



Left Digit Effects in Numerical Estimation across Development

Katherine Williams, Alexandra Zax, Andrea L. Patalano & Hilary Barth

To cite this article: Katherine Williams, Alexandra Zax, Andrea L. Patalano & Hilary Barth (2022) Left Digit Effects in Numerical Estimation across Development, Journal of Cognition and Development, 23:2, 188-209, DOI: [10.1080/15248372.2021.1984243](https://doi.org/10.1080/15248372.2021.1984243)

To link to this article: <https://doi.org/10.1080/15248372.2021.1984243>



Published online: 25 Oct 2021.



Submit your article to this journal [↗](#)



Article views: 85



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)



Left Digit Effects in Numerical Estimation across Development



Katherine Williams, Alexandra Zax, Andrea L. Patalano, and Hilary Barth

Wesleyan University

ABSTRACT

Number line estimation (NLE) tasks are widely used to investigate numerical cognition, learning, and development, and as an instructional tool. Interpretation of these tasks generally involves an implicit expectation that responses are driven by the overall magnitudes of target numerals, in the sense that the particular digits conveying those magnitudes are unimportant. However, recent evidence shows that numbers with similar magnitudes but different leftmost digits are estimated very differently. For example, “798” is placed systematically much too far to the left of “801” in a 0–1000 NLE task by children aged 7–11 and adults. Here we ask whether this left digit effect generalizes to two-digit numerals in a 0–100 NLE task and whether it emerges in younger children. Children aged 5–8 (Study 1, $N = 73$), adults (Study 2, $N = 44$), and children aged 9–11 (Study 3, $N = 27$) completed a standard 0–100 NLE task on a touchscreen tablet. We observed left digit effects for two-digit numerals in children aged 8–11 and adults, with large effect sizes, demonstrating that these effects generalize to smaller numerical ranges. Left digit effects were not apparent in 5- to 7-year-olds, suggesting that these effects do not emerge at younger ages for smaller, more familiar numerical ranges. We discuss developmental emergence of left digit effects in number line estimation and implications within and beyond the field of cognitive development.

Learning how we acquire, process, and use the meanings of numbers throughout life is a key goal of research in cognitive and developmental psychology. One widely used paradigm in this area is number line estimation (NLE). In the most common form of NLE tasks, participants identify the appropriate location of a number on a bounded response line (e.g., with endpoints labeled 0 and 1000). NLE tasks are commonly used to investigate the development of numerical competence and have been linked to formal math skills in children (Ramani & Siegler, 2008; Schneider et al., 2018; Siegler, 2016; Xing et al., 2021). For example, greater accuracy on a NLE task in kindergarten predicts greater performance in exact calculation skills (Xenidou-Dervou, van der Schoot, & van Lieshout, 2015). Number lines also play a causal role in promoting children’s fraction magnitude concepts, and teaching young children to represent fractions on a number line leads to greater success with fraction magnitude comparisons (Fazio, Kennedy, Siegler, & De Smedt, 2016; Gunderson, Hamdan, Hildebrand, &

CONTACT Hilary Barth  hbarth@wesleyan.edu  Department of Psychology, Wesleyan University, 207 High Street, Middletown, CT 06459

Alexandra Zax is now in the Clinical Child Psychology Program at the University of Kansas. Katherine Williams is now in the Developmental Psychology Program at the University of Texas at Austin.

© 2021 Taylor & Francis

Bartek, 2019; see also Hamdan & Gunderson, 2017; Hansen et al., 2015). A recent meta-analysis underscores this link between NLE performance and math competence (Schneider et al., 2018).

Given the widespread use of NLE tasks as diagnostic and instructional tools, and to better understand the link between task performance and math competence, it is important to understand the cognitive processes underlying performance. The present work focuses on a question that explorations of NLE performance have not addressed: do target numerals with equivalent magnitudes generally lead to equivalent estimates, in the sense that the particular digits comprising those numerals are irrelevant? The influence of individual digits on number-related judgments has been considered in some areas of numerical cognition (e.g., Moeller, Nuerk, & Willmes, 2009; Nuerk, Kaufmann, Zoppoth, & Willmes, 2004; Nuerk, Moeller, & Willmes, 2015). However, the role of individual digits comprising multi-digit numbers has generally gone unaddressed in the context of number line estimation.

Both children's and adults' number line estimates can be strongly influenced by the particular leftmost digits of target numerals: numbers with very similar magnitudes but different leftmost digits are estimated differently at least in some contexts. This evidence of a "left digit effect" in NLE comes from studies (Lai, Zax, & Barth, 2018) which compared estimates for numbers between 0 and 1000 with similar magnitudes but different leftmost digits (e.g., 598 and 601). Critically, because these numbers have nearly indistinguishable magnitudes, they should be placed in approximately the same location on the line (e.g., 598 and 601 are three units apart, and three units should be nearly indistinguishable on a 0–1000 scale). School-age children (ages 7–11) and adults reliably placed numbers with similar magnitudes and different leftmost digits too far apart, and they placed the smaller number systematically to the left of the larger number on average. (Estimates of numbers with similar magnitudes and the same leftmost digits, e.g., 848 and 853, did not differ.) A strong left digit effect was observed both for children, who have more limited number experience, and for adults, who have extensive experience (Lai et al., 2018).

These findings also point to a particular developmental progression, at least for the 0–1000 range. The youngest children tested in the study by Lai et al., produced different estimates for numbers like 798 and 801 (different hundreds digits, essentially same magnitudes), but *not* for numbers like 701 and 798 (same hundreds digits, very different magnitudes). That is, these young children did not differentiate three-digit numbers with the same leftmost digits (Lai et al., 2018). The older children and the adults, unsurprisingly, did differentiate numbers like 701 and 798, but their estimates were heavily influenced by the leftmost digits of target numerals. This finding suggests that children first make symbolic 0–1000 number line placements based on hundreds digits, and only later do they make more fine-grained placements.

The evidence for left digit effects in other kinds of tasks and situations is widespread, emphasizing the role of digit-level information in judgments related to number and magnitude. In the domain of price comparisons, adults judge prices with different leftmost digits (e.g., \$5.00 vs. \$4.99) to be meaningfully different, while prices with the same leftmost digit (e.g., \$4.20 vs. \$4.19) are judged similarly (Thomas & Morwitz, 2005). These kinds of effects are not restricted to judgments about prices. Related left digit biases have also been reported in the behavior of prospective home buyers (Beracha & Seiler, 2015),

smokers attempting to quit (MacKillop et al., 2014), and other varied domains of judgment (see also Choi et al., 2019; MacKillop et al., 2014; Olenski, Zimmerman, Coussens, & Jena, 2020).

The extent of digit-level influence on number line performance is only beginning to be explored, with potentially important implications for designing and interpreting symbolic number line estimation studies. Because the numerical magnitudes of target numerals, not their component digits, are commonly the focal consideration when designing studies, target numerals are not chosen with left digit effects in mind, nor do analyses consider these effects. An assumption that numerals with similar magnitudes but different leftmost digits are interchangeable for practical purposes appears to underlie most design decisions and analytical approaches in research incorporating number line estimation. To give one example, researchers often use “contour analyses” to examine accuracy and/or variability of estimates at regions of interest along the number line (such as the midpoint and endpoints). Ashcraft and Moore (2012) did so for a 0–100 number line (averaging the two observations for the target numerals closest to the midpoint, 48 and 52) and for a 0–1000 number line (for 489 and 526). Averaging these observations for the purpose of contour analysis is done under the assumption that placements for these targets are essentially equivalent, because they have similar magnitudes, despite their differing leftmost digits. This approach is not unusual (e.g., Jung, Roesch, Klein, Dackermann, Heller, & Moeller, 2020; Peeters, Degrande, Ebersbach, Verschaffel, & Luwel, 2016; Peeters, Verschaffel, & Luwel, 2017). In our view these examples are not exceptional, but rather are representative of a broader pattern in which most prior work involves similar underlying assumptions. Also, target numeral sampling for studies involving number lines is often limited, with children producing estimates for as few as 10 target numerals in occasional studies (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010) and more commonly approximately 20 target numerals in others (Booth & Siegler, 2006; Laski & Siegler, 2007; Opfer & Siegler, 2007). Where left digit effects are present, they could potentially have an outsized influence on performance and data interpretation.

