greater processing of magnitudes, in this case the whole number landmarks condition. Thus, landmarks can have positive effects on numerical magnitude representations, as well as the negative ones demonstrated in Experiment 1. The key seems to be whether the landmarks increase or decrease encoding of magnitudes and use of strategies that make use of those encodings. The landmarks also promoted accurate estimation by unambiguously indicating the location of the whole numbers. This facilitation was evident in the fact that use of the mixed number strategy led to far more accurate estimation when landmarks were present (PAE = 3%) than when they were not (PAE = 23%).

An alternative interpretation of the Experiment 2 findings was that it was the spatial distribution of the quintile landmarks that promoted accurate processing of fraction magnitudes, rather than the crucial variable being whether the landmarks indicated whole number locations. To test this interpretation, and to replicate and extend the Experiment 1 findings regarding deleterious effects of landmarks other than the midpoint on 0-1 number lines, Experiment 3 compared effects on magnitude processing of the presence or absence of quintile landmarks on 0-1 number lines. The main hypothesis was that the quintile landmarks would interfere with processing of 0-1 fraction magnitudes, although landmarks at the same locations had improved processing of 0-5 fraction magnitudes in Experiment 2.

Experiment 3

Method

The children were 44 fifth graders (mean age = 10.97 years, SD = 0.48 years; 48% females; 73% Caucasian, 11% Asian, 9% Hispanic, 5% African-American, and 2% Native American). The children were from the same four public elementary schools in Norman, Oklahoma as in Experiment 2, though different children participated in the two experiments. The same research assistants as in Experiment 2 conducted the experiment.

The number line task and procedure were the same as in Experiment 2 in both conditions, except that the rightmost endpoint was labeled "1" rather than "5," and in the quintile landmarks condition, the landmarks were labeled 1/5, 2/5, 3/5, and 4/5 (Fig. 1g). Children in both conditions were asked to locate the position of 20 fractions, chosen so that two were from each tenth of the number line (Appendix A). The magnitude comparison task was the same as in Experiment 2, except that the reference fraction was 4/7. The OCCT math score was obtained for each child whose parents gave permission. Average scores on the mathematics portion of the OCCT did not differ for children in the no landmarks and quintile landmarks conditions, M's = 815 (SD = 93) and 806 (SD = 124), $t(38) = .25$, $p > .05$.

Results

Number line estimation

Accuracy. As hypothesized, the quintile landmarks led to less accurate estimates on the 0-1 number line than when no landmarks were present, PAE = 15% (SD = 11%) versus 9% (SD = 7%), $t(42) = 2.17$, $p < .05$, $d = .65$. Linearity of estimates was similar for children in the no landmarks and quintile landmarks conditions, $R^2_{lin} = .81$ (SD = .27) versus .77 (SD = .20), $t(42) = .52$, $p > .05$, as were their slopes, M = .91 (SD = .29) versus .88 (SD = .35), $t(42) = .30$, $p > .05$.

Encoding. Gamma correlations indicated an association between condition and encoding, Fisher's Exact Probability Test, $p < .05$. Estimates of 86% of participants in the no landmarks condition were best fit by overall magnitude, versus 64% in the quintile landmarks condition. Conversely, the numerator was the best predictor of the estimates of 9% of children in the no landmarks condition but 36% of children in the quintile landmarks condition. Thus, the quintile landmarks on 0-1 number lines appeared to reduce encoding of fraction magnitudes and promote encoding of the numerator.

Strategies. The same strategies were used in Experiment 3 as in Experiment 1, which also examined estimates of fractions in the 0-1 range. The relation between each child's frequency of use of a given strategy and the child's PAE was also similar. In the quintile landmarks condition, accurate estimation (low PAE) was correlated with frequency of numerical transformations, subjective segmentation of the number line, and reliance on fraction magnitudes, $r(20) = -.65$, $p < .001$; $r(20) = -.64$, $p < .001$; and $r(20) = -.77$, $p < .001$; respectively. Frequency of reliance on the numerator or denominator (independent components strategy) was again related to inaccurate estimation, $r(20) = .43$, $p < .05$. In the no landmarks group, the correlations with PAE were, r (20) = -.59, p < .01, for use of numerical transformations; r (20) = -.24, p > .10, for use of number line segmentation; r (20) = -.69, p < .001, for use of fraction magnitudes; and r (20) = .50, p < .05, for reliance on the numerator or denominator.

