Cognition 230 (2023) 105257

Contents lists available at ScienceDirect

Cognition

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

Modeling the left digit effect in adult number line estimation

Andrea L. Patalano [*], Kelsey Kayton [a], Hilary Barth

Department of Psychology, Wesleyan University, United States

ARTICLE INFO

Keywords:
Left digit effect
Number line estimation
Numerical cognition
Mathematical modeling

ABSTRACT

Number line estimation tasks are frequently used to study numerical cognition skills. In a typical version, the bounded number line task, target numerals must be placed on a bounded line labeled only at its endpoints (e.g., with 0 and 100). Placements by adults, while highly accurate, reveal a cyclical pattern of over- and underestimation of target numerals. The pattern suggests use of proportion judgment strategies and is well-captured by cyclical power models. Another systematic number line bias that has recently been observed, but has not yet been considered in modeling efforts, is the *left digit effect*. Numerals with different leftmost digits (e.g., 39 and 41) are placed farther apart on a line than is warranted. In the current study ($N = 60$), adult estimates were obtained for all numerals on a 0–100 number line estimation task, and fit of the standard cyclical power model was compared with two modified versions of the model. One modified version included a parameter that underweights the rightward digit's place value (e.g., the ones digit here), and the other used the same parameter to underweight *all* digits' place values. We found that both modifications provided a considerably better fit for individual and median data than the standard model, and we discuss their relative merits and cognitive interpretations. The data and models suggest how a left digit bias might impact estimates across the number line.

1. Introduction

Understanding numerical quantities is an important cognitive skill. A tool frequently used to study, assess, and train numerical magnitude estimation skill is the *number line estimation task* (e.g., Barth & Paladino, 2011; Booth & Siegler, 2008; Brez, Miller, & Ramirez, 2016; Hamdan & Gunderson, 2017; Schneider et al., 2018; Siegler & Opfer, 2003; Siegler & Ramani, 2009; Slusser, Santiago, & Barth, 2013; Xing et al., 2021; Zhu, Cai, & Leung, 2017). A typical version, called the bounded number line task, involves estimating the locations of target numerals (e.g., '82') on a horizontal line labeled only at its endpoints (e.g., with 0 and 100; see Fig. 1).[1] The task is one of proportion judgment in that it involves judging the location of a numeral relative to bounding reference points (e.g., 25 is a quarter of the way between 0 and 100; e.g., Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011; but see Siegler & Opfer, 2003). A central measure of task performance is *overall accuracy error*, which captures the mean absolute difference between one's placements and correct locations (and is often expressed as a percentage of the total range), and is used to assess individual differences in numerical estimation skills. Overall accuracy error has been linked to many measures of numerical competence, including math achievement test scores in children (e.g., Booth & Siegler, 2008; Holloway & Ansari, 2009; Schneider, Grabner, & Paetsch, 2009; Tosto et al., 2017; see Schneider et al., 2018, for review) and decision-related number skills in adults (Patalano et al., 2020; Peters & Bjalkebring, 2015; Schley & Peters, 2014).

Already well documented is a cyclical pattern of placement error in number line estimation, illustrated in its simplest form (a one-cycle curve) in Fig. 3a. This pattern, which is found in a wide range of tasks involving proportion judgment (see Hollands & Dyre, 2000; Zhang & Maloney, 2012), has been successfully modeled using Hollands and Dyre's *cyclical power model*. The model builds on Stevens' Law (Stevens, 1957) which describes the relationship between the estimated magnitude of a physical stimulus (e.g., brightness) and its actual magnitude as

* Corresponding author at: Department of Psychology, Wesleyan University, 207 High Street, Middletown, CT 06459, United States.
E-mail address: apatalano@wesleyan.edu (A.L. Patalano).
[a] Present Address: Department of Psychology, Ohio University, Athens, OH, United States.

[1] We focus here on bounded number line estimation tasks but there are also unbounded versions that elicit different error patterns (Cohen et al., 2018; Cohen & Blanc-Goldhammer, 2011).

https://doi.org/10.1016/j.cognition.2022.105257
Received 24 March 2022; Received in revised form 5 July 2022; Accepted 12 August 2022
Available online 10 October 2022

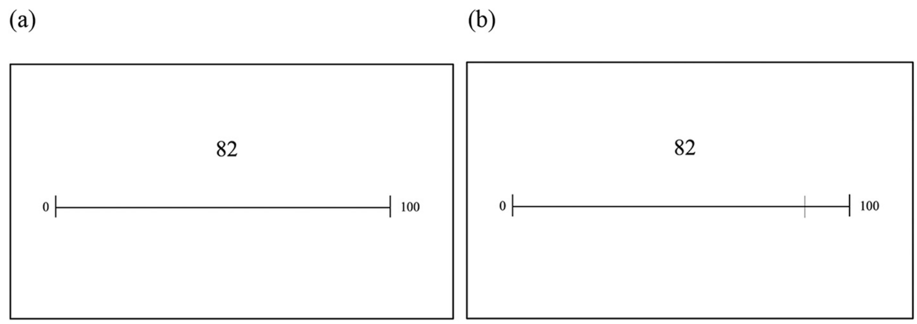

Fig. 1. Schematic of number line estimation display (a) before and (b) after response.
Note. Participants clicked on the horizontal line to estimate the location of the target number. The vertical placement line in the second image was red in color.

a power function $y = x^{\beta}$, where β quantifies the bias in estimates (the curve is concave when $\beta < 1$ and convex when $\beta > 1$).[2] Spence (1990) extended Stevens' Law to proportion judgments by showing that when there are two bounding reference points, estimates are predicted by $y = x^{\beta}/(x^{\beta} + (1 - x)^{\beta})$, with β determining the magnitude and direction of the bias (S-shaped when $\beta < 1$ and an inverse S-shaped when $\beta > 1$). Hollands and Dyre (2000) generalized the model to also explain a range of other patterns, such as multi-cyclical ones (e.g., Fig. 3d), that arise from use of additional reference points during proportion judgment.

The cyclical power model, although developed in the context of physical stimuli rather than numerical magnitudes, was successfully extended by Barth and Paladino (2011) and Cohen and Blanc-Goldhammer (2011) to the number line estimation task. In this context, the bias observed and reflected by β is thought to arise from imprecision in estimates of magnitude from numerals (e.g., Dehaene, Izard, Spelke, & Pica, 2008; Siegler & Opfer, 2003) and also from skill in judging part-whole relationships (e.g., 59 as a proportion of 100; Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011; Cohen, Blanc-Goldhammer, & Quinlan, 2018; Slusser et al., 2013). The observed pattern of bias is sometimes one cycle and sometimes multi-cyclical, with the number of cycles thought to depend on how many additional reference points, besides the labeled endpoints, are used to perform the task (e.g., also using the line's midpoint; Peeters, Sekeris, Verschaffel, & Luwel, 2017; Slusser et al., 2013; Sullivan, Juhasz, Slattery, & Barth, 2011; Zax, Slusser, & Barth, 2019). An individual's placements are typically well fit by either a one-cycle curve (more often inverse S-shaped as in Fig. 3a) or a two-cycle curve (more often S-shaped as in Fig. 3d; Cohen & Blanc-Goldhammer, 2011; Patalano, Saltiel, Machlin, & Barth, 2015; Slusser & Barth, 2017). While theoretical issues remain (e.g., why the pattern is S-shaped versus inverse S-shaped across individuals), the cyclical bias pattern itself is well-established in this context.

The cyclical power model (like other models that attempt to account for number line placements) does not, however, explain a more recently discovered bias in placements that is of a different nature. With the cyclical power model, target numerals like 79 and 81 are predicted to be placed in similar locations on a 0–100 number line because the two numerals have similar overall magnitudes. Recent work, however, has revealed that the individual digits that comprise the numerals are also importantly related to where the targets are placed on the line, beyond their contribution to overall magnitude. Specifically, what Lai, Zax, and Barth (2018) found using a 0–1000 number range was a *left digit effect*, whereby numerals with similar overall magnitudes but distinct leftmost digits (e.g., 299/301) were placed farther apart than is warranted. They found that this pattern did not occur with numerals that have distinct middle digits (e.g., 248/252), suggesting that the bias is driven by the leftmost digit. In studies such as Lai et al., targets are shown individually and in random order; pairing is for data analysis only. This left digit effect is robust and has been replicated using multiple numerical ranges (including 0–100 and 0–1000 ranges; Savelkouls, Williams, & Barth, 2020; Williams, Zax, Patalano, & Barth, 2022), a reverse number line (e.g., 1000–0; Williams, Bradley, Xing, Barth, & Patalano, 2022), a speeded task (Lai et al., 2018; Williams, Paul, Zax, Barth, & Patalano, 2020); and motivational incentives to perform as accurately as possible (Kayton et al., 2022).

To date, explorations of left digit effects in number line estimation have focused on a particular method of indexing these effects: showing that numerals such as 299/301 or 79/81 are on average placed in locations that are farther apart than is warranted. This method does not imply that left digit effects only exist for paired numerals that fall in the vicinity of either side of decades or hundreds boundaries. However, previous studies' focus on the vicinity of these boundaries does leave open the question: does the effect arise only when paired targets are only a few units away from left digit boundaries (e.g., 49/51, with a 0–100 range), or are more distant targets (e.g., 45/55) also placed farther apart than is warranted? Other aspects of left digit effects are also not well understood. For example, one might ask whether below-boundary values are placed farther to the left than warranted and above-boundary values farther to the right, or something else? Lai et al. (2018) did not formally test these possibilities, but their data visualizations suggest the former. Exploratory work by Kayton et al. (2022) goes further, offering a clue to how the left digit effect might arise from placements: it was observed that targets just above a boundary were on average placed nearly in the correct location (in fact, slightly too far to the left), while those just below a boundary were placed considerably farther to the left than warranted, suggesting that it might be the placements of below-boundary values driving the effect. This exploratory finding suggests a possible pattern of bias whereby, within each left digit range (10s, 20s, etc. for the 0–100 overall range), the placements of some or all targets are compressed farther to the left (closer to the placement of the lower boundary of the range) than is warranted. For example, on the 0–100 number line, some or all targets between 21 and 29 might be placed closer to where 20 is placed, and thus farther from where 30 is placed than is warranted (regardless of whether 20 and 30 themselves are placed accurately), with the pattern repeating in each tens range.