In the present work, we address outstanding questions about left digit effects in number line estimation. There is clear evidence for left digit effects in 0–1000 NLE tasks for adults and children aged seven through eleven. We do not yet know whether left digit effects also arise in younger children or whether they arise reliably in different numerical ranges. In Study 1, we address questions of generalization to younger ages and smaller numerical ranges by giving a 0–100 NLE task to children aged five through eight years. In Study 2, we address generalization to smaller numerical ranges in adults, asking whether left digit effects appear in adults’ 0–100 placements as they do in their 0–1000 placements, consistent with a prior report with a smaller sample than that used here (Savelkoul, Williams, & Barth, 2020). In Study 3, we address generalization to smaller numerical ranges in older children, giving the same 0–100 NLE task to children aged nine through eleven. We also consider these findings in combination with previous work to elucidate the developmental progression of left digit effects in the context of number line estimation.

Study 1

Children aged five through eight years completed a standard NLE task with target numerals between 0 and 100. Target numerals included nine pairs of numbers on either side of decade boundaries (e.g., targets “29” and “31” on either side of decade boundary 30). These

numbers functioned as “pairs” in the design and analyses only; they were not presented together during the task. If the particular digits comprising target numerals do not matter for participants’ placements, estimates for target numerals with different tens digits but very similar magnitudes will be placed no farther apart, on average, than is predicted by their absolute difference (e.g., 29 and 31 should be placed approximately two units apart). However, if particular digits do matter and left digit effects are in play, numbers below decade boundaries will be systematically placed too far to the left of those above decade boundaries (e.g., 29 should be placed more than two units to the left of 31).

Method

Participants

Ninety-one children, recruited from a local participant database, participated. Eighteen of these were excluded for not completing the study (four 5-year-olds) or producing estimates uncorrelated with target numerals (twelve 5-year-olds, two 6-year-olds). The final sample ($N = 73$) included 11 5-year-olds ($M_{age} = 5;5$, range = 5;1–5;11, 6 female, 5 male), 21 6-year-olds ($M_{age} = 6;4$, range = 6;0–6;11, 12 female, 9 male), 20 7-year-olds ($M_{age} = 7;4$, range = 7;0–7;11, 8 female, 12 male), and 21 8-year-olds ($M_{age} = 8;4$, range = 8;0–8;11, 12 female, 9 male).

Stimuli

Stimuli were created with the application EstimationLine (https://hume.ca/ix/estimation_line/) and presented on an Apple iPad Air. A black 16.65 cm horizontal line with endpoints labeled 0 and 100 was displayed approximately 5.8 cm above the bottom of the screen. A green rectangular “Go” icon and a red rectangular “Done” icon were located at the lower left hand and lower right corners, respectively (approximately 0.4 mm from the corners of the screen and 1.1 cm above the bottom of the screen). Experimental stimuli included nine pairs of numbers falling on either side of decade boundaries 10 through 90 (8/12, 18/22, 29/31, 38/42, 49/51, 59/61, 69/71, 78/82, 88/92). Target numerals were paired in the study design but were not presented to participants in pairs. The pairs were part of a set of 30 distinct target numerals, presented in a different pseudorandom order to each participant: 2, 4, 8, 12, 17, 18, 22, 23, 29, 31, 36, 38, 42, 47, 49, 51, 58, 59, 61, 62, 69, 71, 74, 78, 82, 86, 88, 92, 97, 99.

Design and procedure

Participants completed one block of the number line task: a practice trial (target numeral “27”) was followed by 30 test trials, with each target numeral presented once. Testing took place in a quiet room. The experimenter first said, “In this game, you will see a number line going from 0 to 100, and I will ask you to show me where you think some numbers should go on the number line. When you decide where the number goes, touch where you think it goes on the screen, and a red line will show up.” On each trial, the experimenter touched the Go icon and “Where is [N]” appeared approximately 5.8 cm from the left of the screen and 12 cm above the bottom of the screen. The experimenter asked, “If 0 goes here [point to left

endpoint] and 100 goes here [point to right endpoint], where would you put [N]?.” Participants touched the line to indicate their response, and the researcher pressed the Done icon to advance to the next trial. After a participant completed three or more trials successfully, on the remaining trials, the experimenter shortened the verbal prompt to, “Where would you put [N]?.” No feedback was provided.

Analyses

Percent absolute error (PAE) was calculated by dividing the difference between a participants’ estimate and the target numeral’s true location by the numerical range, then multiplying by 100 to express a percentage (Booth & Siegler, 2006; Lai et al., 2018; Slusser, Santiago, & Barth, 2013). We excluded individual estimates on a trial-by-trial basis if they differed from the group mean for that target numeral by more than two standard deviations (5-year-olds: $M = 4.24\%$ of trials, 6-year-olds: $M = 3.48\%$ of trials, 7-year-olds: $M = 3.17\%$ of trials, 8-year-olds: $M = 3.97\%$ of trials). Like Lai et al. (2018), we did not make specific predictions about age differences but did include age groups in our analyses.

To determine whether children produced different estimates for numbers around decade boundaries, we computed decade difference scores. In previous work on left digit effects in 0–1000 NLE, researchers supposed that numbers like 399 and 401 would be indistinguishable on a 0–1000 scale and placed in the same location (on average). Accordingly, they interpreted differences between estimates for these numbers (with 399 placed too far to the left of 401) as evidence of left digit effects (Lai et al., 2018). In the present study, we did not make a parallel assumption because numbers like 29 and 31 could possibly be distinguishable on a 0–100 scale: 29 might reliably be placed to the left of 31 because of highly accurate estimation performance, not because of left digit bias. To test for left digit effects in this 0–100 context, we controlled for the exact distance between numbers by subtracting the veridical difference between each pair of target numerals from the difference between their placements (e.g., after calculating “placement for 31 – placement for 29” we subtracted two from the resulting difference). Therefore, for each critical pair, we calculated (*placement for larger number – placement for smaller number*) – (*larger number – smaller number*). This yielded nine decade difference scores per individual (one for each pair: 8/12, 18/22, 29/31, 38/42, 49/51, 59/61, 69/71, 78/82, 88/92). We planned to exclude participants at this point if more than five pairs had missing data due to outlier removal, but no participants required exclusion. Finally, we averaged the scores to obtain a single mean decade difference score for each participant.¹ Decade difference scores significantly greater than zero reflect a left digit effect.

We also computed a separate type of difference score for pairs of numbers with the same leftmost digit but different magnitudes (e.g., 82 and 88). Following Lai et al. (2018), these “high-low” difference scores were used to determine whether participants differentiated numbers with differing magnitudes falling within the same decade. We calculated high-low difference scores by subtracting the estimate for the smaller target numeral in the pair from the estimate for the larger (e.g., placement for 88 – placement for 82). For each participant, we averaged the resulting differences for eight high-low pairs (12/18, 22/29, 31/38, 42/49,

¹Mean decade difference scores were calculated both with and without the pair 8/12, but we report the main analyses excluding this pair (because 8 does not have a tens digit).

51/59, 61/69, 71/78, 82/88) to obtain one mean high-low difference score per participant. Again, no participants required exclusion due to outlier removal. High-low difference scores that do not differ from zero (on average) would show that participants' placements did not distinguish these numbers. This would likely indicate that leftmost tens digits alone influenced estimates if a participant also showed a left-digit effect in decade difference scores. If, however, high-low scores are greater than zero, on average, this would suggest that estimates were not determined only by the leftmost tens place digit.

Results and discussion

Overall accuracy is reported for each age group and is generally consistent with previous reports: for 5-year-olds, PAE = 15.4%; for 6-year-olds, PAE = 11.3%; for 7-year-olds, PAE = 9.0%; for 8-year-olds, PAE = 6.4%. We first asked if young children's performance on the 0–100 NLE task exhibited a left digit effect. The question of interest here is whether the left digit effect is present or absent in each age group, and the corresponding tests were preplanned, thus we focus on discussing and depicting results without correction for multiple comparisons. We also provide information about outcomes with Bonferroni correction for multiple comparisons for interested readers. Independent single-sample *t*-tests (one-tailed) revealed that decade difference scores (age 5: $M = -1.13$, $SD = 9.96$; age 6: $M = 1.25$, $SD = 3.96$; age 7: $M = 0.42$, $SD = 5.54$; age 8: $M = 1.75$, $SD = 3.00$) were significantly greater than zero for 8-year-olds ($t(20) = 2.67$, $p = 0.015$, $d = 0.58$), but not for any other age group (all t s < 1.5 , all p s > 0.10).² Only 8-year-olds' performance therefore showed evidence of a left digit effect in this analysis.³ Left digit effects in the 0–100 NLE range were not observed for 5- to 7-year-olds.⁴ See Figure 1 for participants' mean decade difference scores (generated from individual participant data).

