Comments Regarding *Numerical Estimation Strategies Are Correlated with Math Ability in School-Age Children*

Clarissa A. Thompson[1], Pooja G. Sidney[2], Charles J. Fitzsimmons[1], Marta Mielicki[1], Lauren Schiller[3], Daniel A. Scheibe[1], John E. Opfer[4], Robert S. Siegler[3]

[1]Kent State University

[2]University of Kentucky

[3]Teachers College, Columbia University

[4]The Ohio State University

Author's Note

Preparation of this article was supported in part by U.S. Department of Education, IES Grant # R305U200004 and NSF Grant #1941132 to Clarissa A. Thompson and by IES Grant R305A180514 to Robert Siegler.

Thompson, C. A.,Sidney, P. G., Fitzsimmons, C. J., Mielicki, M., Schiller, L. K., Scheibe, D. A., Opfer, J. E., & Siegler, R. S. (2022). Comments regarding *Numerical estimation strategies are correlated with math ability in school-age children. Cognitive Development, 62,* https://doi.org/10.1016/j.cogdev.2022.101188.

This work was supported by U.S. Department of Education Institute of Education Sciences through grant number R305U200004 to Kent State University. The opinions expressed are of the authors and do not represent views of the Institute of Education Sciences or the U.S. Department of Education.

Comments Regarding *Numerical Estimation Strategies Are Correlated with Math Ability in School-Age Children*

In the target article, Xing and colleagues (2021) claimed that 6- to 8-year-olds who spontaneously referenced the midpoint of 0-100 number lines made more accurate magnitude estimates and scored higher on a standardized math achievement test than other children. Unlike previous studies, however, the authors found no relation between accuracy on the number line estimation task and a dot discrimination task used to asssess the Approximate Number System (ANS). These findings, the authors claim, constitute evidence against the idea that children's numerical magnitude understanding entails representational change. We disagree.

In the literature on the development of numerical magnitude understanding, the "gold standard" assessment is the number-line estimation task (Schneider et al., 2018; Siegler & Opfer, 2003; Siegler et al., 2009). Unlike numerical comparisons ("Which is larger--N1 or N2?") or numerical orderings ("Can you put N1, N2, and N3 in order from smallest to largest?"), number-line estimates tell us *how much* larger the person understands the numbers to be. For example, when placing 15 on a 0-100 number line, the child's estimate tells us how large they think 15 is *in comparison* to 0 and 100.

Proponents of the representational change approach (e.g., Opfer et al., 2011; Siegler & Opfer, 2003) argue that number-line estimation reflects understanding of how numerical magnitudes relate to one another. In contrast, proponents of the proportion judgment approach (e.g., Barth & Paladino, 2011; Slusser et al., 2013) argue that number-line estimation performance reflects children's ability to place estimates on a number line relative to the 0 endpoint (i.e., unbounded model)[1], the 0 and right-most endpoint (i.e., 1-cycle model), or the endpoints and the midpoint of the number line (i.e., 2-cycle model).

Proponents of both approaches agree that estimates improve over the course of development, becoming less variable and more accurate. However, they disagree on how best to interpret patterns of data arising from number-line estimation. Are young children's estimates best fit by a mixed log-linear model (representational change approach; Opfer et al., 2016; Kim & Opfer, 2017, 2020) or one of several cyclical power functions (proportion judgment approach; Barth & Paladino, 2011; Slusser et al., 2013) when model complexity (e.g., number of free parameters) is taken into consideration?

In this rebuttal to Xing et al.'s target article, *Numerical estimation strategies are correlated with math ability in school-age children*, we argue that:

1. Number-line estimates reflect numerical magnitude representations, some task-specific features, and strategic behavior.
2. Proportional reasoning and representational change accounts must be compared head-to-head.
3. Whole numbers are ratios, too. Therefore, research on fractions can clarify mechanisms of developmental change in number-line estimation.

1. Number-line Estimates Reflect Numerical Magnitude Representations, Some Task-Specific Features, and Strategic Behavior

1.1 The Number-Line Estimation Task Aligns with the Mental Number Line.

Humans, and many other animals (Dehaene, 2011), appear to possess a mental number line. In

[1] The unbounded model is hard to reconcile with findings from Siegler and Opfer (2003). Virtually every second grader estimated two-digit numbers linearly in the 0-100 context and logarithmically in the 0-1,000 context. The only way that could happen is if the children noticed the right endpoint.

cultures with orthographies that are oriented left-to-right, the mental number line is oriented with smaller numbers on the left and larger numbers on the right (Dehaene, 2011; Opfer et al., 2010), and this left-to-right orientation aligns with the analogous orientation of the number-line estimation task. Further, size (i.e., when participants compare two symbolic or non-symbolic numerosities that are equidistant from one another, participants are quicker and more accurate when the numerosities are smaller versus larger) and distance effects (i.e., when participants compare numerosities that are more distant from one another, they are quicker and more accurate because these numerosities are easier to discriminate from one another), which have been observed in humans and other animals, follow the Weber-Fechner Law. The Weber-Fechner Law is characterized by overestimates at the low end of the numerical range and compression at the high end of the numerical range (i.e., is logarithmic in nature). Because of this, proponents of the representational change approach argue that the number-line estimation task is a good *proxy* for measuring underlying representations of magnitude.