The unhelpful landmarks again appeared to produce their effects through influencing strategy use. A repeated-measures ANOVA that analyzed percentage of numerical transformation, subjective segmentation, fraction magnitude, and independent components strategies strategy usage showed a main effect of strategy use F(3, 126) = 50.65, p < .0001, η^2 = .48, no main effect of experimental condition, F(1, 42) = .41, p > .05, and a significant strategy use x condition interaction, F(3, 126) = 13.93, p < .0001, η^2 = .13. The quintile landmarks led to less use of numerical transformations, 23% (SD = 22%) versus 38% (SD = 20%), t(42) = 2.22, p < .05, d = .71; less subjective segmentation of number lines, 17% (SD = 23%) versus 50% (SD = 30%), t(42) = 4.22, p < .001, d = 1.23; and less use of fraction magnitudes, 43% (SD = 34%) versus 75% (SD = 27%), t(42) = 3.47, p < .001, d = 1.04. The landmarks did not affect use of the independent components strategy, 10% (SD = 20%) versus 6% (SD = 10%), p > .10.

To summarize, the number line findings from this experiment closely paralleled those from Experiment 1, which also examined estimation on 0-1 number lines. Quintile landmarks led to less accurate estimates, less encoding of fraction magnitudes, and less use of beneficial strategies than no landmarks.

Magnitude comparison

Prior exposure to quintile landmarks on the number lines tended to reduce accuracy on the subsequent magnitude comparisons below that which occurred when no landmarks were present, 73% (SD = 22%) versus 83% (SD = 13%) correct, t(42) = 1.80, p = .08, d = .55. The two measures of each child's accuracy of fraction magnitude representations, number line PAE and percent correct magnitude comparisons, were correlated in both the no landmarks condition, r (20) = -.61, p < .01, and in the quintile landmarks condition, r (20) = -.65, p < .01.

Relations of magnitude knowledge to achievement test performance

In the no landmarks condition, all measures of children's fractions magnitude knowledge were related to their mathematics achievement test scores: correct magnitude comparisons, r (18) = .65, p < .01; and number line PAE, r (18) = -.67, p < .01; linearity, r (18) = .62, p < .01; and slope, r (18) = .51, p < .05. In the quintile landmarks condition, the three number line measures also were related to the mathematics achievement test scores: PAE, r (18) = -.76, p < .01; linearity; r (18) = .72, p < .01; and slope, r (18) = .56, p < .01; but magnitude comparison accuracy was not, r (18) = .32, ns. These results cannot be attributed to children in the no landmarks condition having greater mathematics ability than children in the quintile landmarks group; the two groups did not differ on a mathematics achievement test given before the experiment.

Discussion

The results of Experiment 3 indicated that it was not quintile landmarks per se that yielded superior estimation on the 0-5 number lines in Experiment 2. Rather, it was the relation between the landmarks and the kind of encoding and strategy use that the landmarks promoted. In a 0-1 fractions context, where quintile landmarks reduced encoding of numerical magnitudes and strategies based on fraction magnitudes, the quintile landmarks hindered performance rather than helping it.

Throughout this article, we have used the term "landmark" to refer to the combination of hatch marks and numerical labels that indicate the locations of numbers on the number line. It is possible, however, that hatch marks alone would have a comparable effect, because they suggest ways of segmenting number lines, and frequency of segmentation strategies correlates positively with estimation accuracy. Therefore, in Experiment 4, we randomly assigned children to either a hatch-marks-and-numerical-labels condition or to a hatch-marks-alone condition for a 0-1 number line estimation task with quintile landmarks. Our hypothesis was that the combination of numbers and hatch marks would have a greater deleterious effect than the hatch marks alone, because numerical labels indicating the number of fifths would interfere with processing of the magnitude of the fraction being estimated (none of which were fifths).