We next carried out exploratory analyses to identify any significant left digit effects that might emerge at the level of individual target numeral pairs. We conducted single sample *t*-tests (one-tailed) on decade difference scores for nine critical pairs. We note that these analyses are based on small numbers of trials, with only one difference score per participant. Prior to correction for multiple comparisons, many pairs yielded significant left digit effects (we provide these results and additional descriptive statistics in Table A1 of the Appendix). Following a stringent correction for multiple comparisons within each age group ($\alpha = 0.006$), a significant left digit effect emerged for target numerals around the decade boundary value 90 in 8-year-olds ($p < 0.001$), but no significant effects emerged for 5- to 7-year-olds. Individuals' mean decade difference scores for each pair are shown in Figure 2.

We also asked whether children in each age group reliably differentiated lower from higher numbers within a decade (whether high-low difference scores differed from zero). High-low difference scores (for age 5: $M = 3.90$, $SD = 10.04$; for age 6: $M = 2.12$, $SD = 4.06$;

²When we include 14 participants originally excluded for producing estimates uncorrelated with target numerals, the results remain the same (8-year-olds, $p = 0.015$; all other ages, p s > 0.10).

³Results remain the same if the first decade boundary pair (8/12) is included (8-year-olds, $p = 0.005$; all other ages, $p > 0.3$). After correcting for multiple comparisons ($\alpha = 0.012$), 8-year-olds do not show a left digit effect ($p = 0.014$) except when the pair 8/12 is included ($p = 0.005$).

⁴We did not predefine criteria for excluding individual participants as outliers, but one 7-year-old's difference score was 3.5 standard deviations below the age group average. Thus, we conducted an additional post hoc analysis for 7-year-olds with this participant excluded. Results remained the same, both without the pair 8/12 ($M = 1.11$, $SD = 2.79$, $p = 0.101$), and when it was included ($M = 1.44$, $SD = 3.22$; $t(18) = 1.95$, $p = 0.067$).

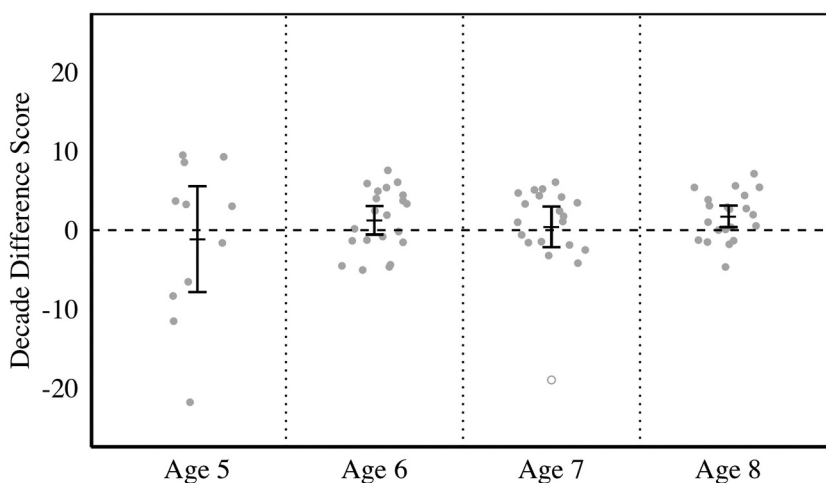


Figure 1. Individuals' mean decade differences scores for each age group. Difference scores reflect the difference in placements for pairs of numbers like 31 and 29 minus their absolute difference (e.g., 2 subtracted from the placement of 31 minus the placement of 29; see Analyses section for details). Black bars show the mean decade difference score for each age group, with 95% confidence intervals. The open circle in the Age 7 data represents the participant treated as an outlier in some analyses (this point is included in the group average; see Results).

for age 7: $M = 4.14$, $SD = 4.06$; for age 8: $M = 3.62$, $SD = 2.78$) were greater than zero for 8-year-olds ($t(20) = 5.98$, $p < 0.001$, $d = 1.30$) and for 7-year-olds ($t(19) = 4.57$, $p < 0.001$, $d = 1.08$), but not for 6-year-olds or 5-year-olds ($ts < 2.0$, $ps > 0.06$). That is, 5-year-olds and 6-year-olds produced indistinguishable estimates for pairs of numbers like 22 and 29 (and also showed no left digit effect). In contrast, 7-year-olds and 8-year-olds produced reliably different estimates for these numbers (with only 8-year-olds showing a left digit effect). This finding suggests two things. First, although 8-year-olds did exhibit left digit effects, their placements were not determined solely by leftmost digits; they also differentiated target numerals within decades. Second, 7-year-olds, but not 5- and 6-year-olds, could differentiate target numerals within decades, but none of these age groups exhibited left digit effects. Thus, in this dataset, the ability to distinguish target numerals within decades emerged in younger age groups than did left digit effects. We return to these points in the General discussion.

Consistent with previous findings, high-low difference scores were negatively correlated with decade difference scores ($r(71) = -0.87$, $p < 0.001$). This is perhaps unsurprising because a child using *only* tens digits would place 29 and 31 far apart but 22 and 29 close together. Although we did not make specific predictions about age differences because the study was not powered to detect age differences through a direct comparison of age groups ($N = 280$ would be needed), we conducted an ANOVA with age group as the independent variable and decade difference score as the dependent variable and found no effect of age: $F(3, 69) = 0.75$, $p = 0.529$. Neither decade difference scores ($p = 0.219$) nor high-low difference scores ($p = 0.713$) were reliably correlated with age in months.

Key findings from this study are as follows. Left digit effects do appear in children's 0–100 number line estimates: 8-year-olds (but not younger children) systematically placed target numerals with the same approximate magnitudes but different tens-place digits farther apart