Number-line estimation performance is typically operationalized as Percent Absolute Error (PAE) -- the absolute value of the linear deviation of a person's number line estimate from the location of the to-be-estimated number divided by the scale of the line (Siegler & Booth, 2004). For example, if a child attempts to estimate the number 15 on a 0-100 number line, but places the mark at the location for 45, PAE = |45-15|/100 * 100 = 30%. Because PAE is a measure of *error*, lower PAE means that the estimates are more accurate.

Barth and Paladino (2011, p. 134), proponents of the proportion judgment model, claimed that, "Number-line tasks can only be properly understood as proportion judgments." However, in the same paper (p. 126), they also said that the 'bias' parameter (β) in their model (described below) was an index of the “function relating psychological to actual magnitude,” and that “participants in a [number-line estimation] task must recall the proper magnitudes associated with the presented numerals.” If not from mental representations of magnitudes, from where are these magnitudes recalled?

Oscillating back to their stance that the number-line estimation task does not tap underlying numerical magnitudes, the target paper by Xing and colleagues maintains that “number-line placements are quite far removed from mental representations of numerical magnitude (p. XX),” yet acknowledge that to successfully perform on the number-line estimation task, “participants must interpret symbols representing numerical quantities and map them to a bounded linear space (p. XX),” and that “accurate performance requires not just estimating the magnitude of one number, but judging the size of the target number relative to the entire interval (p. XX)." These latter claims are consistent with our view that the number-line estimation task taps relative magnitude understanding.

1.2 Context Also Matters. Although *both* accounts model performance on the number-line estimation task as a measure of numerical magnitude understanding, Xing and colleagues claim that proponents of the representational change approach believe that children’s number line estimates are *solely* due to directly mapping between the internal representation of numerical magnitude and the external representation (i.e., placing a hatch mark to indicate the location of a number on a line). This characterization of our position is simply incorrect; we have never made this claim. Nothing of psychological interest only has a single cause. Of course, children’s number-line estimates *also* reflect task features and strategic behavior on the part of the child. We have never believed or claimed that number-line estimates are solely due to such direct mapping from the mind to paper or computer screen.

Indeed, our studies (often replicated by Barth and colleagues) have shown that performance on the number-line estimation task is associated with *many* factors including (a) transformation of magnitudes into easier-to-estimate numbers (e.g., decimals, percentages, "round" numbers; Siegler et al., 2011), (b) spontaneous, and task-supported, segmentation of the number line (Siegler & Opfer, 2003; Siegler & Thompson, 2014), (c) proportional scaling to the physical length of the line (Boyer et al., 2008; Kim & Opfer, 2020; Möhring et al., 2014; Möhring et al., 2018), (d) the number line's numerical endpoints (Barth & Palladino, 2011; Siegler & Opfer, 2003; Thompson & Opfer, 2010), (e) instruction about the number line's midpoint (Opfer et al., 2016; Zax et al. 2019), (f) position of the to-be-estimated number over the midpoint of the line (Dackerman et al., 2018), (g) practice trials (Barth et al. 2016; Opfer & Siegler, 2007; Opfer & Thompson, 2008; Thompson & Opfer, 2016), (h) length of the implied unit (Kim & Opfer, 2020), and (i) order of to-be-estimated numbers (Kim & Opfer, 2018).

Proponents of the proportion judgment approach claim that proponents of the representational change approach believe that the number-line task is a pellucid, *direct* window into participants' underlying numerical representation. But, this is not our view. From the first study using number-line estimation to examine knowledge of magnitudes (Siegler & Opfer, 2003), we have shown and argued that the experimental context (e.g., numerical range in which children are estimating) influences number-line estimation performance, just as context influences *every* cognitive process from color perception to scientific reasoning.

Although contextual factors can impact number-line estimation, the critical point is that the representational change approach is at the right level of abstraction. That is, it is both the most generalizable model (i.e., there is a logarithmic-to-linear shift across *all* variants of the whole-number estimation task) and the best-fitting model of children's estimates (i.e., the mixed log-lin model fits children's estimation data better than competing models). Moreover, accuracy on the number-line estimation task (PAE and degree of logarithmic compression in the mixed log-lin model) generalizes far beyond the task at hand to other numerical tasks purported to tap magnitude understanding. PAE is strongly related to measures of performance on other symbolic and non-symbolic numerical tasks, such as magnitude comparison (Braithwaite & Siegler, 2020; Fazio et al., 2017; Ratcliff & McKoon, 2020; Sidney, et al., 2018; Siegler & Thompson, 2014), ordering (Yu et al., 2020), arithmetic (Siegler et al., 2011), categorization (Laski & Siegler, 2007; Opfer & Thompson, 2008), memory for numbers (Thompson & Siegler, 2010), and standardized achievement test performance (Booth & Siegler, 2006; Fazio et al., 2014; Siegler & Pyke, 2013; Schneider et al. 2017; Siegler & Thompson, 2014).