Experiment 4

Method

Fifth graders were randomly assigned to either the hatch-marks-and-numerical-labels condition (n=9, Mean age = 10.69, SD = .32, 44% males, 44% Caucasian, 22% Hispanic, 22% Native American, and 11% Asian) or to the hatch-marks-alone condition (n = 10, Mean age = 10.45, SD = .30, 40% males, 80% Caucasian, 20% Native American). The children were sampled from four public elementary schools in the same district in Norman, Oklahoma as in Experiments 2 and 3, though the current experiment was conducted two academic years later. One male and one female research assistant conducted the experiment. Verbal reports of strategy use were not collected in this experiment.

Children estimated the location of 20 fractions on 0-1 number lines. Those in the hatch-marks-and-numerical-labels condition were presented a number line divided by equally spaced vertical hatch marks that were labeled 0, 1/5, 2/5, 3/5, 4/5, and 1, as in Experiment 3. Those in the hatch-marks-alone condition were presented hatch marks at the same locations, but with no numerical labels except 0 and 1 at the endpoints. Children in both conditions were told, "Make sure you pay close attention to the different marks and the numbers on the number line when you decide where to place your mark." The procedure for the magnitude comparison task was the same as in Experiment 3. Average scores on the mathematics portion of the OCCT did not differ for children in the hatch-marks-and-numerical-labels condition (M = 789, SD = 127) versus the hatch-marks-alone condition (M = 871, SD = 69), $t(13) = 1.52$, $p > .05$.

Results

Number line estimation

Accuracy. Children estimated more accurately when presented hatch marks without numerical labels than with them (M = 9.1%, SD = 5.99%, versus M = 19.1%, SD = 11.2%, $t(17) = 2.46$, $p < .05$, $d = 1.11$). Number lines without numerical labels also elicited estimates that were more linear ($R^2 = .863$, SD = .18, versus $R^2 = .653$, SD = .22, $t(17)$ 2.29, $p < .05$, $d = 1.04$) and with a slope closer to 1.00 (M = .996, SD = .17, versus M = .750, SD = .22), $t(17) = 2.75$, $p < .05$, $d = 1.25$. PAE, linearity, and slope of children in the hatch marks alone condition were highly similar to those of children in the no landmarks conditions of Experiments 1 and 3, the two other experiments that presented 0-1 number lines, suggesting that children largely ignored the hatch marks without numerical labels. Thus, the landmarks decreased estimation accuracy only when the hatch marks were accompanied by numerical labels.

Encoding. Gamma correlations indicated a trend toward a significant association between condition and encoding, Fisher's Exact Probability Test, $p = .057$. Estimates of 44% of participants in the hatch-marks-and-numerical-labels condition were best fit by overall magnitude, versus 90% in the hatch-marks-alone condition. Conversely, the numerator was the best predictor of the estimates of 10% of children in the hatch-marks-alone condition but 56% of children in the hatch-marks-and-numerical-labels condition. Thus, the hatch marks plus the numerical labels appeared to reduce encoding of fraction magnitudes and promote encoding of the numerator in isolation.

Magnitude comparison

Magnitude comparison accuracy was not influenced by whether children earlier estimated fractions on number lines that contained hatch marks and numerical labels 81% (SD = 11%) or hatch marks alone 85% (SD = 10%) correct, $t(17) = .823$, $p > .05$. The two measures of each child's fraction magnitude representations, number line PAE and percent correct magnitude comparisons, were correlated in both the hatch-marks-and-numerical labels condition, r (7) = -.78 p < .05, and in the hatch-marks-alone condition, r (8) = -.78, p < .01.

Relations of magnitude knowledge to achievement test performance

In the hatch-marks-alone condition, children's fractions magnitude knowledge was not related to their mathematics achievement test scores: correct magnitude comparisons, r (5) = -.24; and number line PAE, r (5) = -.26; linearity, r (5) = .17; and slope, r (5) = .30, p's > .05. In the hatch-marks-and-numerical-labels condition, more accurate performance on the three number line measures and on magnitude comparison were directionally related to the mathematics achievement test scores, though the very small degrees of freedom led to several differences not being significant: number line PAE, r (6) = -.61, p>.05; linearity; r (6) = .82, p = .013; and slope, r (6) = .70, p = .053; magnitude comparison accuracy, r (6) = .46, p > .05.