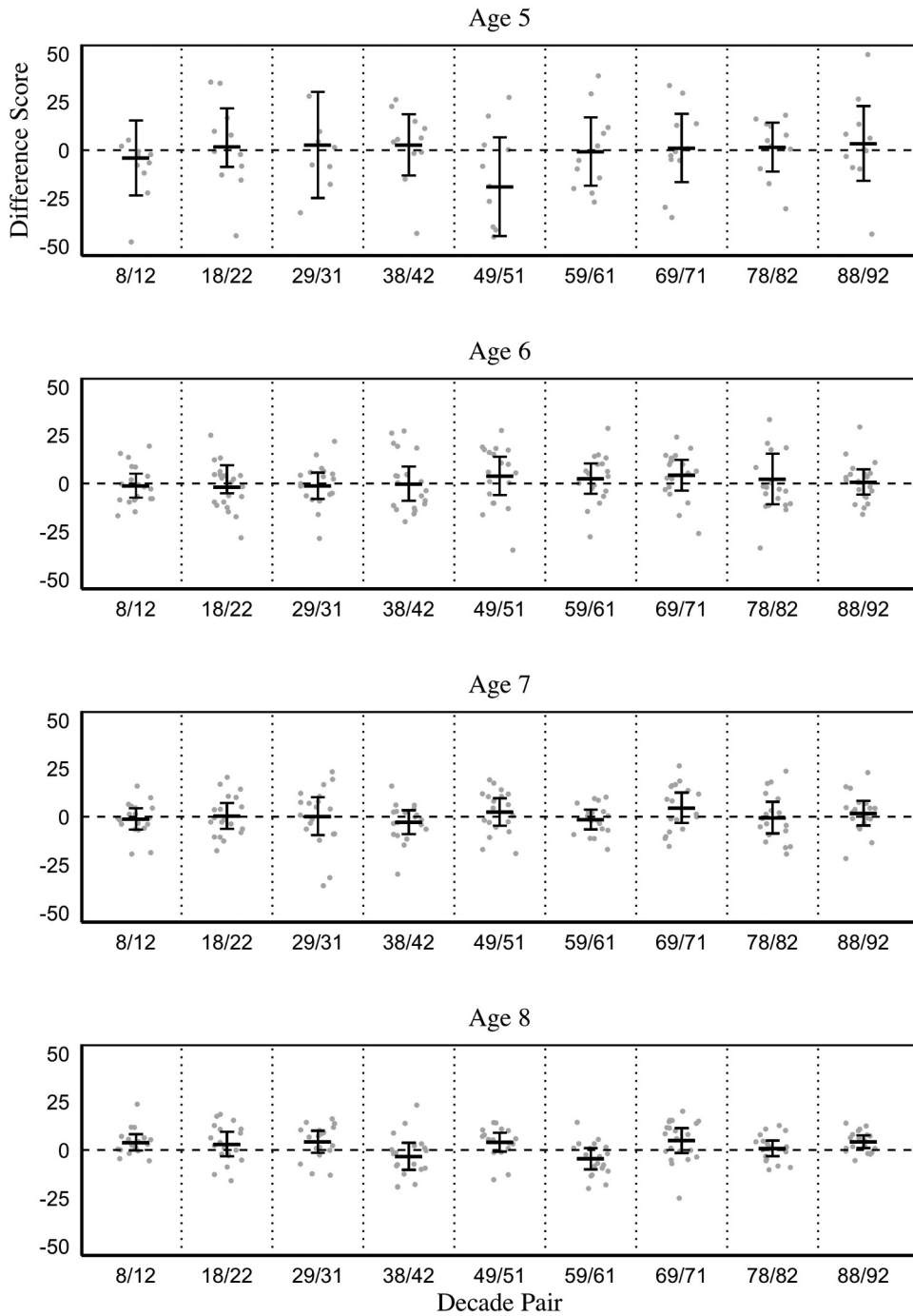


Figure 2. Individuals' mean decade difference score for each pair by age group. Difference scores greater than zero indicate a left digit effect. Black bars show the mean difference score for each decade pair and simultaneous 95% confidence intervals ($\alpha = 0.006$).

on the line than is predicted by their absolute difference. For example, they placed 29 more than two units to the left of 31. Children as young as seven show a large left digit effect in 0–1000 NLE with three-digit numbers (Lai et al., 2018), but in the present study, children younger than eight did not exhibit left digit effects in 0–100 NLE. However, 64% of 5-year-olds ($N = 7$), 81% of 6-year-olds ($N = 17$), and 60% of 7-year-olds ($N = 12$) had positive mean decade difference scores (as did 76% of 8-year-olds, $N = 16$) revealing that many younger children were placing the smaller numeral too far to the left of the larger numeral in our task.

These findings, taken together with previous reports, indicate that left digit effects in NLE are present at approximately age 8 across multiple numerical ranges, that they are not yet present at age 6 (at least in the 0–100 range), and that they are likely beginning to emerge around age 7 or 8 across multiple numerical ranges. Thus, there is evidence for a general left digit effect in number line estimation: the effect arises for two-digit target numerals in 0–100 NLE and three-digit target numerals in 0–1000 NLE, at least in children.

Study 2

Study 2 concerns the generalization of the left digit effect in NLE to a smaller numerical range in adults. Large left digit effects have been observed in two studies of adults' performance across different task conditions with larger numerical ranges (i.e., 0–1000, Lai et al., 2018). However, less is known about similar influences in smaller numerical ranges. We asked if tens-place digits overly influence adults' NLE placements in the highly familiar 0–100 range, as has been found in one previous report using a smaller sample (Savelkouls et al., 2020). To explore the left digit effect in the context of adults' performance, in Study 2, adults completed the same task used in Study 1, with additional blocks of trials.

Method

Participants

Forty-four undergraduates ($M_{age} = 19.66$, range = 18–23 years, 31 female, 13 male) participated in the study for course credit. Exclusion criteria were the same as in Study 1; no participants required exclusion for failing to complete the study or for producing estimates uncorrelated with target numerals.

Stimuli

Stimuli were identical to those used in Study 1.

Design and procedure

Participants completed eight blocks of the 0–100 number line task. Each block consisted of one practice trial (with the numeral 27) followed by 30 test trials. Target numerals were presented once per block in a different pseudorandom order within each block. Written and spoken instructions identical to those in Study 1 were given before each consecutive block.

There was no option to retry, and no feedback was provided. Between blocks, the following appeared on the screen: “Done. Please return the iPad to the Experimenter.” Adults completed the task at their own pace.

Analyses

Analyses were preregistered unless noted otherwise (<https://aspredicted.org/blind.php?x=67vy9e>). Percent absolute error (PAE) was calculated as in Study 1. In keeping with the exclusion criteria from Study 1, we excluded individual estimates on a trial-by-trial basis if they differed from the group mean for a given target numeral by more than two standard deviations ($M = 3.6\%$ of trials). We again calculated mean decade difference scores using eight critical pairs. (We planned to exclude participants from this analysis if more than five pairs were missing due to outlier removal; no participants required exclusion.) We report both individual mean decade difference score for each participant (averaged across all eight blocks of trials) and mean individual decade difference scores within each of block of trials. As in Study 1, mean decade difference scores greater than zero reflect a left digit effect. We also calculated one mean high-low difference score (e.g., the average estimated difference between numbers like 31 and 39) from individual scores for eight high-low pairs (12/18, 22/29, 31/38, 42/49, 51/59, 61/69, 71/78, 82/88; see Study 1 Analyses section for more details). We report one mean high-low difference score for each participant (averaged across eight blocks of trials).

Results and discussion

Overall accuracy aligns with previous reports of adult performance, with an average PAE of 3.6% (across eight blocks of trials). Independent single-sample t -tests (one-tailed) revealed that decade difference scores ($M = 0.74$, $SD = 1.28$) were greater than zero ($t(43) = 3.85$, $p < 0.001$, $d = 0.58$) indicating a strong left digit effect.⁵ Participants’ individual mean decade difference scores are shown in Figure 3.

We again carried out exploratory analyses for left digit effects at the level of individual target numeral pairs by conducting single sample t -tests (one-tailed) on the mean decade difference score (averaged across eight blocks) for nine critical pairs. Decade difference scores were significantly greater than zero (after correcting for multiple comparisons, $\alpha = 0.006$), for target numerals around decade boundaries 40, 70, and 90 ($ts > 2.00$, $ps < 0.001$, $ds = 0.71\text{--}0.89$).⁶ See Figure 4 for individuals’ mean decade difference scores (averaged across eight blocks of trials) for each pair. (We provide descriptive statistics in Table A2 of the Appendix, including reports of pairs that produced significant effects prior to correction for multiple comparisons.)

We also conducted exploratory block-level analyses. Decade difference scores were significantly greater than zero (corrected for multiple comparisons, $\alpha = 0.006$) in Block 1 ($M = 1.08$, $SD = 1.75$) and Block 3 ($M = 1.03$, $SD = 1.92$), all $ts > 3.5$, $ps < 0.001$, $ds = 0.38\text{--}0.62$. This was not the case in any other blocks (all $ts < 2$, $ps > 0.055$). The left

⁵Results remain the same when we include the pair 8/12: ($M = 0.65$, $SD = 1.17$, $t(43) = 3.66$, $p = 0.001$, $d = 0.85$).

⁶Though our original tests were one-tailed, there was one effect that appeared to be in the opposite of the expected direction: decade difference scores were significantly less than zero (after correcting for multiple comparisons, $\alpha = 0.006$) for target numerals around the decade boundary 80 ($t(43) = -4.20$, $p < 0.001$, $d = 0.63$).

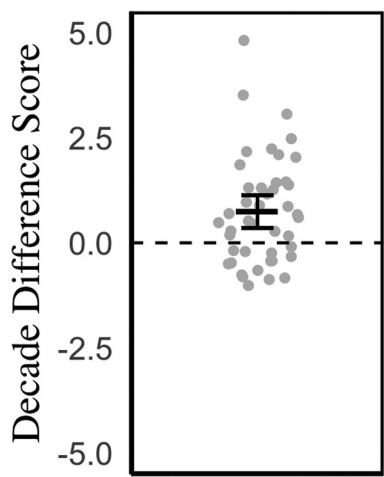


Figure 3. Individuals’ mean decade difference scores. Decade difference scores reflect the difference between placements for pairs of numbers like 22 and 18 minus the absolute difference between them (e.g., four subtracted from the placement of 22 minus the placement of 18; see Study 1 Analyses section for details). Black bars show the group mean decade difference score, with 95% confidence interval.