Number-line estimation is also related to later algebra performance (Bailey et al., 2012; Booth & Newton, 2012; Siegler et al., 2012), which, in turn, is key to successful educational and financial outcomes (NMAP, 2008). Indeed, unlike the cyclic measures used by proponents of the proportion judgment approach, PAE has been shown to be related to measures that have real-world importance, including measures in educational contexts (i.e., understanding of course grading schemes: Scheibe et al., 2021; Thompson et al., in press), health knowledge (Lau et al., 2021; Mielicki et al., 2021; Scheibe et al., 2021; Thompson, Mielicki, et al., 2021; Thompson, Taber, et al., in press, Thompson, Taber, et al., 2021; Woodbury et al., 2021), and financial acumen (Furlong & Opfer, 2009; Kanayet et al., 2014; Taber et al., 2021). This type of predictive validity provides a meaningful way of comparing the usefulness of measures and underlying theories.

1.4 Purported Lack of Relation Between ANS Acuity and PAE. Xing et al. claim, "number-line placements are not primarily determined by the kinds of internal representations of magnitude that are thought to be probed by typical ANS tasks" (p. XX). The basis for their claim was that PAE and the measures of the ANS that they used were not significantly correlated in their study. However, the task they used was limited in that the authors only analyzed 29 of the administered trials, which included numerosities in a truncated range (9 to 14). Such trials are quite easy for children between the ages of 6 and 8.5; even preschoolers are able to reason about numbers in this numerical range (Thompson & Siegler, 2010), and first graders are quite accurate in judging whether non-symbolic arrays of asterisks are greater than or less than 50 (Thompson et al., 2016). Appropriate assessment requires the numerical range of stimuli to match children's estimation abilities (Thompson & Opfer, 2010, 2016; Wall et al., 2016).

Previous studies by researchers not associated with the representational change approach have shown that Approximate Number System acuity and PAE *are* related (Halberda et al., 2012). Our own research has similarly shown that children's performance on measures of ANS acuity, such as dot discrimination, is related to PAE in the 0-1,000 range with correlations ranging from $r = .41$ to $r = .68$ (Booth & Siegler, 2006; Fazio et al., 2014; Thompson & Siegler, 2010). In both Booth and Siegler (2006) and Thompson and Siegler (2010), the dot discrimination task also correlated with another non-symbolic estimation task, the "zips" task, in which children produced line segments of specified lengths.

Beyond data from our own labs, a recent meta-analysis (Schneider et al., 2017) showed reliable associations between non-symbolic and symbolic tasks as well as math achievement. The take-home message from Schneider and colleagues' meta-analysis, similar to the take-home message in work by Fazio and colleagues (2014), was that accuracy on the number-line estimation task, as measured by PAE, is fairly strongly related to outcomes of educational importance. No such evidence, prior to this target article, existed for any measure associated with proportion judgment models.

Replication is crucial to overcome the reproducibility crisis in psychological science. Given the prior contrary findings and the methodological limitations of their study, Xing et al. should, at minimum, replicate their findings with a larger sample of children and more appropriate matching of problem difficulty to age of participants before drawing strong conclusions from them.

2. Proportional Reasoning and Representational Change Accounts Must be Compared Head-to-Head

2.1 Percent Absolute Error as a Measure of Estimation Accuracy. Xing and colleagues claim that an individual's β parameter (i.e., bias), which is a value extracted from the cyclical power-function fit of participants' number-line estimation performance, is associated with math achievement. However, if the authors of the target article make this claim, then the burden of proof is on them to demonstrate that β is a *better* predictor of children's overall mathematics achievement than simple error (PAE) or degree of logarithmic compression (i.e., the lambda parameter in the mixed log-lin model, Opfer et al., 2016). In previous studies (Kim & Opfer, 2017), lambda was a much better predictor of addition and subtraction performance than any of the model parameters of the mixed cyclic power models. This head-to-head model comparison (Opfer et al., 2016)--even if it was not the authors' goal in their study, and we believe it should have been--is especially important given that it has been established from many studies that PAE is strongly correlated with mathematics achievement for both whole numbers and fractions (for a comprehensive review of this literature, see Schneider et al., 2018). Including a correlation table showing *how* and *whether* PAE, lambda, math achievement, and β are interrelated would be helpful for readers to establish whether the current sample of children performed similarly to the multitude of children described in other published samples.

If Xing and colleagues do not directly compare theoretical accounts, their claim that the proportion judgment framework accounts for "the exact nature of" (abstract, p. XX) children's developing number-line estimation skills is highly questionable. As Xing et al. point out, most children who are classified as using multiple reference points (i.e., best fit by the 2-cycle model) in the target article would *also* be well fit by the linear function. This begs the question of whether the authors' regression analyses predicting math achievement from the best-fitting estimation model (i.e., unbounded vs. 1-cycle vs. 2-cycle) add much beyond another demonstration that 'more accurate or linear estimates are related to math achievement,' which has already been well-documented by proponents of the representational change approach.