General discussion

The present study demonstrated both helpful and harmful effects of landmarks with numerical labels on numerical magnitude representations. It also yielded useful information regarding the processes through which landmarks exercise their effects and about relations between numerical magnitude representations and mathematics achievement under conditions that increase or decrease attention to fraction magnitudes. In this concluding section, we discuss these issues and findings and their educational implications.

Effects of landmarks on numerical magnitude representations

Findings from the present study demonstrated that, as with spatial landmarks on spatial tasks, numerically labeled landmarks can help, harm, or leave unchanged children's numerical representations. The key is whether the landmarks promote encodings and strategies involving structurally important parts of numbers' magnitudes, whether the landmarks promote encodings and strategies that reduce attention to magnitudes, or whether the landmarks promote encodings and strategies that are redundant with those that are used spontaneously.

In Experiments 1, 3, and 4, numerically labeled decile, quartile, and quintile landmarks on 0-1 number lines reduced number line estimation accuracy, relative to encountering no landmarks, a midpoint landmark, or a hatch mark unaccompanied by a number. In two of these three experiments, the negative effects of the numerically labeled landmarks also extended to performance on subsequent numerical magnitude comparisons.

Results of Experiment 2 demonstrated that landmarks also can positively influence fraction magnitude representations. Whole number landmarks greatly improved estimation accuracy on 0-5 number lines. The landmarks exercised this positive effect by leading about half of the children who saw them to consistently translate improper fractions into mixed numbers. Children in the whole number landmarks condition who consistently used the mixed number strategy were much more accurate on the numerical magnitude comparison task than peers in the group who did not use it and than peers in the no landmarks condition. The fact that children in the whole number landmarks condition were more accurate on both the number line and magnitude comparison tasks indicated that the landmarks exerted a positive causal influence on both. The fact that children in the whole number landmarks condition who consistently used the mixed number strategy had much higher mathematics achievement test scores than those who did not indicated that children's mathematics knowledge influenced the effects of the whole number landmarks on the children's performance. Other variables, such as general intellectual ability and prior knowledge of fractions, also might be related to the benefits children derive from whole number landmarks.

Also as hypothesized, redundancy of physical and subjective landmarks led to the redundant physical landmarks having no effect. This was demonstrated by the almost identical performance of children in the midpoint landmarks condition of Experiment 1 (PAE = 8%) and peers in the no landmarks conditions of

Experiments 1 and 3 and the hatch-marks-alone condition of Experiment 4 (all three PAEs = 9%). The finding also was consistent with prior studies of purely spatial tasks in which children consistently used the midpoint as a subjective landmark (Lew, 2011); with Ashcraft and Moore's (2012) finding that on number line tasks with whole numbers, children generated a subjective midpoint landmark; and with eye-tracking data (Schneider et al., 2008) where many of children's eye movements centered on the midpoint of the number line. In all cases, children seemed to generate subjective midpoint landmarks, which would make a physical midpoint landmark redundant.

One implication of these findings is that unlike with whole numbers, coding of fraction magnitudes is not automatic. Processing of whole number magnitudes occurs automatically regardless of whether it is relevant to the problem. For example, U.S. third graders (Berch, Foley, Hill, & Ryan, 1999) and Chinese kindergartners (Zhou et al., 2007) are slower to judge that one number is physically larger than another if the physically larger number has the smaller numerical magnitude. Such automatic activation of numerical magnitudes clearly was not present with the fifth graders' fractions magnitude representations in the present experiments. Their number line estimates reflected varied encodings and strategies and fairly often were not based on fraction magnitude. Results of previous studies of number line estimation are consistent with this conclusion. For example, sixth graders in Siegler et al. (2011) required 10 s to estimate fraction magnitudes for numbers with two or three numerals, whereas sixth graders in Siegler & Opfer (2003) required less than 4 s to estimate whole number magnitudes with the same number of numerals. Whether educated adults, mathematicians, or anyone else automatically represents fraction magnitudes remains an open question.