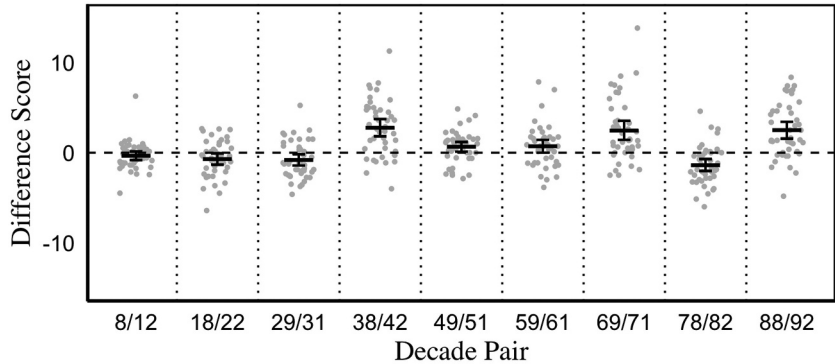


Figure 4. Individuals’ mean decade difference score for each pair. Difference scores greater than zero indicate a left digit effect (see Analyses section for more details). Black bars show the group mean difference score for each decade pair with simultaneous 95% confidence intervals (corrected for multiple comparisons, $\alpha = 0.006$).

digit effect was not uniform across blocks of trials with this stringent correction: the effect emerged for Block 1 and Block 3 but not in Block 2 or Blocks 4–8. One possibility is that the left digit effect diminished over time as adults gained experience and familiarity with the task and the target numerals (because the same 30 target numerals were estimated within each block, resulting in eight estimates of each target numeral throughout the study).

We next assessed participants' average high-low difference scores and found that they were significantly greater than zero, on average, ($M = 6.40$, $SD = 1.13$; $t(43) = 37.50$, $p < 0.001$, $d = 5.66$) indicating that adults reliably differentiated numbers like 22 and 29. Consistent with Study 1, there was a strong negative correlation between high-low difference scores and decade difference scores ($r(42) = -.93$, $p < 0.001$). These findings demonstrate that adults' estimates, unsurprisingly, are not driven solely by the leftmost digits of target numerals.

In Study 2, we asked whether left digit effects appear in adults' NLE placements in the very familiar 0–100 range, as they do in the 0–1000 range (Lai et al., 2018). Recall that if estimates are generated solely from the magnitudes of two-digit target numerals, numbers like 29 and 31 should be placed no more than two units apart. Like the 8-year-olds' performance in Study 1, adults' performance in Study 2 revealed a left digit effect: digits in the tens place heavily influenced adults' estimates (e.g., adults placed 29 more than two units to the left of 31). These data, consistent with prior work (Savelkouls et al., 2020), show that adults exhibit robust left digit effects in NLE, even for two-digit numerals in the highly familiar 0–100 range. As expected though, number line placements were not determined only by leftmost digits: adults successfully differentiated pairs of target numerals within decades (like 7- and 8-year-olds, but not 5- and 6-year-olds, in Study 1).

Study 3

Study 3 concerns the generalization of the left digit effect in smaller NLE ranges to older children, aged nine to 11 years. Though left digit effects have been observed in older children's performance across different task conditions with larger numerical ranges (i.e., 0–1000, Lai et al., 2018), less is known about whether similar influences on performance appear with smaller, more familiar numerical ranges. This was the focus of Study 3: children aged 9–11 completed the same 0–100 NLE task used in Study 1.

Method

Participants

We recruited child participants ($N = 27$) from a local participant database and a local science museum; one additional participant was excluded for not completing the study. Our original target sample size was $N = 60$ participants ($N = 20$ each of 9-, 10-, and 11-year-olds), but in-person data collection was interrupted by the COVID-19 pandemic. Because no further data will be available from our face-to-face iPad-based task in the near future, we include these data here as one sample of children aged 9–11 (deciding post hoc to conduct combined analyses on children aged 9–11, rather than separating the age groups as originally planned). The final sample included seven 9-year-olds ($M_{age} = 9;4$, range = 9;0–9;6, four female, three male), eight 10-year-olds ($M_{age} = 10;4$, range = 10;0–10;9, five female, three male), and 12 11-year-olds ($M_{age} = 11;5$, range = 11;1–11;9, seven female, five male).

Stimuli

Stimuli were identical to those used in Study 1.

Design and procedure

Design and Procedure were identical to those used in Study 1.

Analyses

Analyses were preregistered unless noted otherwise (<https://aspredicted.org/blind.php?x=67vy9e>). Percent absolute error (PAE) was calculated as in Study 1. In keeping with the exclusion criteria from Study 1, we excluded individual estimates on a trial-by-trial basis if they differed from the group mean for a given target numeral by more than two standard deviations ($M = 3.40\%$ of trials). We again calculated mean decade difference scores using eight critical pairs. (We planned to exclude participants from this analysis if more than five pairs were missing due to outlier removal; no participants required exclusion.) Recall that mean decade difference scores significantly greater than zero reflect a left digit effect. We also calculated one mean high-low difference score (e.g., the average estimated difference between numbers like 39 and 31) from individual scores for eight high-low pairs (see Study 1 Analyses section for more details).

Results and discussion

Overall accuracy is generally consistent with previous reports for this age group, PAE = 5.47%. Independent single-sample t -tests (one-tailed) revealed that decade difference scores ($M = 3.58$, $SD = 2.45$) were greater than zero ($t(26) = 7.58$, $p < 0.001$, $d = 1.45$) indicating a strong left digit effect.⁷ See Figure 5 for participants' individual mean decade difference scores.

As in Studies 1 and 2, we carried out exploratory analyses for left digit effects at the level of individual pairs by conducting single sample t -tests (one-tailed) on the mean decade difference score for nine pairs of target numerals. Decade difference scores were significantly greater than zero (after correcting for multiple comparisons, $\alpha = 0.006$) for target numerals around decade boundaries 50, 70, and 90 ($ts > 3.00$, $ps < 0.002$, $ds = 0.59$ – 0.87). No decade difference scores were significantly less than zero. See Figure 6 for individuals' mean decade difference scores for each pair. (We provide descriptive statistics in Table A3 of the Appendix, including reports of pairs that produced significant effects prior to correction for multiple comparisons.)

We next assessed participants' average high-low difference scores and found that they were greater than zero, on average ($M = 3.57$, $SD = 2.19$; $t(27) = 8.46$, $p < 0.001$, $d = 1.63$) indicating that children aged 9 to 11 years reliably differentiated numbers like 22 and 29. This finding is consistent with the conclusion that the leftmost digits of target numerals do not solely drive older children's estimates. Consistent with the findings of Study 1 and Study

⁷Results remain the same when the pair 8/12 is included in the analysis: ($t(26) = 7.59$, $p < 0.001$, $d = 1.46$).

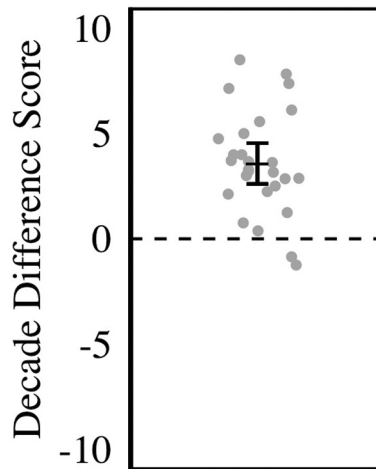


Figure 5. Individual mean decade difference scores. Decade difference scores reflect the difference between placements for pairs of numbers like 22 and 18 minus the absolute difference between them (e.g., four subtracted from the placement of 22 minus the placement of 18; see Study 1 Analyses section for details). Black bars show the group mean decade difference score, with 95% confidence interval.

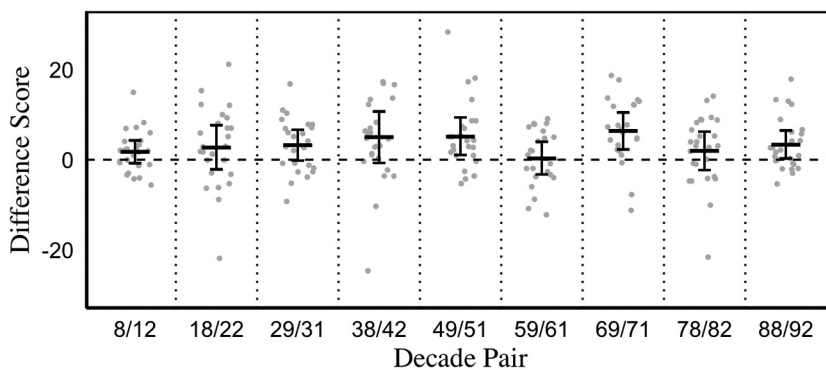


Figure 6. Individuals' mean decade difference score for each pair. Decade difference scores greater than zero indicate a left digit effect (see Analyses section for more details). Black bars show the group mean difference score for each pair with simultaneous 95% confidence intervals (corrected for multiple comparisons, $\alpha = 0.006$).