Moreover, since any complex model with enough free parameters can fit noise (Vanderckhove et al., 2015), absolute model fits often are not very informative (Opfer et al., 2011). A mixed log-lin model (e.g., Opfer et al., 2016) avoids the problem of averaging over different estimation patterns, thereby providing a more robust test of alternate model fits. When the log-lin model has been pitted head-to-head against cyclical power models, it has consistently been favored as the most likely data-generating model (Kim & Opfer, 2020; Opfer et al., 2016). The present study presents no data inconsistent with this finding.

3. Whole Numbers are Ratios, Too

The ages at which the unbounded versus 1-cycle versus 2-cycle models offer best fits of children's estimates seem to be task-specific. There has been no attempt by proponents of the proportion judgment account to explain the causal mechanisms underlying developmental changes. Rather than basing our understanding of the development of proportional reasoning skills on the shaky foundation of an ill-specified model, a simple model of explicit proportional reasoning (i.e., estimating fraction magnitudes on number lines) has been enormously productive.

Any general theoretical account of the development of numerical estimation understanding *must* integrate development of both whole numbers and fractions (McCrink & Wynn, 2007; Siegler et al., 2011; 2013). Whole numbers are inherently proportional (e.g., 18/3 = 6/1 = 24/4 = 6; Opfer et al., 2011; Sidney et al., 2017), and whole-number knowledge can be used to improve children's fraction knowledge (Yu et al., 2020). Therefore, research about fractions and other rational numbers should be considered when attempting to understand causal mechanisms underlying the development of estimation skills.

3.1 Research on Fraction Magnitudes Can Clarify Mechanisms of Developmental Change in Number-Line Estimates. According to the proportion judgment approach, as detailed by Slusser and colleagues (2013), performance on the number-line estimation task is driven by β and the number of spontaneous reference points participants employ during estimation. This analysis raises the question: What underlies bias and the number of reference points used during estimation? Unfortunately, the proportion judgment account does not address, much less answer, this question. Xing et al. simply say, "...from ages 6-8, the ability to spontaneously partition the bounded number line (where "spontaneous partitioning" means partitioning in the absence of any useful cues built into the task) when making proportion estimates...is a key correlate of math competence (p. XX)."

The open questions are, *how* and *why* do children and adults develop additional subjective anchor points over time? Xing et al. claim that, "the link between midpoint strategy use and math scores is clearly attributable to children's own spontaneous strategies that support more accurate proportion estimation (p. XX)," but they do not discuss the mechanism underlying this relation. The representational change account proposes that increasingly precise magnitude understanding, stemming from children's increasing experience with ever larger numerical ranges and types of numbers in addition to use of advanced strategies, explains improved accuracy and production of linear patterns of number-line estimates (Siegler & Opfer, 2003; Thompson & Opfer, 2010). Experimental studies have tested and yielded evidence supportive of this account (e.g., Opfer & Siegler, 2007; Thompson & Opfer, 2010).

Additionally, Xing and colleagues say (p. XX), "Once an age is reached at which most or all participants in the sample have the ability to use such a strategy, of course, the presence of the strategy will not indicate a difference in math competence." This was precisely the conclusion drawn by Siegler and Thompson (2014) when they found that fifth graders did not benefit from a labeled landmark at the midpoint of a 0-1 fraction number line. Siegler and Thompson's explanation was that these children already spontaneously segmented the line at its midpoint to aid in making their estimates, and therefore the midpoint landmark was redundant. Furthermore, those children who were randomly assigned to the 0-5 whole-number landmarks group who knew how to transform improper fractions into mixed numbers had higher standardized math achievement scores than those who did not use the strategy. This was not the case for those children who made estimates in the 0-5 no landmarks condition. Taken together, these results indicated that spontaneous strategy use, in conjunction with researcher-generated landmarks, was associated with higher math achievement. If Xing and colleagues consider findings from research involving children's and adults' use of landmarks to make more precise fraction estimates, their own findings regarding the effects of using midpoints on estimation accuracy would not seem as novel. Likewise, if they only consider findings from research involving whole numbers relevant, research dating back nearly two decades (Siegler & Opfer, 2003) indicated that participants were more accurate on estimates that were close to subjective landmarks.

Xing et al. take it as evidence that the 2-cycle model fits children's data better than the representational change account because children spontaneously use landmarks. But, given the substantial overlap in the linear model and the 2-cycle model, the fit of the cyclical power function seems to be a statistical artifact. In fact, the authors did not disclose the percentage of children who were best fit by the linear function. Proponents of the representational change approach have generated a plethora of evidence (see Supplemental Table) consistent with the idea that children spontaneously segment the number line when estimating whole-number and fraction magnitudes. The two approaches are in agreement that such subjective landmarks influence children's whole-number estimation performance.