Another implication of the present findings concerns what the number line task measures. Some have claimed that number line estimation does not measure numerical magnitude representations, arguing that the task instead assesses understanding of proportionality (Barth & Paladino, 20111; Slusser, Santiago, & Barth, 2013). However, the large differences between accuracy of number line estimates on tasks where the proportional reasoning requirements are identical indicates that this view is misguided. For example, in the present study, the proportional reasoning requirements in the no landmark conditions of different experiments were the same; one fraction needed to be placed in each tenth of equal length number lines that differed only in whether the number at the right end of the line was 1 or 5. Despite this equivalence, accuracy of estimates with different ranges of fractions differed dramatically—PAE of 9% with 0-1 number lines in Experiments 1, 3, and 4 versus PAE of 23% with 0-5 number lines in Experiment 2. This finding converges with prior findings that children's number line estimates with smaller whole numbers are much more accurate than the same children's estimates with larger whole numbers (e.g., Siegler & Opfer, 2003) and that first graders' estimates of whole numbers are much more accurate than sixth graders' estimates with fractions (Laski & Siegler, 2007; Siegler & Pyke, 2013). Like any task, multiple influences affect number line estimation, and understanding of proportions might be one of them, but there can be no question that the task does measure numerical magnitude knowledge.

A related methodological question concerns whether the number line task measures magnitude representations as opposed to magnitude estimation strategies. The fact that instructions, problem sets, and compatibility effects influence performance even on tasks that are said to be pure measures of numerical representations, such as whole number magnitude comparison, calls this distinction into question (Nuerk, Kaufmann, Zoppoth, & Willmes, 2004). The most justified conclusion seems to be that all tasks are influenced by strategies and that none yields pure measures of representations independent of strategy use.

How landmarks exercise their effects

The current findings indicate that numerically labeled landmarks exert their effects through their impact on encoding and strategy use. This was entirely consistent with Lew's (2011) conclusion after reviewing findings on the effects of landmarks on spatial processing, a conclusion that motivated the present study.

Especially striking, in Experiment 2, whole number landmarks on a 0-5 number line increased encoding of improper fractions as mixed numbers, which allowed use of the mixed numbers to guide number line estimates, a strategy that yielded highly accurate estimates. Consistent use of the mixed number strategy and the encodings that supported it appeared to reflect an insight that some children had and others did not. Roughly half of the children in the whole number landmarks condition consistently used the mixed number strategy with fractions greater than one. This was reflected in their verbal explanations citing the mixed number strategy on 98% of trials with improper fractions and also in 90% or more of their estimates being between the correct pair of whole numbers.

Our interpretation that landmarks influence numerical estimation through their effects on encoding and strategy use is consistent with superficially conflicting previous findings with decimals. In Rittle-Johnson et al. (2001) and Schneider et al. (2009), decile landmarks improved estimation accuracy for decimal fractions on 0-1 number lines. In contrast, in the present study, decile landmarks reduced estimation accuracy on 0-1 number lines. The reason was that decile landmarks increase encoding of tenths and use of strategies based on them, which is useful for estimating decimals but not for estimating the large majority of fractions.

Relations between fraction magnitude representations and mathematics achievement

Accuracy of magnitude representations was consistently related to mathematics achievement test scores. This relation was found for two different measures of magnitude representations: number line estimation and magnitude comparison. Similar relations have been found in many studies of whole number representations (e.g., Geary, 2011; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Halberda, Mazzocco, & Feigenson, 2008), and in at least two previous studies of fraction representations (Siegler & Pyke, 2013; Siegler, et al., 2011).

The current findings went beyond previous findings in demonstrating that these relations are stronger when experimental conditions promote attention to numerical magnitudes than when they do not. Landmarks that promoted greater attention to fractions' magnitudes tended to lead to stronger and more consistent relations between magnitude representations and achievement test scores. Our interpretation is that conditions that promoted magnitude encoding maximized relations of number line estimation and magnitude comparison to mathematics achievement, because then each child's magnitude knowledge was the limiting factor on their performance. In contrast, conditions that interfered with magnitude encoding weakened the relation, because under those conditions, individual differences in number line estimation and magnitude comparison would reflect ability to inhibit the distracting landmarks, as well as magnitude knowledge. This interpretation is speculative and clearly requires testing, but at a minimum the phenomenon seems worthy of further exploration.