Because past studies have largely used only subsets of target numerals from the number range under study (e.g., 38 numerals from the 0–1000 range; Lai et al., 2018), past data cannot be easily used for modeling small-scale changes in placement error (across neighboring numerals). In the present study, we administer a 0–100 number line estimation task in which target numerals consist of *all* numerals in the range with the goal of better understanding the left digit effect. We use the collected data: (1) to replicate the phenomenon of the left digit effect on a 0–100 number line for target pairs used in past work; (2) to expand the analysis in order to further characterize the effect; and (3) to develop and test several modified versions of the cyclical power model that incorporate the proposed downward compression of placements within each tens range.

[2] A scaling parameter was also included in the original model.

2. Method

2.1. Participants and procedure

Participants were 60 adults (44 women, 16 men) who were recruited from Prolific Academic and completed the study online for small monetary compensation. Their ages ranged from 18 to 60 (M = 25.9). All participants completed, at their own pace, a total of 101 trials of a 0–100 bounded number line task. The trials consisted of all numerals from 0 to 100 (presented once each) in a different randomized order for each participant. Exclusion criteria, central inferential statistics, and mathematical models and model testing procedures were preregistered (https://aspredicted.org/4sz6s.pdf) unless otherwise noted.

2.2. Number line estimation task

2.2.1. Task description

The program was designed using lab.js software (lab.js. org; Henninger, Shevchenko, Mertens, Kieslich, & Hilbig, 2019) and distributed through Open Lab (Shevchenko, 2022). Participants completed a screen calibration task at the outset to ensure a consistent display size across computers (see Li, Joo, Yeatman, & Reinecke, 2020). On each trial of the number line task, participants were presented with a target numeral (e.g., 22; 1 cm tall) in the center of the screen, 4 cm above a black horizontal line (20 cm long) (see Fig. 1a). The horizontal line had two vertical lines as endpoints (1.1 cm long). The endpoints were labeled with 0 and 100 (each 0.8 cm tall). When participants selected the location of the target numeral with a mouse click, a vertical red line appeared (0.8 cm long) as show in Fig. 1b. A "Next" button then appeared that could be clicked to go on to the next trial. Participants were instructed to give their number line responses as quickly and accurately as possible. The click location was recorded as a number on the line to the nearest tenth of a unit (e.g., 45.7).

3. Results

3.1. Exclusions

In addition to being preregistered, all exclusion criteria were as in Lai et al. (2018) and Patalano, Williams, Weeks, Kayton, and Barth (2022). A participant's estimate for a target number was identified as an outlier for all analyses, except computation of overall accuracy error, if it differed from the group mean for that target by more than two standard deviations (4.09% of trials were removed). One participant with more than three tens pairs missing (as described in Section 3.3) due to outlier removal was excluded from all analyses. We also preregistered excluding any participant not completing the task or for whom the correlation between placements and target values was $r < 0.5$, but there were none. A total of 59 participants (out of the original 60) were in the final dataset.

3.2. Percent absolute error

To check that performance on the task was similar to what has been found in past work, we computed a common measure of overall accuracy error called *percent absolute error = (|target placement – correct target location|/100) x 100)*, summed over relevant trials. We found that percent absolute error was low (M = 3.81%, SD = 1.03, range = 1.99–6.54), comparable to what has been found in past studies of adults in 0–100 number line tasks (e.g., 3.60% using all trials; Williams, Zax et al., 2022). In exploratory analyses, we also calculated percent absolute error using two subsets of trials: one version used the half of trials closest to tens boundaries (0–2, 8–12, 18–22 through 98–100) and the other used the half farthest from tens boundaries (3–7, 13–17, 23–27 through 93–97). We computed these because the question often arises as to the extent to which percent absolute error varies as a function of whether boundary trials are heavily represented in the stimulus set. If targets near boundaries are most susceptible to a left digit bias, one might predict error to be greater when there are more of these targets in the set. We found that percent absolute error for targets near boundaries (M = 3.60%) was actually lower than for targets far from boundaries (M = 4.03%; $t(58) = 5.15$, $p < .001$, $d = 0.67$), but that the direction reversed when targets around reference points on the line (targets 0–2, 48–52, and 98–100, which are straightforward to place accurately) were excluded. In the latter case, percent absolute error was higher for targets near boundaries (M = 4.35%) than for targets far from boundaries (M = 4.03%; $t(58) = -3.43$, $p = .001$, $d = 0.45$), with a medium effect size. The finding suggests that target selection matters in that overall accuracy error is low near reference points, but is otherwise higher for targets close to (versus far from) tens boundaries.

3.3. Average tens difference scores

We next computed a standard measure of the left digit effect for two-digit targets called the *average tens difference score*. The target values we used to compute the measure were the same as those used in past adult studies with a 0–100 range, even though the present study design allows for other possibilities as well. There were eight pairs of values (paired for the purposes of data analysis only: 18/22, 29/31, 38/42, 49/51, 59/61, 69/71, 78/82, 88/92; Vaidya et al., 2022, Williams, Zax et al., 2022). Following past work, we computed the *average tens difference score = ((larger target placement – smaller target placement) – (correct larger target location – correct smaller target location))*, averaged over all of the tens pairs. An average tens difference score above 0 is interpreted as evidence for a left digit effect, because the difference is greater than what would be expected based on the magnitudes of the numbers alone.

As expected, the average tens difference score (M = 0.76, SD = 1.87, range = −3.96–5.68) was significantly different from (greater than) 0, $t(58) = 3.11$, $p = .003$, two-tailed, $d = 0.41$, consistent with a left digit effect. The magnitude of the tens difference score was within the range of means reported in past work (Ms = 0.74–1.67; Vaidya et al., 2022; Williams, Zax et al., 2022), and was in the predicted direction for 6 out of 8 pairs (all except 18/22 and 59/61; see Fig. 2a), and 36 out of the 59 participants (61%; binomial test, $p = .118$, two-tailed). The number of participants showing the expected pattern was close to the ~70% seen in past work with the 0–100 number line (e.g., Williams, Zax et al., 2022), and lower than the ~85% seen with the 0–1000 number line (e.g., Patalano et al., 2022; Williams et al., 2020). We suspect most people are susceptible to the bias, but that the effect is smaller and more often obscured by other sources of error at the individual level, with the 0–100 scale. However, see Williams et al. (2020) for a discussion of possible individual differences in the left digit effect as well.

To demonstrate that a difference score >0 is specific to numbers that cross a left digit boundary, and that it does not also arise for other pairs of numbers, we also computed an *average fives difference score* for eight pairs of two-digit values that cross a fives boundary (13/17, 24/26, 33/37, 44/46, 54/56, 64/66, 73/77, 83/87) that we matched to the tens pairs on distances between values in each pair.[3] In this case, we would *not* expect the five difference score to be >0 since the pairs do not cross a left digit boundary. If anything, we would expect it to be <0 because the placements might be spaced too close together (more compressed as one moves towards the lower end of the tens range). We found that the average fives difference score was in fact significantly <0 (M = −0.84, SD = 2.10, range = −5.58–3.06; $t(58) = -3.07$, $p = .003$, two-tailed, d = 0.41), showing no evidence of a left digit effect. This pattern extended to 7 out of 8 pairs (all except 44/46; see Fig. 2b) and to 38 out of 59

[3] We report the findings using the eight pairs surrounding boundaries from 15 to 85 but the results would not change if we were to instead use the eight boundaries from 25 to 95 (i.e., the ones above rather than below each tens pair).

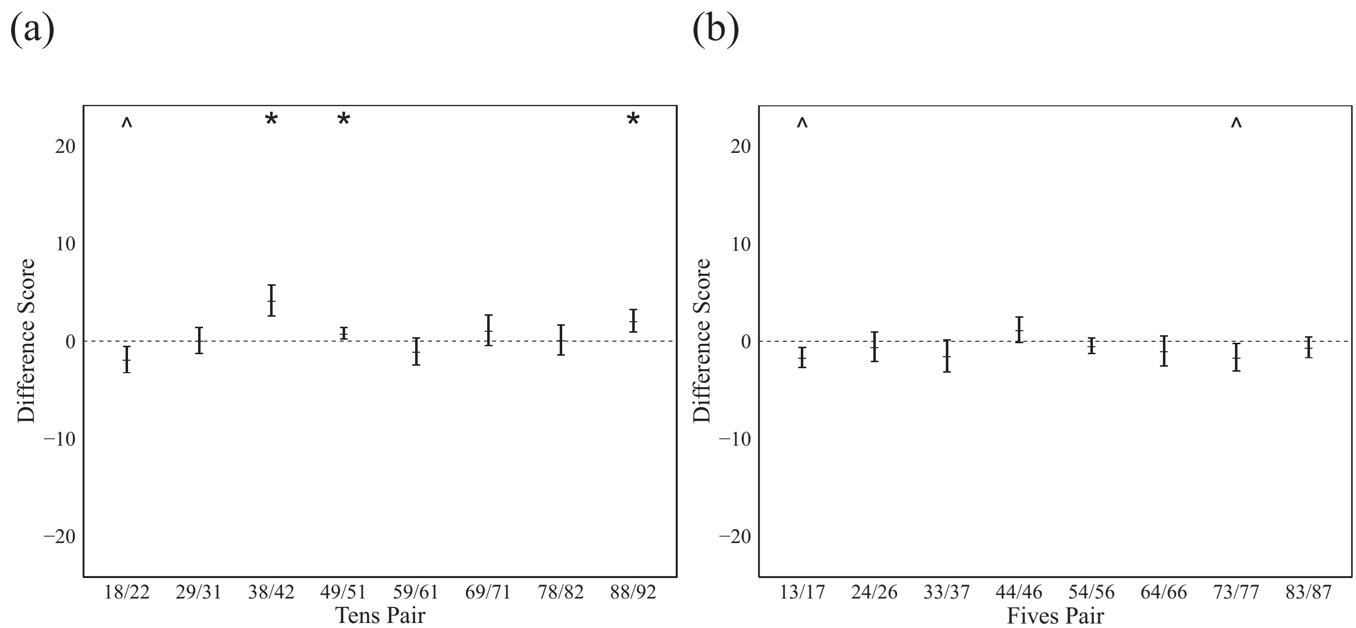

Fig. 2. (a) Average tens difference scores and (b) average fives difference scores by target pair.
*$p < .05$, two-tailed, positive direction; ˆ$p < .05$, two-tailed, negative direction. *Note.* Bars reflect 95% confidence intervals. Tens difference scores were in the predicted direction (>0) for 6/8 pairs, consistent with a left digit effect. Fives difference scores were in the reverse direction (<0) for 7/8 pairs, evidence that the left digit effect does not extend to fives pairs.