4. Conclusions

Numbers are inherently relational (Sidney & Thompson, 2019); there are intricate links connecting space and number (Geary et al., 2021; Opfer et al., 2010; McCrink & Opfer, 2014; McCrink & Wynn, 2009; Thompson et al., 2017); use of segmentation strategies is linked to number-line estimation precision (Ashcraft & Moore, 2012; Fitzsimmons et al., 2021; Fitzsimmons et al., 2020; Opfer & Siegler, 2007; Schneider et al., 2009; Siegler & Opfer, 2003; Sidney et al., 2018); and estimation performance is correlated with overall math achievement (Fazio et al., 2014; Sidney et al., 2018; Siegler & Pyke, 2013; Siegler & Thompson, 2014; Siegler et al., 2011), health knowledge (Lau et al., 2021; Mielicki et al., 2021; Scheibe et al., 2021; Thompson, Mielicki, et al., 2021; Thompson et al., in press; Thompson, Taber, et al., 2021; Woodbury et al., 2021), and financial acumen (Furlong & Opfer, 2009; Kanayet et al., 2014; Taber et al., 2021). To fairly assess whether the proportion judgment model provides added value to the literature beyond what has already been published by proponents of the representational change approach, the two approaches must be compared head-to-head in future research, and models must account for the large body of data generated by proponents of both approaches.

Supplemental Table 1: Non-exhaustive list of relevant studies involving the relation between number-line estimation PAE and its association with math achievement, dot discrimination, and segmentation strategies

Citation	**Whole Numbers (WN), Fractions (FR), or Decimal (D) Stimuli and Numerical Ranges**	**Ages of Participants**	**Relevant Findings**
Rittle-Johnson et al., 2001	D (0-1)	11 yrs (5th grade)	Researcher-generated landmarks and reminders helped children find the location of decimals on number lines
Siegler & Opfer, 2003	WN (0-100, 0-1,000)	7-11 yrs (2nd, 4th, 6th grades), college-age adults	Participants were the most accurate around subjective quartile landmarks
Booth & Siegler, 2006	WN (0-100, 0-1,000)	5-9 yrs (K-4th grade)	PAE correlated with math achievement; number line estimation was correlated with dot discrimination and production of line segments
Opfer & Siegler, 2007	WN (0-1,000)	8-10 yrs (2nd-4th grade)	Feedback on important landmarks (e.g., 5, 150, 725) improved estimation precision

Booth & Siegler, 2008	WN (0-100)	7 yrs (1st grade)	PAE correlated with novel arithmetic problem performance
Opfer & Thompson, 2008	WN (0-1,000)	7 yrs (1st - 2nd grade)	Feedback on an important landmark (i.e., 150) improved estimation precision and transferred to magnitude categorization; children were most accurate around the point of feedback
Schneider et al., 2008	WN (0-100)	6-8 yrs (1st - 3rd grade)	Eye-tracking data indicated that young children's attention was spontaneously directed to the midpoint and endpoints of number lines
Thompson & Opfer, 2008	WN (0-1,000) and FR (1/1 -1/1,440 min)	6-8 yrs (1st - 3rd grade)	Feedback on an important landmark (i.e., 150) improved estimation precision and transferred to fraction number line estimation precision
Thompson & Opfer, 2010	WN (0–1,000, 0–10,000, 0–100,000_	7-12 yrs (2nd - 6th grade), college-age adults	Progressive alignment of increasingly larger numerical ranges allows children to scale up their number line estimates
Thompson & Siegler, 2010	WN (0-20, 0-1,000)	5 - 8 yrs (pre-K - 2nd)	Number line estimation correlated with magnitude comparison, memory for numbers, dot discrimination, and production of line segments

Siegler, Thompson, & Schneider, 2011	FR (0-1, 0-5)	11-13 yrs (6th and 8th grades)	Estimation precision was related to spontaneous segmentation of number lines, magnitude comparison, and arithmetic performance
Ashcraft & Moore, 2012	WN (0-100, 0-1000)	6-10 yrs (1st - 5th grade), college-age adults	Spontaneous use of midpoint strategy
Siegler & Pyke, 2013	FR (0-1, 0-5)	12-14 yrs (6th and 8th grades)	Estimation precision was related to standardized math achievement
Fazio et al., 2014	WN (0-1,000) and FR (0-1)	10 yrs (5th grade)	Symbolic and non-symbolic number line estimation and magnitude comparison was correlated with overall math achievement
Siegler & Thompson, 2014	FR (0-1)	10 yrs (5th grade)	Estimation precision was related to magnitude comparison and standardized math achievement; researcher-generated landmarks as well as spontaneous segmentation was related to more precise estimates
Fazio et al., 2016	FR (0-1)	10 yrs (4th - 5th grade)	Researcher-generated landmarks to help children find the location of fractions on number lines
Opfer, Thompson, &	WN (0-1,000)	7 yrs (1st - 2nd	Feedback on important landmarks (e.g., 500)