Educational implications

Most instructional decision-making requires going beyond research on the particular decision in question and relying on general principles. The general principle that emerges from the present study is that to predict the effect of numerical landmarks, consider the encodings and strategies that the landmarks are likely to promote and the fit of those encodings and strategies to tasks of interest.

The usefulness of this principle can be illustrated through predictions regarding effects of landmarks on combinations of numerical notation and range that have not been studied. For example, for whole number ranges that start with 0 and end with a power of 10 (0-10, 0-100, 0-1,000, etc.), dividing the range into deciles that correspond to the most significant digit in that range should promote useful encodings and strategies, especially if the correspondence between the landmarks and the most significant digit in the number being estimated is highlighted. Similarly, the same landmarks that are useful with positive numbers should be useful with negatives, because the same encoding and strategies can be used (Tzelgov & Ganor-Stern, 2009). To cite a third example, for common fractions between 0 and 1, encouraging division of number lines into the number of equal size units indicated by the denominator should be useful (e.g., when the fraction 4/7 is being estimated, encourage children to divide the number line into seven equal size units and to count out four of the sevenths units). Spontaneous segmentation of this type on 0-1 number lines was correlated with accurate number line estimation in Siegler, et al. (2011), probably because, like the whole number segmentation of 0-5 number lines in the present study, it promoted division of the lines into meaningful units that could be counted to arrive at accurate estimates. Classroom discussions of why certain landmarks are useful for estimating the magnitudes of positive whole numbers, negative whole numbers, common fractions, and decimals, and why all segments in a given number line need to be equal, seem likely to help students better understand both the notational systems and the magnitudes of specific numbers within those systems.

Acknowledgements:

We thank Callie Hammond, David Martin, Audra Miller, Holly Cole, Xuan Nguyen, Christopher Parker, Sara Whisenhunt, Lana Roskin, and Sara Haas for data collection and coding as well as the administrators, parents, and teachers at Peters Township, Northgate, and Brentwood School Districts in Pennsylvania and the Norman Public School District in Oklahoma. The research described in this article was funded in part by Grants R305A080013 and R305H050035 from the Institute of Education Sciences (IES), Grant R342C100004: 84.324C from the IES Special Education Research & Development Centers, the Teresa Heinz Chair at Carnegie Mellon Univ., the Siegler Center of Innovative Learning at Beijing Normal Univ., and the Univ. of Oklahoma's Junior Faculty Summer Fellowship.

References

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience, 9*, 278-291.

Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*(2), 246-267.

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

Bastable, V., Murray, M., Bloomfield, K., Russell, S. J., Schifter, D., Cochran, K., . . . Horowitz, N. (2012). *Investigations in number, data, and space: An innovative approach to teaching and group learning mathematics.* New York:: Pearson.

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.

Clearfield, M. W. (2004). The role of crawling and walking experience in infant spatial memory. *Journal of Experimental Child Psychology, 89*(3), 214-241.

Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin & Review, 18*, 331-8.

Cramer, K., & Henry, A. (2002). Using manipulative models to build number sense for addition of fractions. *National Council of Teachers of Mathematics 2002 Yearbook: Making Sense of Fractions, Ratios, and Proportions* (pp. 41-48). Reston, VA: National Council of Teachers of Mathematics.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371-396.

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

de Hevia, M.D., & Spelke, E.S. (2010). Number-space mapping in human infants. P*sychological Science, 21*, 653-660.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105,* 683-700.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539-1552.

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*(4), 1343-1359.

Givvin, K. B., Stigler, J. W., & Thompson, B. J. (2011). What community college developmental mathematics students understand about mathematics, Part II: The interviews. *The MathAMATYC Educator*, *Volume 2,* No. 3, 4-18.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*(7213), 665-668. doi: Doi 10.1038/Nature07246

Hecht, S. A. (1998). Toward an information-processing account of individual differences in fraction skills. *Journal of Educational Psychology, 90*, 545-559.

Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology, 102*(4), 843-858.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*, 435-448.

Hubbard, T. L., & Ruppel, S. E. (2000). Spatial memory averaging, the landmark attraction effect, and representational gravity. *Psychological Research/ Psychologische Forschung, 64*, 41-55.

Huttenlocher, J., Hedges, L. V., & Vevea, J. L. (2000). Why do categories affect stimulus judgment? *Journal of Experimental Psychology: General, 129*(2), 220-241.

Huttenlocher, J., Newcombe, N., & Sandberg, E. (1994). The coding of spatial location in young children. *Cognitive Psychology, 27*, 115-147.

IES Fractions Practice Guide. (2010). Developing effective fractions instruction: A practice guide (NCEE #2010-009). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. See http://ies.ed.gov/ncee/wwc/publications/practiceguides/. (Panel: Siegler, R., Carpenter, T., Fennell, F., Geary, D., Lewis, J., Okamoto, Y., Thompson, L., & Wray, J.).

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*, 45-58.

Kloosterman, P. (2010). Mathematics skills of 17-year-old students in the United States: 1978–2004. *Journal for Research in Mathematics Education, 41*, 20–51.

Lew, A. R. (2011). Looking beyond the boundaries: Time to put landmarks back on the cognitive map? *Psychological Bulletin, 137*(3), 484-507.

Lew, A. R., Bremner, G., & Lefkovitch, L. P. (2000). The development of relational landmark use in six- to twelve-month-old infants in a spatial orientation task. *Child Development, 71*, 1179-1190.

Lourenco, S. F., & Longo, M. R. (2010). General magnitude represenation in human infants. *Psychological Science, 21*(6), 873-881.

Martin, W. G., Strutchens, M. E., & Elliott, P. C. (2007). *The learning of mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and holes: Gaps in rational number sense in children with vs. without mathematical learning disability. *Developmental Science, 11*, 681–691.

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

National Council of Teachers of Mathematics (NCTM) (2007). *Second handbook of research on mathematics teaching and learning.* Washington, DC: National Council of Teachers of Mathematics.

Nieder, A., & Miller, E. K. (2004). Analog numerical representations in Rhesus monkeys: Evidence for parallel processing. *Journal of Cognitive Neuroscience, 16*, 889-901.

Opfer, J. E., & DeVries, J. M. (2008). Representational change and magnitude estimation: Why young children can make more accurate salary comparisons than adults. Cognition, 108, 843–849.

Opfer, J. E., Siegler, R. S., Young, C. J. (2011). The powers of noise-fitting: reply to Barth and Paladino. *Developmental Science, 14*, 1194-1204.

Preacher, K. J. (2002, May). Calculation for the test of the difference between two independent correlation coefficients [Computer software]. Available from http://quantpsy.org.

Rieser, J. J., Hill, E. W., Talor, C. R., Bradfield, A., & Rosen, S. (1992). Visual experience, visual field size, and the development of nonvisual sensitivity to the spatial structure of outdoor neighborhoods explored by walking. *Journal of Experimental Psychology: General, 121*(2), 210-221.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*, 346-362.

Santens, S., & Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition, 108*, 263-270.

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.

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., . . . Stern, E. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development, 23*(3), 409-422. doi: Doi 10.1016/J.Cogdev.2008.07.002

Schneider, M., & Siegler, R. S. (2010). Representations of the magnitudes of fractions. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1227-1238.

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

Siegler, R. S., Fazio, L. K., Bailey, D. H., & Zhou, X. (2013). Fractions: The new frontier for theories of numerical development. *Trends in Cognitive Science, 17*, 13-19.

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

Siegler, R. S., & Pyke, A. A. (2013). Developmental and individual differences in understanding fractions. *Developmental Psychology, 49*, 1994-2004.

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

Stigler, J. W., Givvin, K. B., & Thompson, B. (2010). What community college developmental mathematics students understand about mathematics. *The MathAMATYC Educator*, *10*, 4-16.