2, we found a strong negative correlation between high-low difference scores and decade difference scores: $r(25) = -0.842$, $p < 0.001$. Neither decade difference scores ($p = 0.584$) nor high-low difference scores ($p = 0.244$) were reliably correlated with age in months.

In Study 3, we asked if left digit effects appear in 9- to 11-year-olds' NLE placements in the 0–100 range, as they do in the 0–1000 range (Lai et al., 2018). Like the 8-year-olds in Study 1 and adults in Study 2, older children in Study 3 showed clear evidence of a left digit effect: digits in the tens place heavily influenced estimates (e.g., they placed 29 more than two units to the left of 31). Thus, 9- to 11-year-olds exhibit a robust left digit effect in their

number line estimation performance, even for two-digit numbers in the highly familiar 0–100 numerical range. Also, unsurprisingly, number line placements were not determined by leftmost digits alone: older children successfully differentiated pairs of target numerals within decades (like 7- and 8-year-olds in Study 1 and adults in Study 2).

Combined exploratory analyses: studies 1–3

The findings above provide evidence of a left digit effect in children's and adults' 0–100 NLE performance. Children aged 8–11 years (Study 1 and Study 3) and adults (Study 2) placed target numerals with similar magnitudes, but different leftmost digits, too far apart on a 0–100 number line, and they placed the smaller numeral systematically to the left of the larger numeral. Was this effect seen in most individuals, or are these findings driven by only a few individuals? To attempt to answer this question, we treated each participant as an individual experiment and asked what percentage of participants showed a reliable left digit effect at the individual level. We defined individual left digit effects in the following (post hoc) manner. Using individual participants' data, for each pair of target numerals falling around decade boundaries 20 through 90, we asked if the smaller number was placed to the left of the larger number and, if so, whether the estimated difference (e.g., how many units apart 29 and 31 were placed on the line) was greater than the absolute difference (e.g., 29 and 31 have an absolute difference of two units). If there is a left digit effect, the smaller numeral will be placed to the left of the larger numeral, *and* the estimated difference will be greater than the absolute difference far more often than is predicted by chance. Sign tests revealed no significant individual-level left digit effects when defined in this manner for 5- to 8-year-olds (Study 1) or 9- to 11-year-olds (Study 3). Only six adults (Study 2) showed individual left digit effects defined in this way. However, given the small number of trials included in this analysis, especially for children, these results should be interpreted with caution: a different study design would be better suited to testing for individual-level left digit effects.

General discussion

The current work investigates several open questions about left digit effects in children's and adults' number line estimation (NLE) performance. In previous studies, adults and children aged seven to eleven showed large and robust left digit effects for three-digit numbers in a 0–1000 symbolic NLE task. Their estimates were heavily influenced by the specific digits in target numerals that instantiated indistinguishable numerical magnitudes, producing dramatically different estimates for numbers like 699 and 701, for example. In the present studies, we asked whether the left digit effect exists more broadly. Specifically, we asked whether children ranging from 5 to 11 years of age and adults demonstrate similar left digit effects for two-digit numbers in a standard 0–100 NLE task.

We first asked if the left digit effect in NLE previously observed for three-digit numbers in the 0–1000 range also exists for two-digit numbers in the 0–100 range. Across three studies, we tested younger children (aged 5–8), older children (aged 9–11), and adults on a standard 0–100 NLE task. If left digit effects appear in 0–100 NLE placements, numbers like 29 and 31 with different leftmost digits will be placed more than two units apart (and 29 will be placed systematically to the left of 31). Second, we asked if the left digit effect in NLE might emerge before the age of seven. Children younger than seven were not tested on the

0–1000 range (Lai et al., 2018) because they have little familiarity with three-digit numbers, but left digit effects might emerge in younger children when tested on two-digit numbers in the more familiar 0–100 numerical range.

In Study 1, 8-year-olds, but not younger children, made 0–100 NLE placements that were influenced by the leftmost digits of target numerals. Eight-year-olds placed numbers like 29 too far to the left of numbers like 31, showing that the leftmost digits of target numerals influenced their estimates. The same was true of adults in Study 2 and older children, aged 9–11, in Study 3.⁸ These findings provide clear evidence that left digit effects appear in number line tasks with two-digit numbers in the 0–100 range, just as they do for three-digit numbers in the 0–1000 range (Lai et al., 2018). Children younger than eight (Study 1) did not make placements that were heavily influenced by leftmost digits in this task. Prior studies report left digit bias in the 0–1000 range at age 7 and age 8; the present work provides no evidence that left digit effects arise at younger ages even when children are tested with a smaller, more familiar numerical range.

These findings differ somewhat from the findings in the 0–1000 NLE range. In particular, Lai et al., identified a particular developmental progression in which the youngest children (age 7 in one study; age 8 in a second study) appeared to make 0–1000 NLE placements based *only* on the leftmost digit in three-digit numbers: they showed large left digit effects but did not differentiate numbers that differed considerably in magnitude and shared a leftmost digit (like 701 and 798) at all. Older children still showed large left digit effects, as did adults, but by age 9, children could successfully differentiate three-digit numbers that differed considerably in magnitude and shared a leftmost digit. We did not observe a parallel developmental progression for the 0–100 NLE here. In the present studies, 5- and 6-year-olds showed neither a left digit effect in 0–100 NLE nor an ability to reliably distinguish target numerals within a decade that shared a leftmost digit (e.g., 22 and 29). Eight-year-olds (but not 7-year-olds) showed left digit effects, and their placements were not solely based on leftmost digits: both 7- and 8-year-olds placed numbers like 22 and 29 in different locations on the line, showing that they successfully distinguished within-decade target numerals with the same leftmost digit (e.g., 22/29). Nine- to 11-year-olds and adults also showed robust left digit effects and successfully distinguished within-decade target numerals. Thus, performance on the 0–100 NLE task did not show the same developmental progression as performance on the 0–1000 NLE task: for three-digit numbers, left digit (hundreds) effects precede within-hundreds differentiation, but for two-digit numbers, left digit (decades) effects do not precede within-decade differentiation. We consider the description of these developmental progressions somewhat preliminary because, to date, they are based on a small number of studies.

Left digit effects in NLE appear to emerge at approximately age 7 or 8 both for two- and three-digit target numerals across multiple numerical ranges (at least 0–100 and 0–1000) and they persist into adulthood with large effect sizes, even for highly familiar numbers (two-digit target numerals in the 0–100 range). Of course, grade level and exposure to specific school mathematical curricula may contribute to the emergence of left digit effects. Future work could delineate the specific concepts and skills that support this emergence.

⁸Adopting a more stringent criterion in Study 1 (i.e., correcting for multiple comparisons) would lead to a slightly different conclusion: the left digit effect in our data set is present in children aged 9–11, but not before age 9. However, this would be a highly conservative correction given the predictions of the study.

For example, place value concepts can be difficult for children to master until their late elementary years (Fuson, 1990). One possibility is that the emergence of left digit effects depends on a degree of facility with the place-value structure of Arabic numerals that is not generally present in children before age 7. Critically, place-value understanding assumes successful mastery of the value of individual digits in multidigit numbers based on their position (from left to right) and also that each position has a value ten times larger than the position to its right. However, young children have partial knowledge of multidigit numbers and their place-value structure before receiving explicit instruction. For example, preschoolers know that larger single-digit numbers (e.g., 7, 8, 9) are larger quantities than small single-digit numbers (e.g., 1, 2, 3; Sarnecka & Gelman, 2004) and that number of digits is related to magnitude (e.g., $101 > 99$; Mix, Prather, Smith, & Stockton, 2014). Preschoolers also know that place-value matters in multidigit numbers (i.e., in English, the first digit spoken corresponds to the leftmost digit, and the magnitude of each place decreases from left to right) and can use their knowledge of the temporal-spatial patterns of spoken number words to produce written number forms (i.e., the left-to-right place-value positions of 100s, 10s, and 1s; see Byrge, Smith, & Mix, 2014). If young children first master an understanding of the units position, this might explain why 7-year-olds could differentiate numbers like 22 and 29 but did not show a left digit effect for numbers like 29 and 31. Preschoolers' ability to map number words to written digits is greater than their ability to match number names to quantities for both 2- and 3-digit numbers, suggesting that the latter ability might develop at a later age, but there are large individual differences in what children know (Yuan, Prather, Mix, & Smith, 2019).