Kim, 2016		grade)	resulted in cyclical power function fit of estimation data
Thompson & Opfer, 2016	WN (0-1,000)	5-8 yrs (K-2nd grade), college-age adults	Feedback on an important landmark (i.e., 150) improved estimation precision and transferred to memory for numbers
Fazio et al., 2017	FR (0-5)	College-age adults	Spontaneous segmentation strategies related to precise number line estimates and magnitude comparison
Thompson, Morris, & Sidney, 2017	WN (0-100)	7 yrs (1st grade)	Estimation precision was related to a non-symbolic estimation task (i.e., finding a page in a book without page numbers)
Sidney, Thalluri, et al., 2018	FR (0-5)	College-age adults	Spontaneous segmentation strategies related to precise number line estimates
Sidney, Thompson, & Rivera, 2019	FR (no endpoints labeled)	11 yrs (5th - 6th grade)	Researcher-generated landmarks on number lines helped children solve fraction division problems relative to other visual models
Braithwaite & Siegler, 2020	FR (0-1)	9-12 yrs (4th - 6th grade)	Segmenting number lines into unit fractions was related to magnitude comparison accuracy
Fitzsimmons, Thompson, et al., 2020	FR (0-1_	College-age adults	Spontaneous segmentation strategies related to precise number line estimates

Fitzimmons, Morehead, et al., 2021	WN (0-1,000, 1,000-1,000,000,000)	7 yrs (1st - 2nd grade), college-age adults	Spontaneous segmentation strategies related to precise number line estimates

References

Ashcraft, M. H., & Moore, A. M (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology, 111*(2), 246-267. https://doi.org/10.1016/j.jecp.2011.08.005

Bailey, D. H., Hoard, M. K., Nugent, L., & Geary, D. C. (2012). Competence with fractions predicts gains in mathematics achievement. *Journal of Experimental Child Psychology, 113,* 447-455. https://doi.org/10.1016/j.jecp.2012.06.004

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Developmental Science, 14*(1), 125-135. https://doi.org/10.1111/j.1467-7687.2010.00962.x

Barth, H., Slusser, E., Cohen, D., & Paladino, A. (2011). A sense of proportion: Commentary on Opfer, Siegler and Young. *Developmental Science, 14*(5), 1205-1206. https://doi.org/10.1111/j.1467-7687.2011.01081.x

Barth, H., Slusser, E., Kanjilia, S., Garcia, J. Taggert, J., & Chase, E. (2016). How feedback improves children's numerical estimation. *Psychonomic Bulletin & Review, 23*(4). https://doi.org/10.3758/s13423-015-0984-3

Booth, J. L., & Newton, K. J. (2012). Fractions: Could they really be the gatekeeper's doorman? *Contemporary Educational Psychology, 37*, 247-253. https://doi.org/10.1037/xlm0000546

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*(1), 189-201. https://doi.org/10.1037/0012-1649.41.6.189

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology, 44*(5), 1478-1490. https://doi.org/10.1037/a0013110

Braithwaite, D. W., & Siegler, R. S. (2020). Putting fractions together. *Journal of Educational Psychology*. 1-16. https://doi.org/10.1037/edu0000477

Dehaene, S. (2011). *The Number Sense* (2nd ed.). New York, NY: Oxford University Press.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of number in the animal and human brain. *Trends in Neurosciences, 21*(8), 355-361. https://doi.org/10.1016/S0166-2236(98)01263-6

Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of differenttypes of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123,* 53-72. https://doi.org/10.1016/j.jecp.2014.01.013

Fazio, L. K., DeWolf, M., & Siegler, R. S. (2017). Strategy use and strategy choice in fraction magnitude comparison. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 42,* 1-16. https://doi.org/10.1037/xlm0000153

Feigenson, L., Dehaene S., Spelke E. (2004). Core systems of number. *Trends in Cognitive Science, 8*(7), 307-14. doi: 10.1016/j.tics.2004.05.002.

Fitzsimmons, C. J., Morehead, K., Thompson, C. A., Buerke, M., & Dunlosky, J. (2021). Can feedback, correct, and incorrect worked examples improve numerical magnitude estimation precision? *Journal of Experimental Education*. Advance online publication. https://doi.org/10.1080/00220973.2021.1891009

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2020). Do adults treat equivalent fractions equally? Adults' strategies and errors during fraction reasoning. *The Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(11), 2049-2074. https://doi.org/10.1037/xlm0000839.

Furlong, E. E., & Opfer, J. E. (2009). Cognitive constraints on how economic rewards affect cooperation. *Psychological Science, 20*(1), 11-16. https://doi.org/10.1111/j.1467-9280.2008.02244.x

Geary, D. C., Hoard, M. K., Nugent, L., & Scofield, J. E. (2021). In-class attention, spatial ability, and mathematics anxiety predict across-grade gains in adolescents' mathematics achievement. *Journal of Educational Psychology, 113*(4), 754-769. https://doi.org/10.1037/edu0000487

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., Germine, L. (2012). Number sense across lifespan. *Proceedings of the National Academy of Sciences, 109*(28) 11116-11120. DOI: 10.1073/pnas.1200196109

Kanayet, F. J., Opfer, J. E., & Cunningham, W. A. (2014). The value of numbers in economic rewards. *Psychological Science, 25*(8), 1534-1545. https://doi.org/10.1177/0956797614533969

Kim, D., & Opfer, J. E. (2017). A unified framework for bounded and unbounded numerical estimation. *Developmental Psychology, 53*(6), 1088-1097.

https://doi.org/10.1037/dev0000305.