Tudusciuc, O., & Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proceedings of the National Academy of Sciences of the USA, 104*, 14513–14518.

Tzelgov, J., & Ganor-Stern, D. (2009). The representation of negative numbers: Exploring the effects of mode of processing and notation. *Quarterly Journal of Experimental Psychology, 62*, 605-624.

Zhou, X. L., Chen, Y., Chen, C. S., Jiang, T., Zhang, H. C., & Dong, Q. (2007). Chinese kindergartens' automatic processing on numerical magnitude in Stroop-like tasks. *Memory & Cognition, 35*(3), 464-470.

Zorzi, M., Priftis, K., & Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature, 417*, 138.

Appendix: Number line and magnitude comparison problems presented in experiments 1-4.

NL Exp. 1	MC Exp. 1		NL Exp. 2	MC Exp. 2		NL Exp. 3	MC Exp. 3		NL Exp. 4	MC Exp. 4
1/19	5/8 vs. 4/7		1/4	5/2 vs. 7/3		1/19	6/11 vs. 4/7		1/19	1/19 vs. 4/7
1/8	4/7 vs. 2/3		2/5	9/4 vs. 5/2		1/11	5/8 vs. 4/7		1/11	1/11 vs. 4/7
1/6	3/7 vs. 5/9		7/9	5/2 vs. 11/4		1/8	4/7 vs. 9/14		1/10	4/7 vs. 1/10
3/16	4/7 vs. 3/8		4/5	15/8 vs. 5/2		1/6	4/9 vs. 4/7		2/15	2/15 vs. 4/7
3/13	3/7 vs. 5/8		11/9	19/6 vs. 5/2		2/9	4/7 vs. 5/7		2/9	2/9 vs. 4/7
2/7	2/9 vs. 3/7		14/11	5/2 vs. 9/5		4/15	4/7 vs. 5/12		3/13	4/7 vs. 3/13
6/17	4/7 vs. 7/9		9/5	10/3 vs. 5/2		1/3	4/7 vs. 7/9		3/10	4/7 vs. 3/10
3/8	2/3 vs. 3/7		15/8	5/2 vs. 7/2		6/17	6/17 vs. 4/7		6/17	6/17 vs. 4/7
5/12	3/7 vs. 1/6		9/4	5/2 vs. 11/3		5/12	4/7 vs. 1/3		5/12	4/7 vs. 5/12
4/9	5/6 vs. 4/7		7/3	5/2 vs. 14/11		4/9	7/8 vs. 4/7		7/16	4/7 vs. 7/16
6/11	1/9 vs. 3/7		5/2	5/2 vs. 11/9		6/11	4/7 vs. 4/15		5/10	5/10 vs. 4/7
4/7	4/7 vs. 2/9		11/4	4/5 vs. 5/2		4/7	4/7 vs. 8/9		4/7	4/7 vs. 5/8
5/8	3/7 vs. 7/9		19/6	7/9 vs. 5/2		5/8	4/7 vs. 11/12		5/8	4/7 vs. 9/14
9/14	1/6 vs. 4/7		10/3	5/2 vs. 17/4		9/14	2/9 vs. 4/7		9/14	7/10 vs. 4/7
5/7	5/6 vs. 3/7		7/2	13/3 vs. 5/2		5/7	13/14 vs. 4/7		7/10	4/7 vs. 3/4
11/15	1/9 vs. 4/7		11/3	5/2 vs. 9/2		7/9	4/7 vs. 1/6		3/4	4/7 vs. 11/13
5/6			17/4	5/2 vs. 2/5		7/8	1/8 vs. 4/7		11/13	7/8 vs. 4/7
11/13			13/3	14/3 vs. 5/2		8/9	1/11 vs. 4/7		7/8	9/10 vs. 4/7
8/9			9/2	1/4 vs. 5/2		11/12	4/7 vs. 1/19		9/10	18/19 vs. 4/7
17/18			14/3			13/14			18/19	

Note: Number Line Estimation Task is abbreviated NL and the Magnitude Comparison Task is abbreviated MC.