participants (64%; binomial test, $p = .036$, two tailed). This finding shows that the tens difference score >0 cannot be explained by a more global pattern of bias (such as the known S-shaped curve) because, if the latter were true, one would expect the pattern to extend to the fives difference score as well. The fives difference score pattern seen here is consistent with the proposed downward compression of placements within a tens range in that it shows that pairs around fives boundaries are, in fact, placed closer together than is warranted.

One question presented in the introduction was that of whether the tens difference score is reliably >0 only for targets close to the tens boundary or if it extends to other pairs that cross such a boundary (e.g., 52 and 68 cross the boundary of 60 but are rather far from it). This is also partly a methodological question: Can any pairs that cross a left digit boundary be used to assess the left digit effect? To address this question, we computed average tens difference scores for symmetrical pairs at various distances from the tens boundaries (with each target being from one unit away to nine units away; e.g., 59/61, 58/62, through 51/69), as shown in Table 1. What can be seen is that the tens differences scores are reliably >0 for all pairs through those with five-ending targets (e.g., 55/65) and then the effect is no longer present after that point (e.g., it is not present for targets such as 54/66). This pattern suggests that the effect may extend to targets that are within about five units from the boundary. The finding is important in practice for informing stimulus selection when designing studies.

The analyses so far illustrate where the left digit effect emerges but do not yet speak to whether it is driven to a greater extent by above-boundary or below-boundary placements. For example, is the tens difference score for 59/61 driven by the difference in placements between 59 and 60, between 60 and 61, or both? To address this question, in Table 1, we broke down each tens difference score into lower and upper components. For example, for 59/61, we computed a lower difference score for 59/60 as *(placement of 60 – placement of 59) – (correct location of 60 – correct location of 59)* and an upper distance score for 60/61 as *(placement of 61 – placement of 60) – (correct location of 61 – correct location of 60).* If the left digit effect is in fact driven by the placement distance for 59 and 60 rather than for 60 and 61 (i.e., if it is driven by numerals with different leftmost digits), lower tens difference scores should be >0, but this should not be the case for upper difference scores. This is the pattern that we observed. As shown in Table 1, lower tens difference scores were all reliably >0, while upper tens difference score were all reliably <0 (both patterns consistent with compression of placements within a tens range towards the left on the number line). A left digit effect emerged (a tens difference score reliability >0) whenever the lower (positive) difference score was much greater in absolute magnitude than the upper (negative) difference score.

Table 1
Average tens difference scores (and upper and lower components) as a function of distance between paired targets.

Distance (units per direction)	Example	Tens difference score (e.g. 59/61)	Lower difference score (e.g. 59/60)	Upper difference score (e.g. 60/61)
1	59/61	0.99*	1.73*	−0.84ˆ
2	58/62	1.07*	2.14*	−1.09ˆ
3	57/63	0.82*	2.35*	−1.33ˆ
4	56/64	0.83*	2.33*	−1.53ˆ
5	55/65	0.47*	1.95*	−1.67ˆ
6	54/66	−0.10	1.94*	−2.18ˆ
7	53/67	−0.30	1.48*	−1.73ˆ
8	52/68	−0.24	1.49*	−1.82ˆ
9	51/69	−0.16	0.99*	−0.87ˆ

Note. The averages includes all pairs around each of nine tens boundaries (pairs around 60 are just one example). The table shows a left digit effect for distances of up to five units in each direction (e.g., 55/65), arising when positive lower difference scores that are not fully offset by negative upper ones.
*$p < .005$, two-tailed, positive direction.
ˆ$p < .005$, two-tailed, negative direction.

3.4. Modeling of placement data

To model placement data, we used standard one-cycle and two-cycle versions of the cyclical power model, and we tested two different types of modifications to accommodate the left digit effect. The first type of new model that we developed (all new models have both one and two-cycle versions) is referred to as the *modified (cyclical power) model.* This model is different from the standard model in that we added a transformation of the rightmost digit of a numeral to accommodate the left digit effect. A power function for the transformation was preregistered (this is the *power-modified* version of the model), and we also tested a multiplier function (the *multiplier-modified* version of the model) after evaluation of the data. We then also developed and tested a second type of model called an *expanded (cyclical power) model*. This model is different from the modified model in that we extended the digit transformation from just the rightmost digit to *all* digits that comprise a numeral. We used only the power function (and thus also refer to the model as a *power-expanded* model) for the transformation because the power function is more strongly motivated in this context, as will be described later. Note that while tables and graphs show findings for all models together, in the text we first describe and present findings for the modified model, and then we describe and present findings for the expanded model, consistent with the manner in which we developed these models.

3.4.1. Descriptions of standard and modified cyclical power models

Recall that the standard cyclical power model was previously developed (Hollands & Dyre, 2000) and used to capture the global cyclical pattern shown in Fig. 3a and d. The standard (unmodified) cyclical power equations are shown first below, where y = predicted target placement, x = target value, β = index of curvature, UB = upper boundary, and LB = lower boundary (starting value was $\beta = 1$):

$$\text{One}-\text{cycle}: y = \left((x\text{–LB})^{\beta} \Big/ \left((x\text{–LB})^{\beta} + (\text{UB}–x)^{\beta}\right)\right) \bullet 100$$

where $\beta > 0;\ \text{LB} = 0,\ \text{UB} = 100$

$$\text{Two}-\text{cycle}: y = \left((x\text{–LB})^{\beta} \Big/ \left((x\text{–LB})^{\beta} + (\text{UB}–x)^{\beta}\right) \bullet 0.5\right) \bullet 100 + \text{LB}$$

where $\beta > 0$; for $0 \leq x \leq 50$, $\text{LB} = 0,\ \text{UB} = 50$; for $50 < x \leq 100$, $\text{LB} = 0,\ \text{UB} = 50$

For the modified version of the one-cycle and two-cycle model, we made two changes. First, the target value, rather than being represented by x, now has separate input for its tens place value (x_t) and its ones place value (x_o). For example, 25 would be input as $x_t = 20$ plus $x_o = 5$. Second, we added a function $f(x_o)$ that allows for an underweighting of the ones place value before it is added to the tens place value, resulting in lower placements of targets within each tens range. The below equations show the function written generically as $f(x_o)$; we follow these equations with the two specific functions that we used in the present work.

$$\text{One}-\text{cycle modified}: \quad y = \left(((x_t + f(x_o))\text{–LB})^{\beta} \Big/ \left(((x_t + f(x_o))\text{–LB})^{\beta} + (\text{UB}–(x_t + f(x_o)))^{\beta}\right)\right) \bullet 100$$

$$\text{Two}-\text{cycle modified}: \quad y = \left(((x_t + f(x_o))\text{–LB})^{\beta} \Big/ \left(((x_t + f(x_o))\text{–LB})^{\beta} + (\text{UB}–(x_t + f(x_o)))^{\beta}\right) \bullet 0.5\right) \bullet 100 + \text{LB}$$

Given no strong prediction about the precise form of the underweighting, we tested the models using a power function and again using a multiplier function (starting value was $\delta = 1$). We treated the parameter as essentially means of weighting the relative contribution of the ones digit to the overall magnitude of the multidigit numeral, consistent with frequent description of the left digit effect as arising from the weighting of digits during the conversion of numerical symbols to magnitudes (e.g., Thomas & Morwitz, 2005).

$$\text{Power function}: f(x_o) = {x_o}^{\delta}$$

$$\text{Multiplier function}: f(x_o) = \delta x_o$$

If the exponent δ were set to 1 in these functions, the modified cyclical models would be identical to the standard model. However when $\delta < 1$, the ones value of the target contributes less than it should to the target's overall magnitude, resulting in lower placements of targets within a tens range. Only the predicted placements of tens boundary targets (like 10, 20, 30) do not change in the modified models relative to the standard ones because, when the input is 0, $f(x_o)$ returns 0. The overall result is compression of placements within each tens range towards placement of the lower boundary and away from the upper boundary (e.g., placing 25 closer to 20 and farther from 30), as illustrated in Fig. 3b and e for the multiplier-modified model. The power and multiplier functions are similar except that the former predicts greater downward compression. Note that while we did not formally set $\delta \leq 1$, we assume that people typically do not treat the magnitude of any numeral between 0 and 100 as greater than that of the next larger numeral (e.g., that 99 is not treated as larger than 100). Because this assumption is sometimes violated when $\delta > 1$, we expect δ to be ~1 when there is no left digit effect, otherwise <1.