Kim, D., & Opfer, J. E. (2020). Compression is evident in children's unbounded and bounded numerical estimation: Reply to Cohen and Ray (2020). *Developmental Psychology, 56*(4), 853-860. https://doi.org/10.1037/dev0000886

Lau, N.T.T., Wilkey, E.D., Soltanlou, M., Lagacé Cusiac, R., Peters, L., Tremblay, P., Goffin, C., Starling-Alves, I., Ribner, A.D., Thompson, C., Van Hoof, J., Bahnmueller, J., Alvarez, A., Bellon, E., Coolen, I., Ollivier, F. & Ansari D. (In Principle Acceptance, Stage 1 Registered Report). Numeracy and COVID-19: Examining interrelationships between numeracy, health numeracy, and behaviour. *Royal Society Open Science.*

Link, T., Nuerk, H-C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *The Quarterly Journal of Experimental Psychology, 67*(8), 1597-1613. https://doi.org/10.1080/17470218.2014.892517

McCrink, K., & Opfer, J. E. (2014). Development of spatial-numerical associations. *Current Directions in Psychological Science, 23*(6), 439-445. https://doi.org/10.1177/0963721414549751

McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology, 103*, 400-408. https://doi.org/10.1016/j.jecp.2009.01.013

Mielicki, M., Fitzsimmons, C. J., Schiller, L., Scheibe, D., Taber, J., Sidney, P. G., Matthews, P., Waters, E., Coifman, K., & Thompson, C. A. (2021). Adults' health-related problem solving is facilitated by number lines, but not risk ladders and icon arrays [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

Möohring, W., Frick, A., & Newcombe, N. S. (2018). Spatial scaling, proportional thinking, and numerical understanding in 5- to 7-year-old children. *Cognitive Development, 45,* 57-67. https://doi.org/10.1016/j.cogdev.2017.12.001

Möhring, W., Newcombe, N. S., & Frick, A. (2014). Zooming in on spatial scaling: Preschool children and adults use mental transformations to scale spaces. *Developmental Psychology, 50*(5), 1614-1619. https://doi.org/10.1037/a0035905

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for Research in Mathematics Education, 30*(20), 122-147. https://doi.org/10.2307/749607

NMAP (2008). *Foundations for success: Report of the National Mathematics Advisory Board Panel:* U.S. Government Printing Office.

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology, 55*, 169-195. https://doi.org/10.1016/j.cogpsych.2006.09.002

Opfer, J. E., & Thompson, C. A. (2008). The trouble with transfer: Insights from microgenetic changes in the representation of numerical magnitude. *Child Development, 79*(3), 788-804. https://doi.org/10.1111/j.1467-8624.2008.01158.x

Opfer, J. E., Thompson, C. A., & Furlong, E. (2010). Early development of spatial-numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science, 13,* 761-771. https://doi.org/10.1111/j.1467-7687.2009.00934.x

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: reply to Barth and

Paladino. *Developmental Science, 14*(5), 1194-1204. https://doi.org/10.1111/j.1467-7687.2011.01070.x

Opfer, J. E., Thompson, C. A., & Kim, D. (2016). Free versus anchored numerical estimation: A unified approach. *Cognition, 149*, 11-17. https://doi.org/10.1016/j.cognition.2015.11.015

Ratcliff, R., & McKoon, G. (2020). Decision making in numeracy tasks with spatially continuous scales. *Cognitive Psychology, 116*, 1-21. https://doi.org/10.1016/j.cogpsych.2019.101259

Rittle-Johnson, B. (2017). Developing mathematics knowledge. *Child Development Perspectives, 11*(3), 184-190. https://doi.org/10.1111/cdep.12229

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. https://doi.org/10.1037/0022-0663.93.2.346

Scheibe, D. A., Fitzsimmons, C. J., & Thompson, C. A. (2021). Number lines can facilitate students' understanding of college course grading schemes. (Manuscript in preparation). Department of Psychological Sciences, Kent State University.

Scheibe, D. A., Fitzsimmons, C. J., Mielicki, M., Taber, J., Sidney, P., Coifman, K., & Thompson, C. A. (2021). When it is healthy to be confident: What factors predict adults' metacognitive judgments on COVID-related math problems before and after their engagement with an educational intervention? [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

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. https://doi.org/10.1037/a0013840

Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., Jacobs, A. M., & Stern, E. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cognitive Development, 23*(3), 424-437. https://doi.org/10.1016/j.cogdev.2008.07.002

Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science, 20*(3). https://doi.org/10.1111/desc.12372

Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development, 89*(5), 1467-1484. https://doi.org/10.1111/cdev.13068

Sidney, P. G., & Thompson, C. A. (2019). Implicit analogies in learning: Supporting transfer by warming up. *Current Directions in Psychological Science, 28*(6), 619-625. https://doi.org/10.1177/0963721419870801

Sidney, P. G., Thalluri, R., Buerke, M. L., & Thompson, C. A. (2018). Who uses more strategies? Linking mathematics anxiety to adults' strategy variability and performance on fraction magnitude tasks. *Thinking and Reasoning, 25*(1), 94-131. https://doi.org/10.1080/13546783.2018.1475303