The current work provides clear evidence for left digit effects in two-digit numbers in children's and adults' 0–100 NLE performance. One limitation of these findings is that they cannot characterize potential individual differences in left digit effects or provide clues as to their source. Future work might explore the relationship between the left digit effect and individual differences in formal measures of math or verbal performance. In a recent study, the left digit effect was observed in both speeded and unspeeded versions of a 0–1000 NLE task, and the left digit effect was not predicted by SAT math scores but was related to SAT verbal scores in the speeded task only (Williams, Paul, Zax, Barth, & Patalano, 2020). Additional research, however, is needed to better understand the relationship between the left digit effect and STEM-related skills and abilities and whether the specific skills closely associated with the left digit effect in NLE vary by numerical range and developmental age. In ongoing work we are investigating whether the left digit effect is influenced by practice and/or performance feedback. Our post hoc analyses of adults' performance over eight blocks of trials suggested that the left digit effect in Study 2 may have diminished over time. Previous work has found improvements in overall NLE accuracy and reduced overall estimation bias after brief feedback on a small selection of target numerals in children (Barth et al., 2016; Opfer & Siegler, 2007) and adults (Eyler, Cordes, Szymanski, & Fraenkel, 2018). It remains an open question whether left digit effects can be reduced or eliminated with brief accuracy feedback or if a more intricate form of feedback or intervention is required.

How could theoretical accounts of multidigit number processing address left digit effects? Several theories have been put forth to account for observed behaviors in number and price comparison tasks (see Nuerk et al., 2015 for discussion). For example, the holistic model of numerical processing proposes

that numerical symbols are represented by converting numerical symbols into internal approximate magnitudes along a mental number line; thus, it is the precision of this number to magnitude conversion that guides responses (Dehaene, Dupoux, & Mehler, 1990). Thomas and Morwitz (2005) applied a similar idea to account for left digit bias, suggesting that the magnitude encoding for multidigit numbers begins even before we finish reading the number (i.e., as soon as we read the digit “2” in “2.99”). According to this account, the left-to-right processing of numerical symbols (i.e., nine-ending prices such as \$2.99) affects the magnitude conversion process, and the encoded holistic magnitude gets anchored on the leftmost digit (i.e., 2), distorting the perceived price magnitude toward the leftmost digit of a multidigit number (e.g., the perceived magnitude of \$2.99 is significantly lower than the perceived magnitude of \$3.00).

In contrast to the above proposals, other researchers have posited that multidigit numbers are processed in a decomposed fashion and broken down into individual digit representations (i.e., hundreds, decades, units) for comparison based on place-value structure (e.g., Verguts & De Moor, 2005). For example, evidence from number comparison studies involving both compatible (e.g., 42 vs 57, where $4 > 5$ and $2 < 7$) and incompatible digit-level information (e.g., 47 vs 62, where $4 > 6$, but $7 > 2$) suggests that separate digit-level representations contribute to task performance (see Nuerk, Weger, & Willmes, 2001). Nuerk and Willmes (2005) proposed a hybrid model incorporating holistic and decomposed processing that assumes both holistic and individual digit level magnitudes are accessed and then compared in parallel to guide task responses.

While we did not design the current studies to test competing models of multidigit number processing, any model of numerical processing relevant to NLE tasks should account for the present findings. For example, a recent model developed by Dotan and Dehaene (2020) assumes that individual digits are first quantified independently and weighted according to their roles (e.g., hundreds, tens, ones), and then merged to a whole-number quantity representation that informs judgments (e.g., placements in NLE). It could be that the left digit effect in number line estimation arises from the weighting of digit-based magnitudes used to inform whole-number quantities; this is another possible direction for future research.

Given the widespread use of NLE tasks as both research and instructional tools, it is important that we understand the role of digit identity in estimation performance. The findings here are novel: the left digit effect is a broad phenomenon in NLE that applies to two-digit numbers in the 0–100 range as it does to three-digit numbers in the 0–1000 range, emerging as early as age 7 or 8 and continuing into adulthood. Theories of number line estimation will need to accommodate the findings here that the identity of a two-digit target numeral’s leftmost digit (rather than just the numeral’s overall magnitude) influences NLE performance. These findings further provide additional evidence for a new kind of variation in numerical estimation performance across individuals and developmental periods that remains unexplored. In addition, remaining questions about the generalizability and potential sources of this family of effects have important implications for our understanding of the numerical and spatial skills that might contribute to the links between numerical estimation and more formal mathematics learning and achievement.

Acknowledgments

Thank you to the participating children and families; to Wesleyan Cognitive Development Lab researchers Sophie Charles, Lauren Barragan, Esha Bhandari, Kerry Brew, Jenny Chelmow, Taylar Clark, Sarah Corner, Jamie Hom, Meghana Kandlur, Ilana Ladis, Maxine Lai, Maya Layne, Jordan Feingold-Link, Carolina Montano, Sarah Aduke Ohiomah, Aime Arroyo-Ramirez, Sheri Reichelson, Elizabeth Shackney, Steven Fields, Amanda Fiorentino, Praise Owoyemi, Joanna Paul, Ray Alexander, and Katie Vasquez for their assistance with recruitment and data collection; and to Leah Vaidya for assistance with manuscript preparation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by NSF DRL-1920445 and benefited from NSF DRL-1561214, both to HB and ALP.

References

- Ashcraft, M. H., & Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *Journal of Experimental Child Psychology*, 111(2), 246–267. doi:10.1016/j.jecp.2011.08.005
- Barth, H., Slusser, E., Kanjlia, S., Garcia, J., Taggart, J., & Chase, E. (2016). How feedback improves children's numerical estimation. *Psychonomic Bulletin & Review*, 23(4), 1198–1205. doi:10.3758/s13423-015-0984-3
- Beracha, E., & Seiler, M. J. (2015). The effect of pricing strategy on home selection and transaction prices: An investigation of the left-most digit effect. *JSTOR Journal of Housing Research*, 24(2), 147–162. doi:10.1080/10835547.2015.12092101
- Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, 46(2), 545–551. doi:10.1037/a0017887
- Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, 42(1), 189–201. doi:10.1037/0012-1649.41.6.189
- Byrge, L., Smith, L. B., & Mix, K. S. (2014). Beginnings of place value: How preschoolers write three-digit numbers. *Child Development*, 85(2), 437–443. doi:10.1111/cdev.12162
- Choi, J., Jessica, L., Samper, A., Morwitz, V. G., Campbell, M. C., & Bolton, L. E. (2019). The influence of health motivation and calorie ending on preferences for indulgent foods. *Journal of Consumer Research*, 46(3), 606–619. doi:10.1093/jcr/ucz002
- 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. doi:10.1037/0096-1523.16.3.626
- Dotan, D., & Dehaene, S. (2020). Parallel and serial processes in number-to-quantity conversion. *Cognition*, 37, 104387.
- Eyler, R. F., Cordes, S., Szymanski, B. R., & Fraenkel, L. (2018). Use of feedback to improve mental number line representations in primary care clinics. *BMC Medical Informatics and Decision Making*, 18(1), 40. doi:10.1186/s12911-018-0618-6
- Fazio, L., Kennedy, C., Siegler, R., & De Smedt, B. (2016). Improving children's knowledge of fraction magnitudes. *PLoS ONE*, 11(10), 1–14. doi:10.1371/journal.pone.0165243
- Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction*, 7(4), 343–403. doi:10.1207/s1532690xci0704_4