3.4.2. Fitting of standard and modified cyclical power models to data

To model placement data using the described models, we first assigned each participant to a one-cycle group or a two-cycle group based on which standard cyclical power model better fit their responses (i.e., which produced a lower Bayesian Information Criterion, or BIC, score for that individual).[4] We used this approach because the purpose of the work was not to compare fit of these versions of the model (that is, we assumed that some participants would be better fit by a one-cycle version, and some by the two-cycle version, depending on their task strategy). Rather, the goal was to assess whether people who use each strategy are better fit by the standard model versus the corresponding modified model. Further, by grouping participants from the outset, we were able to compare models based on their fit to median responses knowing that the individuals used to compute the medians would be the same across models (e.g., the one-cycle standard and modified models would be tested using the same median responses). What we found was that more participants were better fit by a standard one-cycle model (n = 42) than by a two-cycle model (n = 17), consistent with past work (e.g., Cohen & Blanc-Goldhammer, 2011; Patalano et al., 2020; but see Slusser & Barth, 2017, for a reverse pattern in a very large number range), and all participants were better fit by these models than by a simple identity model.

Within the one-cycle and two-cycle groups separately, we then fit modified models to individual participants' data (which we had already done for the standard model), and also fit both standard and modified models to median placements. As shown in Table 1, both the power and multiplier versions of the modified model fit considerably better than the standard model, with the multiplier-modified version fitting somewhat better than the power-modified one. For the one-cycle group, for the multiplier-modified model fit to median placements, ΔBIC was 81 relative to the standard model, and 83% of participants were individually better fit by the multiplier-modified model than by the standard model. Similarly, but to a lesser degree, for the two-cycle group, for the multiplier-modified model fit to median placements, ΔBIC was 36 relative to the standard model, and 76% of participants were individually better fit by the multiplier-modified model than by the standard model. Descriptive statistics (e.g., mean parameter estimates) for models

[4] The conclusions do not change with use of AIC instead of BIC; we report the latter, at a reviewer's suggestion, because the latter generally better compensates for the number of parameters in the model.

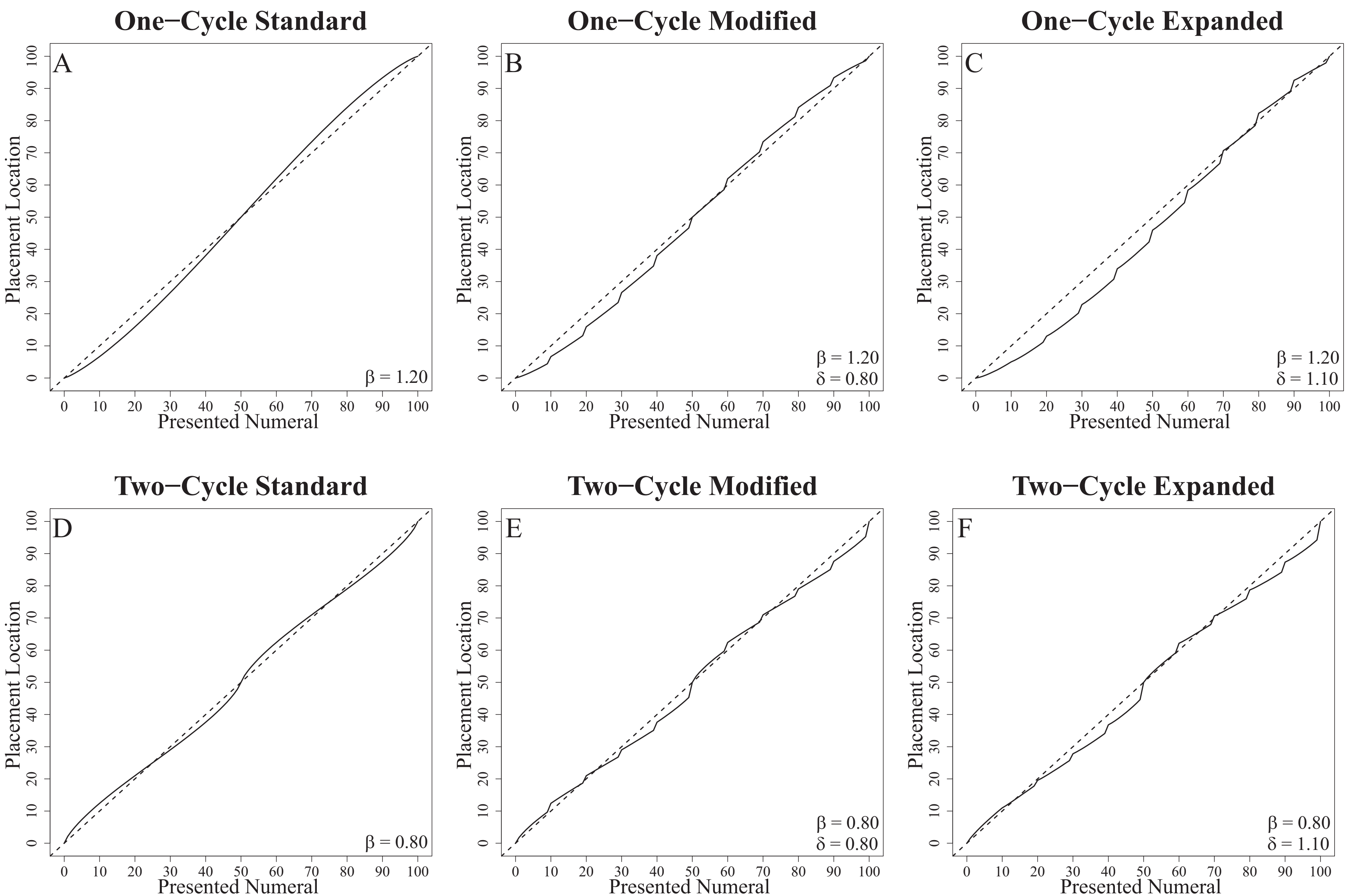

Fig. 3. Predictions of cyclical power model and proposed modified and expanded models.
Note. The multiplier-modified models reflect compression of values within each tens range towards the lower boundary of the range (producing the jumps from 29 to 30, 39 to 40, etc.). (The power-modified version would look similar except that the drops would be steeper.) The expanded models additionally show a compression of tens ranges towards the lower boundary of the full number line (producing lower placements across the 0–100 line).

Table 2
Fits of standard and modified versions of each cyclical model to median placements and to individual participants' data.

	Using group medians				Using individual data	
Model type	n	β	δ	BIC	Preferred over standard model[a]	δ consistent with left digit effect[b]
Identity	59	–	–	214	–	–
One-cycle standard	42	1.16	–	175	–	–
One-cycle power-modified	42	1.16	0.81	109	83%	100%
One-cycle multiplier-modified	42	1.16	0.68	94	83%	98%
One-cycle power-expanded	42	1.15	1.05	82	86%	98%
Two-cycle standard[c]	17	0.90	–	146	–	–
Two-cycle power-modified	17	0.91	0.88	118	71%	100%
Two-cycle multiplier-modified	17	0.91	0.78	110	76%	94%
Two-cycle power-expanded	17	0.92	1.07	90	82%	100%

Note. Each participant was included in either the one-cycle or two-cycle group based on which standard model better fit their data (no participants were better fit by the identity model).

[a] Percentage of participants whose data were better fit by the revised model than by the standard one.

[b] Left digit effect is consistent with $\delta < 1$ for modified model, and $\delta > 1$ for the expanded one.

[c] Participants were placed in the two-cycle group based on the fit of the standard model (per our preregistered procedure), but 5/17 participants were better fit by a one-cycle modified model than by any two-cycle models, and 3/17 were better fit by a one-cycle expanded model than by any two-cycle models. Thus, the table findings may slightly underestimate the number of participants best fit by modified and expanded models.

fit to individual data are in Appendix A.

As also shown in Table 2, the parameter estimates for all models are highly similar, and their values are as expected. For the parameter β, which captures degree of curvature globally, parameter estimates for median placements (one-cycle $\beta = 1.16$; two-cycle $\beta = 0.90$ are similar to those found in past work (e.g., βs = 1.10 and 0.96 respectively, in Slusser & Barth, 2017; see also Patalano et al., 2020). For the parameter δ, which captures the underweighting of the ones digit (when $\delta < 1$) in the modified models, parameter estimates are <1 for median placements (e.g., $\delta = 0.68$ for the one-cycle modified model using the multiplier function) and for nearly all individual participants, consistent with a leftward compression of placements within each tens range. The parameter estimates do not differentiate among models, but do provide further evidence that the preferred model based on BIC values (i.e., the modified model that uses the multiplier function) fits as expected with regard to parameter estimates.

We further illustrate the fit of the modified model graphically. We focus here on the version using the multiplier function though the pattern would be similar for the one using the power function. Fig. 4illustrates the fit of the standard and modified one-cycle models to median placements for participants best fit by the one-cycle model, while Fig. 5illustrates the same for the two-cycle model. Note that there are two types of graphs shown: placement graphs plot placement location (i.e., participant response) against correct location of target numeral, while error graphs replace placement location with placement error. As shown in Fig. 4b, the standard one-cycle model does not fit well because it predicts similar magnitudes of error on both halves of the number line. The modified one-cycle model (Fig. 4d) fits better because it predicts larger errors on the first half of the line (where cyclical bias and left digit bias have an additive effect) and smaller errors on the second half (where the two biases essentially cancel one another out), consistent with the data. The same is true of the modified two-cycle model (e.g., Fig. 5b and d) to a lesser extent. While the peaks in the median placements do not always align precisely with model predictions, and errors between targets 30 and 40 are larger than predicted, the modified models provide a considerable improvement in fit.