Sidney, P. G., Thompson, C. A., Matthews, P. G., & Hubbard, E. M. (2017). From continuous magnitudes to symbolic numbers: The centrality of ratio. *Behavioral and Brain Sciences, 40*(190). https://doi.org/10.1017/S0140525X16002284

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*(2), 428-444. https://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*(3), 237-243. https://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*(10), 1994-2004. https://doi.org/10.1037/a0031200

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127*(4), 377-397. https://doi.org/10.1037/0096-3445.127.4.377

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

Siegler, R. S., Thompson, C. A., & Opfer, J. E. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many time scales. *Mind, Brain, and Education, 3*(3), 143-150. https://doi.org/10.1111/j.1751-228X.2009.01064.x

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

https://doi.org/10.1016/j.cogpsych.2011.03.001

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*(7), 691-697. https://doi.org/10.1177/0956797612440101

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*(1), 13-19. https://doi.org/10.1016/j.tics.2012.11.004

Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology: General, 142*(1), 193-208. https://doi.org/10.1037/a0028560

Taber, J. M., Updegraff, J., Sidney, P. G., O'Brien, A., & Thompson, C. A. (2021). *Experimental tests of how hypothetical monetary lottery incentives influence vaccine-hesitant U.S. adults' intentions to vaccinate* [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

Thompson, C. A., & Opfer, J. E. (2008). Costs and benefits of representational change: Effects of context on age and sex differences in symbolic magnitude estimation. *Journal of Experimental Child Psychology, 101*(1), 20-51. https://doi.org/10.1016/j.jecp.2008.02.003

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child*

Development, 81(6), 1768-1786. https://doi.org/10.1111/j.1467-8624.2010.01509.x

Thompson, C. A., & Opfer, J. E. (2016). Learning linear spatial-numeric associations improves accuracy of memory for numbers. *Frontiers in Psychology, 7*(24), 1-9. https://doi.org/10.3389/fpsyg.2016.00024

Thompson, C. A., & Siegler, R. S. (2010). Linear numerical-magnitude representations aid children's memory for numbers. *Psychological Science, 21*(9), 1274-1281. https://doi.org/10.1177/0956797610378309

Thompson, C.A., Fitzsimmons, C. J., & Scheibe, D. A. (in press). Students' ability to calculate their final course grade may not be as easy as you think: Insights from mathematical cognition. *Scholarship of Teaching and Learning in Psychology*.

Thompson, C. A., Mielicki, M., Rivera, F., Fitzsimmons, C. J., Scheibe, D. A., Sidney, P. G., Taber, J. M., & Waters, E. A. (2021). *Leveraging math cognition to combat health innumeracy* [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

Thompson, C. A., Morris, B. J., & Sidney, P. G. (2017). Are books like number lines? Children spontaneously encode spatial-numeric relationships in a novel spatial estimation task. *Frontiers in Psychology, 8*, 1-10. https://doi.org/10.3389/fpsyg.2017.02242

Thompson, C. A., Ratcliff, R., & McKoon, G. (2016). Individual differences in the components of children's and adults' information processing for simple symbolic and non-symbolic numeric decisions. *Journal of Experimental Child Psychology, 150*, 48-71. https://doi.org/10.1016/j.jecp.2016.04.005

Thompson, C. A., Taber, J. M., Fitzsimmons, C., & Sidney, P. (in press). Strategy reports involving attention to math are associated with accurate responses on a numeric health decision-making problem. Journal of Numerical Cognition.

Thompson, C. A., Taber, J. Sidney, P. G., Fitzsimmons, C. J., Mielicki, M., Matthews, P., Schemmel, E., Simonovic, N., Foust, J., Aurora, P., Stanley Seah, T. H., Disabato, D., Coifman, K. (2021). *Math matters during a pandemic: A novel, brief educational intervention combats whole number bias to improve health decision-making and predicts COVID-19 risk perceptions and worry across 10 days* [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

Vanderkerckhove, J., Matzke, D., & Wagenmakers, E.-J. (2015). Model comparison and the principle of parsimony. *The Oxford Handbook of Computational and Mathematical Psychology.*

Wall, J. L., Tompson, C. A., Dunlosky, J., & Merriman, W. E. (2016). Children can accurately monitor and control their number-line estimation performance. Developmental Psychology, 52(10), 1493-1502. https://doi.org/10.1037/dev0000180

Woodbury, L., Fitzsimmons, C. J., Taber, J. M., Schiller, L., & Thompson, C. A. (2021). *Impact of visual displays on health-risk estimates: Icon arrays can lead to less precise estimates of hypothetical side effects* [Manuscript submitted for publication]. Department of Psychological Sciences, Kent State University.

Yu, S., Kim, D., Mielicki, M., Fitzsimmons, C., Thompson, C. A., & Opfer, J. E. (2020). From integers to fractions: Developing a coherent understanding of proportional magnitude. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society.*

Zax, A., Slusser, E., Barth, H. (2019). Spontaneous partitioning and proportion estimation in children's numerical judgments. *Journal of Experimental Child Psychology, 185*, 71-94. 10.1016/j.jecp.2019.04.004