- Gunderson, E. A., Hamdan, N., Hildebrand, L., & Bartek, V. (2019). Number line unidimensionality is a critical feature for promoting fraction magnitude concepts. *Journal of Experimental Child Psychology*, 187, 104657. doi:10.1016/j.jecp.2019.06.010
- 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. doi:10.1037/dev0000252
- Jung, S., Roesch, S., Klein, E., Dackermann, T., Heller, J., & Moeller, K. (2020). The strategy matters: Bounded and unbounded number line estimation in secondary school children. *Cognitive Development*, 53, 100839. doi:10.1016/j.cogdev.2019.100839
- Lai, M., Zax, A., & Barth, H. (2018). Digit identity influences numerical estimation in children and adults. *Developmental Science*, 21, 12657. doi:10.1111/desc.12657
- Lai, M., Zax, A., & Barth, H. (2018). Digit identity influences numerical estimation in children and adults. *Developmental Science*, 21, 12657. doi:10.1111/desc.12657
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, 78(6), 1723–1743. doi:10.1111/j.1467-8624.2007.01087.x
- MacKillop, J., Amlung, M. T., Blackburn, A., Murphy, J. G., Carrigan, M., Carpenter, M. J., & Chaloupka, F. (2014). Left-digit price effects on smoking cessation motivation. *Tobacco Control*, 23(6), 501–506. doi:10.1136/tobaccocontrol-2012-050943
- Mix, K. S., Prather, R. W., Smith, L. B., & Stockton, J. D. (2014). Young children's interpretation of multidigit number names: From emerging competence to mastery. *Child Development*, 85(3), 1306–1319. doi:10.1111/cdev.12197
- Moeller, K., Nuerk, H.-C., & Willmes, K. (2009). Internal number magnitude representation is not holistic, either. *European Journal of Cognitive Psychology*, 21, 672–685. doi:10.1080/09541440802311899
- Nuerk, H.-C., Kaufmann, L., Zopoth, S., & Willmes, K. (2004). On the development of the mental number line: More, less, or never holistic with increasing age? *Developmental Psychology*, 40, 1199–1211. doi:10.1037/0012-1649.40.6.1199
- Nuerk, H.-C., Moeller, K., & Willmes, K. (2015). Multi-digit number processing—Overview, conceptual clarifications, and language influences. In *Oxford handbook of mathematical cognition* (pp. 106–139). New York, NY: 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–B33. doi:10.1016/S0010-0277(01)00142-1
- Nuerk, H.-C., & Willmes, K. (2005). On the magnitude representations of two-digit numbers. *Psychology Science*, 47, 51–72.
- Olenski, A. R., Zimmerman, A., Coussens, S., & Jena, A. B. (2020). Behavioral heuristics in coronary-artery bypass graft surgery. *New England Journal of Medicine*, 382(8), 778–779. doi:10.1056/NEJMc1911289
- Opfer, J., & Siegler, R. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, 55(3), 169–195. doi:10.1016/j.cogpsych.2006.09.002
- Peeters, D., Degrande, T., Ebersbach, M., Verschaffel, L., & Luwel, K. (2016). Children's use of number line estimation strategies. *European Journal of Psychology of Education*, 31(2), 117–134. doi:10.1007/s10212-015-0251-z
- Peeters, D., Verschaffel, L., & Luwel, K. (2017). Benchmark-based strategies in whole number line estimation. *British Journal of Psychology*, 108(4), 668–686. doi:10.1111/bjop.12233
- Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*, 79(2), 375–394. doi:10.1111/j.1467-8624.2007.01131.x
- Sarnecka, B. W., & Gelman, S. A. (2004). Six does not just mean a lot: Preschoolers see number words as specific. *Cognition*, 92(3), 329–352. doi:10.1016/j.cognition.2003.10.001
- Savelkoul, S., Williams, K., & Barth, H. (2020). Linguistic inversion and numerical estimation. *Journal of Numerical Cognition*, 6, 263–274. doi:10.5964/jnc.v6i3.273

- 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), 1–18. doi:10.1111/cdev.13068
- Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science*, 19(3), 341–361. doi:10.1111/desc.12395
- Slusser, E. B., Santiago, R. T., & Barth, H. C. (2013). Developmental change in numerical estimation. *Journal of Experimental Psychology. General*, 142(1), 193–208. doi:10.1037/a0028560
- 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. doi:10.1086/429600
- Verguts, T., & De Moor, W. (2005). Two-digit comparison. *Experimental Psychology*, 52(3), 195–200. doi:10.1027/1618-3169.52.3.195
- 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 and Behavior*, 00, e01877. doi:10.1002/brb3.1877
- Xenidou-Dervou, I., van der Schoot, M., & van Lieshout, E. C. D. M. (2015). Working memory and number line representations in single-digit addition: Approximate versus exact, nonsymbolic versus symbolic. *The Quarterly Journal of Experimental Psychology*, 68(6), 1148–1167. doi:10.1080/17470218.2014.977303
- Xing, C., Zax, A., Zax, A., Taggart, J., Bass, I., & Barth, H. (2021). Numerical estimation strategies are correlated with math ability in school-age children. *Cognitive Development*, 60, 101089. in press doi:10.1016/j.cogdev.2021.101089
- Yuan, L., Prather, R. W., Mix, K. S., & Smith, L. B. (2019). Preschoolers and multi-digit numbers: A path to mathematics through the symbols themselves. *Cognition*, 189, 89–104. doi:10.1016/j.cognition.2019.03.013

Appendix

Table A1. Study 1 descriptive statistics for mean decade difference scores for each pair by age group.

Variable	8/12	18/22	29/31	38/42	49/51	59/61	69/71	78/82	88/92
Age 5									
<i>N</i>	11	11	11	11	11	11	11	11	11
<i>M</i>	−4.02	1.86	2.69	2.76	−18.9	−0.75	1.10	1.65	3.50
<i>SD</i>	23.3	23.1	32.9	19.2	30.8	21.3	21.3	15.2	23.3
<i>t</i>	−0.57	0.27	0.27	0.48	−2.05	−0.12	0.17	0.36	0.50
<i>p</i>	0.710	0.398	0.396	0.321	0.966	0.545	0.433	0.363	0.314
Age 6									
<i>N</i>	20	21	20	21	19	19	19	19	20
<i>M</i>	−1.15	−1.85	−1.19	−0.172	3.91	2.55	4.31	2.27	0.730
<i>SD</i>	10.1	12.0	11.0	14.7	15.7	12.4	12.53	20.73	10.62
<i>t</i>	−0.51	−0.71	−0.48	−0.05	1.09	0.89	1.498	0.477	0.307
<i>p</i>	0.692	0.756	0.683	0.521	0.146	0.191	0.076	0.320	0.381
Age 7									
<i>N</i>	19	19	19	19	19	18	18	18	19
<i>M</i>	−1.17	0.41	2.30	−1.33	3.69	−1.59	5.52	−0.45	1.88
<i>SD</i>	8.63	10.6	13.2	7.46	10.2	8.03	11.8	12.6	10.1
<i>t</i>	−0.59	0.17	0.74	−0.76	1.53	−0.81	1.91	−0.15	0.81
<i>p</i>	0.718	0.435	0.234	0.770	0.072	0.786	0.037*	0.559	0.214
Age 8									
<i>N</i>	19	19	18	19	19	19	21	20	20
<i>M</i>	3.99	3.07	4.30	−3.27	4.10	−4.53	4.94	0.88	4.34
<i>SD</i>	6.66	9.97	8.82	11.1	7.79	8.69	10.69	6.53	5.36
<i>t</i>	2.61	1.34	2.07	−1.29	2.29	−2.27	2.12	0.60	3.62
<i>p</i>	0.009*	0.098	.027*	0.893	0.017*	0.982	0.023*	0.277	0.001**

Reported *t*-tests were right tailed.

* significant at $p < 0.05$ level (uncorrected)

** significant at $p < 0.005$ level (alpha level with correction for multiple comparisons)

Table A2. Study 2 descriptive statistics for individual decade difference scores for each decade pair.

Variable	8/12	18/22	29/31	38/42	49/51	59/61	69/71	78/82	88/92
<i>N</i>	44	44	44	44	44	44	44	44	44
<i>M</i>	−0.34	−0.70	−0.80	2.78	0.66	0.71	2.49	−1.37	2.50
<i>SD</i>	1.56	2.04	2.06	3.12	1.75	1.34	3.50	2.16	3.10
<i>t</i>	−1.46	−2.29	−2.56	5.92	2.49	2.02	4.72	−4.20	5.35
<i>p</i>	0.924	0.987	0.993	<0.001**	0.008*	0.025*	<0.001**	0.999	<0.001**

Reported *t*-tests were right tailed.

* significant at $p < 0.05$ level (uncorrected)

** significant at $p < 0.005$ level (alpha level with correction for multiple comparisons)

Table A3. Study 3 descriptive statistics for mean decade difference scores for each pair.

Variable	8/12	18/22	29/31	38/42	49/51	59/61	69/71	78/82	88/92
<i>N</i>	25	25	24	26	25	23	24	26	26
<i>M</i>	1.76	2.76	3.21	5	5.2	0.35	6.38	1.96	3.38
<i>SD</i>	4.57	8.84	6.09	10.48	7.53	6.29	7.26	7.74	5.72
<i>t</i>	1.93	1.56	2.58	2.43	3.45	0.27	4.30	1.29	3.02
<i>p</i>	0.033*	0.066	0.008*	0.011*	0.001**	0.397	<0.001**	0.104	0.002**

Reported *t*-tests were right tailed.

* significant at $p < 0.05$ level (uncorrected)

** significant at $p < 0.005$ level (alpha level with correction for multiple comparisons)