3.4.3. Additional exploratory work: Expanded cyclical power model

In conducting our planned modeling work, we observed that that the standard cyclical power model did not predict the observed large underplacement on the left half of the number line relative to the much smaller overplacement on the right half of the line. For this reason, we considered one additional new model, an expanded cyclical power model.[5] This model builds on the modified model already presented but is also able to accommodate the asymmetrical curvature pattern also observed. Recall that in the modified model, a transformation of the ones digit reduced the contribution of this digit's value to the overall value of the target numeral. In the expanded model we present now, we used a power transformation of *all* digits, rather than only the rightmost digit. So, for example, for '25', instead of only the 5 being raised by δ, the 20 (the tens place value) would also be raised by δ. The utility of such a transformation is that in addition to accommodating downward compression within each tens range, it can also accommodate compression of all tens ranges towards the lower end of the number line, similar to what was observed in the data. This expanded model (so named because it is an expansion of our modified model) has the potential to better fit the data, and it is in the spirit of *decomposed theories* of numerical cognition in which the magnitude associated with each digit is accessed individually (e.g., Moeller, Huber, Nuerk, & Willmes, 2011; Verguts & De Moor, 2005).

The equation for the one- and two-cycle versions of the expanded model are shown below. All abbreviations and starting values are the same as for the modified model unless otherwise noted. What is new about the equations is as follows. First, there is now a transformation not just of the ones (x_o) place value for each numeral, but also of tens (x_t) and hundreds (x_h) place values (the latter is included just to accommodate '100'). Second, we previously did not include any transformation of the ones digit for lower and upper boundaries (LB and HB) because 0, 50, and 100 (the boundaries used here) all had a 0 for the ones digit and so would not be affected by any transformation. Now, because tens and hundreds digits are also transformed, it is important to explicitly include transformation of boundary values in the equation. Third, the last symbol in the two-cycle equation, which was previously LB, has been replaced with CS, standing for 'cycle start'. The reason for the change is that, in the cyclical model, it is assumed that the cycle is repeated for each half of the line, so the starting points of the cycles must be 0 and 50 if the pattern is to repeat at the halfway point. Essentially, in

[5] We thank Dale J. Cohen for suggesting this model.

the below model, CS represents the true physical start or midpoint of the line rather than a numeral.

$$\text{One}-\text{cycle expanded}: y = \left(\left((f(x_h)+f(x_t)+f(x_o))-(f(LB_h)+f(LB_t)+f(LB_o))\right)^{\beta} \Big/ \left(\left((f(x_h)+f(x_t)+f(x_o))-(f(LB_h)+f(LB_t)+f(LB_o))\right)^{\beta} + \left((f(UB_h)+f(UB_t)+f(UB_o))-(f(x_h)+f(x_t)+f(x_o))\right)^{\beta}\right)\right) \bullet 100$$

$$\text{Two}-\text{cycle expanded}: y = \left(\left((f(x_h)+f(x_t)+f(x_o))-(f(LB_h)+f(LB_t)+f(LB_o))\right)^{\beta} \Big/ \left(\left((f(x_h)+f(x_t)+f(x_o))-(f(LB_h)+f(LB_t)+f(LB_o))\right)^{\beta} + \left((f(UB_h)+f(UB_t)+f(UB_o))-(f(x_h)+f(x_t)+f(x_o))\right)^{\beta}\right) \bullet 0.5\right) \bullet 100 + CS$$

For the transformation, we chose to use the power function here ($f(x) = x^{\delta}$) even though the multiplier function was preferred earlier. We did this because, in the present context, the power function is more straightforwardly motivated than the multiplier function. With the earlier modified model, the transformation could be thought of as a relative weighting of the ones and tens values in the estimation of overall magnitude, and the relative weighing could be achieved in many ways including through a power or multiplier function. However, with the expanded model, the transformation may be better thought of as similar to the magnitude bias associated with the whole numeral and represented by β, now simply extended to individual digits (although still using a separate parameter δ since the transformation may not be the same for digits). Unlike the modified model in which $\delta < 1$ produced downward compression, in the expanded model downward compression occurs when $\delta > 1$. This is because both the target value and the whole (e.g., 75 as a part of 100) are transformed. When $\delta > 1$, the transformation results in a greater increase in the estimate of the whole than of the target, and so the target is predicted to be placed lower on the line (it is a smaller proportion of the range) than when $\delta = 1$. While we did not formally set $\delta \geq 1$ here, we again (as with the modified model) assume that people do not treat the magnitude of any numeral between 0 and 100 as greater than that of the next larger numeral. Because this assumption can be violated when $\delta < 1$ here, we expect that δ will be ~ 1 when there is no left digit effect, otherwise $\delta > 1$.

3.4.4. *Fitting of the expanded cyclical power model to data*

To model placement data using the described models, within the one-cycle and two-cycle groups separately, we computed median placements and fit the power-expanded model to both median placements and individual participants' data. As shown in Table 2, the expanded model fit the data considerably better than the standard model. For the one-cycle group, for the expanded model fit to median placements, ΔBIC was 93 relative to the standard model, and 86% of participants were individually better fit by the expanded model than by the standard model. Similarly, but to a lesser degree, for the two-cycle group, for the expanded model fit to median placements, ΔBIC was 56 relative to the standard model, and 82% of participants were individually better fit by the expanded model than by the standard model. Descriptive statistics for models fit to individual data are in Supplementary Materials.

As also shown in Table 2, the parameter estimates for the model are as expected. For the parameter *β,* which captures degree of curvature globally, parameter estimates for median placements (one-cycle $\beta = 1.15$; two-cycle $\beta = 0.92$ are similar to those found in the standard and modified models. For the parameter *δ,* which is consistent with a left digit effect when $\delta > 1$ in the expanded model, parameter estimates are >1 for median placements ($\delta = 1.05$ for the one-cycle model and $\delta = 1.07$ for the two-cycle model), and for nearly all individual participants, consistent with a leftward compression of placements both within each tens range and across the whole number range. The parameter estimates do not differentiate among models, but do provide further evidence that the expanded model fits as expected with regard to parameter estimates.

The power-expanded model also fit better than the multiplier-modified model (the best fitting model to this point). For the one-cycle group, the ΔBIC associated with the expanded model fit to median placements was 12 (relative to the multiplier-modified model), and 67% of participants were better fit by the expanded model. Similarly, for the two-cycle group, the ΔBIC associated with the expanded model fit to median placements was 20, with 94% of participants better fit by the expanded model than by the multiplier-modified model. The asymmetrical shape of the curve of the expanded model (now produced largely by the transformation of the tens digit rather than the ones digit) generally matched the data well. However, as suggested by Figs. 4 and 5, the expanded model's improved fit is, in part, in number ranges already quite well fit by the modified model (e.g., close to 0 and to 100), rather than in ranges in which the multiplier-model fit poorly (e.g., in the 30–40 range). And, in some ranges, such as the 40–80 range of the one-cycle model, the expanded model considerably underestimated placements. In sum, the expanded model did provide a better fit overall, but some challenges remain for future work in terms of replicating and explaining some of the more extreme response errors in performance.

At this time, to ensure our findings were not specific to the procedure we used to compare models, we also simply counted the number of participants for whom each model was preferred, looking across all versions of one- and two-cycle models (i.e., not breaking down participants by cyclical group). The findings, as shown in Table 3, are that the expanded model was the preferred model for 73% of all participants, consistent with other findings.

3.5. *Additional exploratory analyses*

Given that we knew the size of each participant's tens difference score, and the estimated δ parameter for each participant's best fitting model between the one-cycle modified and two-cycle modified model, we considered the relationship between these two measures. If the model well captures the left digit effect, individuals with larger tens

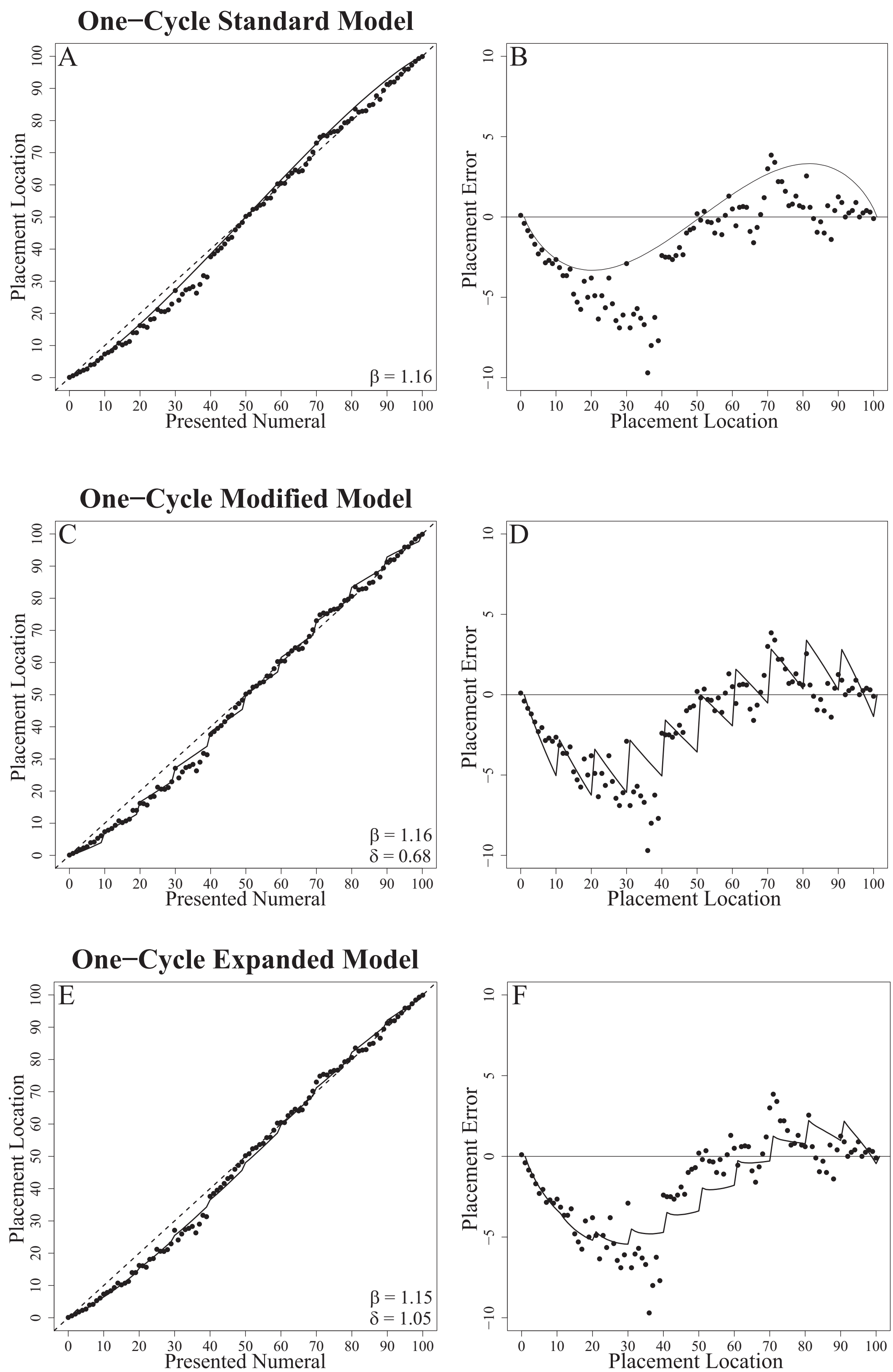

Fig. 4. One-cycle standard, multiplier-modified, and expanded model fit to median placements.
Note. The above graphs fit the models to the median responses for the group best fit by a one-cycle model ($n = 42$). Relative to the standard model, ΔBIC was 81 for the multiplier-modified model, and ΔBIC was 93 for the expanded model, indicating improved fits.

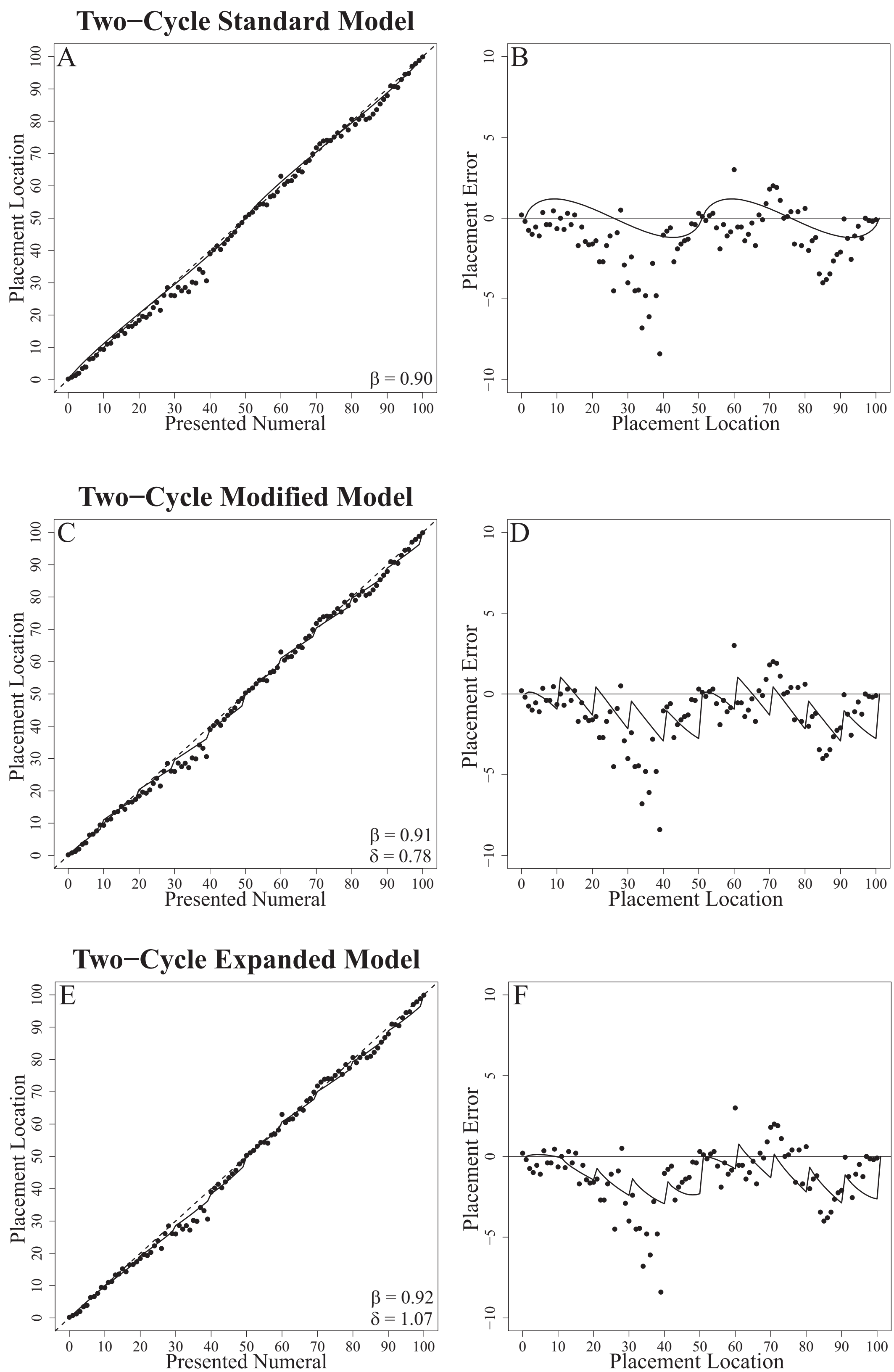

Fig. 5. Two-cycle standard, multiplier-modified, and expanded model fit to median placements.
Note. The above graphs fit the models to the median responses for the group best fit by a one-cycle model ($n = 17$). Relative to the standard model, ΔBIC was 36 for the multiplier-modified model, and ΔBIC was 56 for the expanded model, indicating improved fits.

Table 3
Number of participants best fit by each model.

Model type	Best Fit
One-cycle standard	6
One-cycle power-modified	1
One-cycle multiplier-modified	6
One-cycle power-expanded	26
Two-cycle standard	3
Two-cycle power-modified	0
Two-cycle multiplier-modified	0
Two-cycle power-expanded	17

Note. The majority (73%) were best fit by the power-expanded model.

difference scores might have lower (closer to 0) δ values. The correlation, although suggestive, was not statistically significant for either the multiplier-modified model ($r(58) = -0.16$, $p = .242$), or the power-modified model ($r(58) = -0.21$, $p = .144$), and was close to 0 for the power-expanded model ($r(58) = 0.06$, $p = .649$; for $|\text{skewness}| < 1.10$ for all variables); a larger sample would be needed to detect a relationship of these sizes. We also tested whether the participants better fit by the two-cycle model might have a smaller left digit effect than those best fit by the one-cycle model, consistent with the two-cycle pattern reflecting use of a more sophisticated placement strategy. The difference was also not statistically significant (*M*s = 0.68 vs. 0.79 respectively); $t(57) = 0.20$, $p = .846$.

4. Discussion

In the present study, we collected number line estimation data for all target numerals between 0 and 100 on a bounded number line. Our goals were to replicate and further understand the phenomenon of the left digit effect in the 0–100 context as measured in prior work through average tens difference scores, to assess whether target placements are compressed to the left for each tens range, and to develop and test modifications of the cyclical model to accommodate the proposed compression. What we found, first, was a replication of the phenomenon of the left digit effect: targets with different leftmost digits were placed farther apart than warranted. Second, we expanded our understanding of the effect in multiple ways, including that the phenomenon extends to pairs as many as five units away from boundaries (e.g., 45/55), and that the effect is driven by larger than warranted differences in placements of below-boundary and boundary targets, whereas placements of boundary and above-boundary targets are actually less different than warranted. Third, we found that our two new versions of the standard cyclical power model—a modified version that underweights the ones digit of targets and an expanded version that transforms all digits—both predicted placements considerably better than the standard model, with the expanded version providing the better fit overall.

Recall that the standard cyclical power model of Hollands and Dyre (2000) was developed to explain estimates of physical magnitudes in proportional relationships (e.g., relative length of a tone, proportion of red to blue dots in a display), and does so quite successfully. In the model, the parameter β indexes the degree and direction of cyclical bias, and the bias is attributed to the use of imprecise magnitude estimates in proportion judgments. What warranted modifying the model here was the past extension of the model to symbolic magnitude representations (Barth & Paladino, 2011; Cohen & Blanc-Goldhammer, 2011), and the discovery of a bias specific to these representations (Lai et al., 2018), namely, the left digit effect. The versions of cyclical power models developed in the present work, in which digits are input individually so that they can be independently transformed (e.g., by δ), thus have more specific application than the original cyclical power model. However, the modifications are particularly important for modeling patterns of left digit bias, as the latter are most easily observed as deviations of placements from a cyclical pattern (rather than from an identity line).

4.1. Modified and expanded cyclical power models

Our modified cyclical power model, the first model we tested, fit better than the standard model and explained a number of patterns in placement data. In addition to predicting the large-scale cyclical pattern and the compression of placements within each tens range, the model also predicted the seemingly greater S-shaped curvature on the left side of the number line than on the right side (a pattern also observed in Lai et al., 2018). The model explains this pattern as arising from the additive effects of the two biases: when both biases (cyclical bias and left digit bias) are in the direction of underestimation, underplacement is very large. However, when one bias is in the direction of overestimation and the other is towards underestimation, the biases essentially offset one another and placements remain close to the identity line. While the new model predicts these central patterns, there were some local patterns not fully predicted, most notably large underplacements of targets in the 30 to 40 range.

The second model we tested, the expanded cyclical power model, fit the data somewhat better than the modified model. Like the modified model, the expanded model also predicted the large-scale cyclical pattern, the downward compression of placements within each tens range, and the asymmetrical S-shaped curve. Unlike the modified model, the expanded model explains the asymmetrical pattern as arising largely from an overestimation of the magnitudes of place values of numerals, including boundary numerals, resulting in a lowering of the cyclical curve across the number line (e.g., 20/100 = 20%, but $20^{1.1}/100^{1.1}$ = 17%). In this model, the left digit effect arises not because the ones digit is multiplied by a $\delta < 1$ but, rather, because the overweighting of the tens value leads to large increases in predicted placement across left digit boundaries. While this model does fit better than the modified model, and better captures the overall shape of the global curve, areas for improvement remain (again including the 30–40 range). We further consider both types of models in the remainder of the discussion.

4.2. Interpretation of the parameter δ

The modified and expanded models were motivated by the description of the left digit effect as arising from the weighting of digits during the conversion of numerical symbols to magnitudes (e.g., Thomas & Morwitz, 2005). Thomas and Morwitz suggested that the left digit effect might require a simple modification of a holistic cognitive model of number processing (Brysbaert, 1995; Dehaene, Dupoux, & Mehler, 1990) to accommodate the weighting. In holistic models, each multi-digit number is recognized as a whole and mapped onto a single internal quantity. Such models are consistent with some number line data, including finger-process-tracing findings of early finger movement towards an imprecise holistic magnitude (Dotan & Dehaene, 2013; see also, e.g., Ganor-Stern, Pinhas, & Tzelgov, 2009; Reynvoet & Brysbaert, 1999). However, because holistic models do not explicitly represent individual digits or place value, they cannot be straightforwardly modified to accommodate the left digit effect.

Holistic models have been found to be insufficient for explaining other findings as well, such as a well-known decade-unit compatibility effect in number comparison tasks in which judgment (of the larger overall magnitude) is faster when tens- and ones-digit comparisons suggest the same response (e.g., 42 vs. 57) over different ones (e.g., 57 vs. 62; Nuerk, Weger, & Willmes, 2001; Nuerk, Weger, & Willmes, 2004). A range of models that accommodate such findings include decomposed models in which only quantities associated with individual digits are activated and manipulated (Moeller et al., 2011; Verguts & De Moor, 2005), hybrid models in which both digit and holistic magnitudes are available (Nuerk et al., 2001; Nuerk & Willmes, 2005), and models in which digit magnitudes are integrated into an overall magnitude (Dotan & Dehaene, 2020). It has also been proposed that one might adopt either a holistic or a decomposed strategy based on task demands (e.g., Dotan & Dehaene, 2013). The present work does not speak to whether digit-level magnitudes are directly available and used to guide placements versus that they are used to compute overall magnitude but, at minimum, shows that digit-level information contributes to placements.

A recent number-to-quantity model (Dotan & Dehaene, 2020; see also McCloskey, 1992; McCloskey, Sokol, & Goodman, 1986) is useful for reflecting on interpretations of δ. This model integrates components of holistic and decomposed approaches, and explicitly represents place value in the conversion of numerals into overall magnitude. Specifically, digits are multiplied by place units following base ten rules (e.g., 25 = 2*10 + 5*1), before these component magnitudes (which are also available to guide task performance) are merged into a whole number quantity. While the model does not incorporate imprecision, there are at least two places where it might be added: place weighting and component magnitude estimation. Imprecision in place weighting (e.g., replacing the *1 with *0.8) is consistent with the findings that, although place rules are explicitly learned, place value information is also implicitly acquired (Yuan, Prather, Mix, & Smith, 2020) and frequently accessed automatically (e.g., García-Orza, Estudillo, Calleja, & Rodríguez, 2017; Kallai & Tzelgov, 2012; Nuerk, Moeller, & Willmes, 2015). Imprecision in component magnitude estimation is not novel in that it is similar to the transformation of overall magnitudes by β, except at the level of components (e.g., for 25, both 20 and 5 might be transformed).

Each possibility fits well with one of our revised models. For our modified model, especially the multiplier-modified version, the account that fits most naturally is imprecision in place weighting. In this model, the representation of a numeral such as 25 (characterized earlier as $20 + \delta * 5$) can be rewritten more generally as $2^*\delta_t + 5^*\delta_o$, where $\delta_t = 10$ but $\delta_o < 1$. This model is less well characterized as imprecision in component magnitude estimation in that the bias extends only to the rightward digit rather than, as one might expect, to all place values. For the expanded model, the more straightforward account is imprecision in component magnitude estimation, where each component magnitude is raised to δ (e.g., $20^\delta + 5^\delta$).

4.3. Further testing and extending the models

The current version of the modified cyclical power model is specific to the 0–100 number range in that inputs to the model are tens and ones place values. One question is that of how this particular model might be generalized, especially to the commonly used 0–1000 range. It is likely that both the tens and ones place values are underweighted given that the left digit effect could not be explained by the ones digit alone (the difference score is much larger than in the present study; see, e.g., Kayton et al., 2022). A more general version of the multiplier-modified model might be written as: $d_1{}^*u_1 + d_2{}^*u_2{}^*\delta + d_3{}^*u_3{}^*\delta\ldots$ where d is the digit in a particular position (d_1 being the leftmost digit), u is the number of place units for that position, and δ is the weighting parameter. Although all rightward digits are weighted the same amount in the example, it is an empirical question as to whether the weights should be the same.

The current version of the expanded model, in contrast, already readily generalizes to numerals of any length (each place value is simply raised to δ). However, we have not addressed many questions that remain about this model that might be desirable to test in the future. For example, we kept the magnitude estimation bias δ the same for all the digits, but used a different parameter for the overall magnitude β. Given similarity of estimates of δ and β (e.g., 1.05 vs. 1.15 respectively for the one-cycle version of the extended model), one might ask if a single parameter might instead be used with little change in fit. We also made the assumption that the curve for the two-cycle model crosses at the midpoint of the line given that the midpoint is represented spatially rather than as a presented numeral, but this could be tested as well. For both models, it will be valuable to also collect processing tracing data such as eye movements (e.g., Sullivan et al., 2011) and finger tracking (e.g., Dotan & Dehaene, 2013) to better understand when and how the effect might emerge over the time course of responding.

Given the better fit of the expanded model over the multiplier-modified model, one might wonder why we continue to advance both models. A primary reason is that the expanded model was largely motivated by the asymmetrical S-shaped curve (where underplacement is larger than overplacement). However, this pattern has also been found using similar spatial tasks (Barth, Lesser, Taggart, & Slusser, 2015; Zax et al., 2019; see also Crawford & Duffy, 2010) where one is shown an unlabeled line with a hashmark and must reproduce the hashmark's location on a new line. That there are no numerals in this task suggests that the asymmetry in the curve may have a single source unrelated to the translation of numerals to magnitudes. There are many ways to capture asymmetry in a model besides the approach taken in our expanded model, perhaps most simply by using a different estimate of β for each part of the line. It thus remains possible that the multiplier-modified model, revised to include a different means of producing the asymmetry seen in the data, has the potential to extend to explain more phenomena than the expanded model. This is something that will be important to consider further in future work.

While the focus of the present work is number line estimation, left digit effects are also seen in complex judgment contexts, in domains ranging from consumer behavior to medical treatment decisions (see Patalano et al., 2022, for review). A study particularly relevant to the present work demonstrated that while the relationship between the set market value of a used car and its odometer reading is generally linear, large discontinuities exist at left digit boundaries (Lacetera, Pope, & Sydnor, 2012). The compression towards the lower boundary, strikingly similar to the pattern here, was well-modeled using an attentional weight to underweight rightward digits. The common finding raises many interesting questions about the relationship between that pattern and the processes that gave rise to it and the ones seen in the number line task.

In sum, rather than being an isolated phenomenon driven by the placements of a few numerals, the left digit effect in number line estimation emerges from a systematic leftward compression of placements both across the line as a whole and within each left digit range. Research on the left digit effect contributes to understanding number line estimation behavior, as well as to the process of conversion of symbols to magnitudes and to our understanding of behavior in the many judgment contexts in which left digit effects emerge.

CRediT authorship contribution statement

Andrea L. Patalano: Conceptualization, Methodology, Formal analysis, Writing – original draft, Funding acquisition. **Kelsey Kayton:** Software, Investigation, Formal analysis, Visualization, Writing – original draft. **Hilary Barth:** Conceptualization, Writing – original draft, Funding acquisition.

Acknowledgements

This work was supported by NSF DRL-1920445 and benefited from NSF DRL-1561214, both to HB and ALP. The authors declare that they have no competing interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2022.105257.

References

Barth, H., Lesser, E., Taggart, J., & Slusser, E. (2015). Spatial estimation: A non-Bayesian alternative. *Developmental Science, 18*(5), 853–862. https://doi.org/10.1111/desc.12264

Barth, H., & 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

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*(4), 1016–1031. https://doi.org/10.1111/j.1467-8624.2008.01173.x

Brez, C. C., Miller, A. D., & Ramirez, E. M. (2016). Numerical estimation in children for both positive and negative numbers. *Journal of Cognition and Development, 17*(2), 341–358. https://doi.org/10.1080/15248372.2015.1033525

Brysbaert, M. (1995). Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General, 124*(4), 434–452. https://doi.org/10.1037/0096-3445.124.4.434

Cohen, D. J., & Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychonomic Bulletin & Review, 18*(2), 331–338. https://doi.org/10.3758/s13423-011-0059-z

Cohen, D. J., Blanc-Goldhammer, D., & Quinlan, P. T. (2018). A mathematical model of how people solve most variants of the number-line task. *Cognitive Science, 42*(8), 2621–2647. https://doi.org/10.1111/cogs.12698

Crawford, L. E., & Duffy, S. (2010). Sequence effects in estimating spatial location. *Psychonomic Bulletin & Review, 17*(5), 725–730. https://doi.org/10.3758/PBR.17.5.725

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 626–641. https://doi.org/10.1037/0096-1523.16.3.626

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. https://doi.org/10.1126/science.1156540

Dotan, D., & Dehaene, S. (2013). How do we convert a number into a finger trajectory? *Cognition, 129*(3), 512–529. https://doi.org/10.1016/j.cognition.2013.07.007

Dotan, D., & Dehaene, S. (2020). Parallel and serial processes in number-to-quantity conversion. *Cognition, 204*, Article 104387. https://doi.org/10.1016/j.cognition.2020.104387

Ganor-Stern, D., Pinhas, M., & Tzelgov, J. (2009). Short article: Comparing two-digit numbers: The importance of being presented together. *Quarterly Journal of Experimental Psychology, 62*(3), 444–452. https://doi.org/10.1080/17470210802391631

García-Orza, J., Estudillo, A. J., Calleja, M., & Rodríguez, J. M. (2017). Is place-value processing in four-digit numbers fully automatic? Yes, but not always. *Psychonomic Bulletin & Review, 24*(6), 1906–1914. https://doi.org/10.3758/s13423-017-1231-x

Hamdan, N., & Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: Evidence from a fraction intervention. *Developmental Psychology, 53*(3), 587–596. https://doi.org/10.1037/dev0000252

Henninger, F., Shevchenko, Y., Mertens, U., Kieslich, P. J., & Hilbig, B. E. (2019). Lab.Js: A free, open, online study builder. *PsyArXiv*. https://doi.org/10.31234/osf.io/fqr49

Hollands, J. G., & Dyre, B. P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review, 107*(3), 500–524. https://doi.org/10.1037/0033-295X.107.3.500

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Kallai, A. Y., & Tzelgov, J. (2012). The place-value of a digit in multi-digit numbers is processed automatically. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(5), 1221–1233. https://doi.org/10.1037/a0027635

Kayton, K., Williams, K., Stenbaek, C., Gwiazda, G., Bondhus, C., Green, J., … Patalano, A. L. (2022). Summary accuracy feedback in number line estimation does not reduce the left digit effect. *Memory & Cognition*. https://doi.org/10.3758/s13421-022-01278-2. Advance online publication.

Lacetera, N., Pope, D. G., & Sydnor, J. R. (2012). Heuristic thinking and limited attention in the car market. *American Economic Review, 102*(5), 2206–2236. https://doi.org/10.1257/aer.102.5.2206

Lai, M., Zax, A., & Barth, H. (2018). Digit identity influences numerical estimation in children and adults. *Developmental Science, 21*(5), Article e12657. https://doi.org/10.1111/desc.12657

Li, Q., Joo, S. J., Yeatman, J. D., & Reinecke, K. (2020). Controlling for participants' viewing distance in large-scale, psychophysical online experiments using a virtual chinrest. *Scientific Reports, 10*(1), 1–11. https://doi.org/10.1038/s41598-019-57204-1

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*(1–2), 107–157. https://doi.org/10.1016/0010-0277(92)90052-J

McCloskey, M., Sokol, S. M., & Goodman, R. A. (1986). Cognitive processes in verbal-number production: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: General, 115*(4), 307. https://doi.org/10.1037/0096-3445.115.4.307

Moeller, K., Huber, S., Nuerk, H. C., & Willmes, K. (2011). Two-digit number processing: Holistic, decomposed or hybrid? A computational modelling approach. *Psychological Research, 75*(4), 290–306. https://doi.org/10.1007/s00426-010-0307-2

Nuerk, H. C., Moeller, K., & Willmes, K. (2015). Multi-digit number processing: Overview, conceptual clarifications, and language influences. In R. C. Kadosh, & A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 106–139). Oxford University Press.

Nuerk, H. C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition, 82*(1), B25–33. https://doi.org/10.1016/s0010-0277(01)00142-1

Nuerk, H. C., Weger, U., & Willmes, K. (2004). On the perceptual generality of the unit-decade compatibility effect. *Experimental Psychology, 51*(1), 72–79. https://doi.org/10.1027/1618-3169.51.1.72

Nuerk, H. C., & Willmes, K. (2005). On the magnitude representations of two-digit numbers. *Psychology Science, 47*(1), 52–72.

Patalano, A. L., Saltiel, J. R., Machlin, L., & Barth, H. (2015). The role of numeracy and approximate number system acuity in predicting value and probability distortion. *Psychonomic Bulletin & Review, 22*(6), 1820–1829. https://doi.org/10.3758/s13423-015-0849-9

Patalano, A. L., Williams, K., Weeks, G., Kayton, K., & Barth, H. (2022). The left digit effect in a complex judgment task: Evaluating hypothetical college applicants. *Journal of Behavioral Decision Making, 35*(1), Article e2247. https://doi.org/10.1002/bdm.2247

Patalano, A. L., Zax, A., Williams, K., Mathias, L., Cordes, S., & Barth, H. (2020). Intuitive symbolic magnitude judgments and decision making under risk in adults. *Cognitive Psychology, 118*, Article 101273. https://doi.org/10.1016/j.cogpsych.2020.101273

Peeters, D., Sekeris, E., Verschaffel, L., & Luwel, K. (2017). Evaluating the effect of labeled benchmarks on children's number line estimation performance and strategy use. *Frontiers in Psychology, 8*, 1082. https://doi.org/10.3389/fpsyg.2017.01082

Peters, E., & Bjalkebring, P. (2015). Multiple numeric competencies: When a number is not just a number. *Journal of Personality and Social Psychology, 108*(5), 802–822. https://doi.org/10.1037/pspp0000019

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition, 72*(2), 191–201. https://doi.org/10.1016/S0010-0277(99)00048-7

Savelkouls, S., Williams, K., & Barth, H. (2020). Linguistic inversion and numerical estimation. *Journal of Numerical Cognition, 6*(3), 263–274. https://doi.org/10.5964/jnc.v6i3.273

Schley, D. R., & Peters, E. (2014). Assessing "economic value" symbolic-number mappings predict risky and riskless valuations. *Psychological Science, 25*(3), 753–761. https://doi.org/10.1177/0956797613515485

Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in grades 5 and 6. *Journal of Educational Psychology, 101*(2), 359–372. https://doi.org/10.1037/a0013840

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

Shevchenko, Y. (2022). Open lab: A web application for running and sharing online experiments. *Behavior Research Methods*, 1–8.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*(3), 237–250. https://doi.org/10.1111/1467-9280.02438

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—But not circular ones—Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology, 101*(3), 545–560. https://doi.org/10.1037/a0014239

Slusser, E., & Barth, H. (2017). Intuitive proportion judgment in number-line estimation: Converging evidence from multiple tasks. *Journal of Experimental Child Psychology, 162*, 181–198. https://doi.org/10.1016/j.jecp.2017.04.010

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

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance, 16*(4), 683–692. https://doi.org/10.1037/0096-1523.16.4.683

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64*(3), 153–181. https://doi.org/10.1037/h0046162

Sullivan, J. L., Juhasz, B. J., Slattery, T. J., & Barth, H. C. (2011). Adults' number-line estimation strategies: Evidence from eye movements. *Psychonomic Bulletin & Review, 18*(3), 557–563. https://doi.org/10.3758/s13423-011-0081-1

Thomas, M., & Morwitz, V. (2005). Penny wise and pound foolish: The left-digit effect in price cognition. *Journal of Consumer Research, 32*(1), 54–64. https://doi.org/10.1086/429600

Tosto, M. G., Petrill, S. A., Malykh, S., Malki, K., Haworth, C., Mazzocco, M. M., & Kovas, Y. (2017). Number sense and mathematics: Which, when and how? *Developmental Psychology, 53*(10), 1924–1939. https://doi.org/10.1037/dev0000331

Vaidya, L., Eisen, S., Zax, A., Williams, K., Patalano, A. L., & Barth, H. (2022). *Left digit effects in number line estimation are robust across task format [unpublished manuscript]*. Department of Psychology, Wesleyan University.

Verguts, T., & De Moor, W. (2005). Two-digit comparison: Decomposed, holistic, or hybrid? *Experimental Psychology, 52*(3), 195–200. https://doi.org/10.1027/1618-3169.52.3.195

Williams, K., Bradley, K., Xing, C., Barth, H., & Patalano, A. L. (2022). *Potential moderators of the left digit effect in numerical estimation [unpublished manuscript]*. Department of Psychology, Wesleyan University.

Williams, K., Paul, J., Zax, A., Barth, H., & Patalano, A. L. (2020). Number line estimation and standardized test performance: The left digit effect does not predict SAT math score. *Brain & Behavior, 10*(12), e01877. https://doi.org/10.1002/brb3.1877
Williams, K., Zax, A., Patalano, A. L., & Barth, H. (2022). Left digit effects in numerical estimation across development. *Journal of Cognition and Development, 23*(2), 188–209. https://doi.org/10.1080/15248372.2021.1984243
Xing, C., Zax, A., George, E., Taggart, J., Bass, I., & Barth, H. (2021). Numerical estimation strategies are correlated with math ability in school-aged children. *Cognitive Development, 60*. https://doi.org/10.1016/j.cogdev.2021.101089
Yuan, L., Prather, R. W., Mix, K. S., & Smith, L. B. (2020). Number representation drive number line estimates. *Child Development, 91*(4), e952–e957. https://doi.org/10.1111/cdev.13333
Zax, A., Slusser, E. B., & Barth, H. (2019). Spontaneous partitioning and proportion estimation in children's numerical judgments. *Journal of Experimental Child Psychology, 185*, 71–94. https://doi.org/10.1016/j.jecp.2019.04.004
Zhang, H., & Maloney, L. Y. (2012). Ubiquitous log odds: A common representation of probability and frequency distortion in perception, action, and cognition. *Frontiers in Neuroscience, 6*, 1. https://doi.org/10.3389/fnins.2012.00001
Zhu, M., Cai, D., & Leung, A. W. S. (2017). Number line estimation predicts mathematical skills: Difference in grades 2 and 4. *Frontiers in Psychology, 8*, 1576. https://doi.org/10.3389/fpsyg.2017.01576