# Influence of Benchmarks on the Left Digit Effect in Number Line Estimation

Prakriti Mittal (with Kelsey Kayton, Hilary Barth, & Andrea L. Patalano)
Department of Psychology, Wesleyan University

## Introduction

Number line estimation tasks are widely used as assessment tools and as reliable predictors of math outcomes.[1] On a typical task, participants are asked to estimate the location of Arabic numerals on a bounded number line. Recent evidence reveals a novel source of error in number line estimation performance, as described below.

**Left Digit Effect:** A numerical bias in which numbers with nearly identical magnitudes but different left digits are estimated to be significantly different from each other, while numbers with an identical difference in magnitude but same left digit are not.[2]

*e.g., 701 is estimated as significantly larger than 699*
*but 651 is estimated as about the same as 649*

Previous studies, using motivational incentives and accuracy feedback,[3] have been unsuccessful in reducing the left digit effect. Training with benchmarks (spaced reference points on the line) has been shown to reduce accuracy error generally[4] but it is not yet known whether benchmarks specifically reduce the left digit effect.

Here, we investigate whether a benchmark intervention might reduce the left digit effect, and also whether the left digit effect in number line estimation is related to a similar effect in complex judgment.

## Study 1: Benchmark Number Line Estimation

Does a benchmark intervention on a number line estimation task lead to a reduction in the left digit effect?

## Method

**Participants.** A total of 165 undergraduates were randomly assigned to one of the two conditions described below.

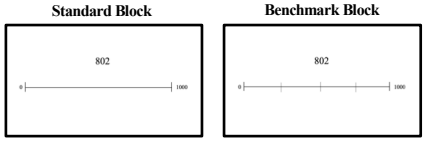
**No-Benchmark Condition:** (three blocks of 60 trials each)
Participants clicked with a mouse on the 0-1000 line to indicate the location of each target numeral.

**Benchmark Condition** (same but benchmarks in 2nd block)
In the second block, participants are shown unlabelled benchmarks (vertical lines) were placed at 250, 500, and 750.

Target numerals were grouped (for analyses only) as follows:
- **Hundreds pairs**: 8 pairs around 100's boundaries (899, 901)
- **Fifties pairs**: 9 pairs around 50's boundaries (449, 451)
- Non-boundary values (942)

Hundreds pairs were critical for assessing the left digit effect, and fifties pairs served as controls; non-boundary values were used to compute overall accuracy error.

**Figure 1.** *Number line without versus with benchmarks*




## Measures and Predictions

For each pair of target numerals, we calculated a difference score: *(placement of larger numeral – placement of smaller numeral)*. We then calculated one average hundreds difference score and one average fifties difference score per participant per block.

**hundreds difference scores > 0 indicates a left digit effect**

## Measures and Predictions (continued)

- **If benchmarks reduce the left digit effect:** Across blocks, hundreds difference scores will decrease more in the Benchmark than the No-Benchmark Condition.
- **If benchmarks *do not* reduce the left digit effect:** Any improvement in hundreds difference scores across blocks will be the same in both conditions.

**Overall accuracy error**

We calculated percent absolute error (PAE): $|placement\ of\ numeral - correct\ location|/1000$. A smaller PAE indicates lower error.

- **If benchmarks reduce the overall error:** Across blocks, PAE will decrease more in the Benchmark than the No-Benchmark Condition.
- **If benchmarks *do not* reduce the left digit effect:** Any improvement in PAE across blocks will be the same in both conditions.

## Results

**Left digit effect**

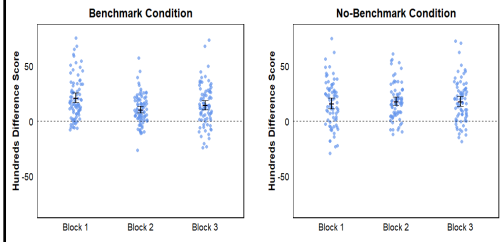
A left digit effect was present: the hundreds difference score was reliably greater than 0 in all blocks for both the Benchmark ($ts > 7, ps < .001, ds = 0.77 – 1.04$) and No-Benchmark Condition ($ts > 6, ps < .001, ds = 0.78 – 1.10$). The fifties difference scores were not reliably greater than 0 in any block for the Benchmark ($ts < 1, ps > .790$) or No-Benchmark Condition ($ts < 1, ps > .660$).

**Left digit effect by condition**

The benchmark intervention led to a sustained reduction in the left digit effect as there was a condition by block interaction for the hundreds difference scores, $F(2,312) = 6.07$, $MSE = 242.19$, $p = .003$, $\eta_p^2 = .04$.

The presence of benchmarks also led to better overall accuracy on the task as there was a condition by block interaction for PAE $F(2, 312) = 54.00$, $MSE < 0.001$, $p < .001$, $\eta_p^2 = .26$.

**Figure 2.** *Hundreds difference score by block and condition*




## Study 2: Consumer Product Judgments

digit effect in perceptions of costliness of consumer products?

## Method

**Procedure.** Participants saw four desks one at a time (in one of four counterbalanced orders) and rated the costliness of each compared to the industry standard (costing $500). Ratings were on a line from "Inexpensive" to "Expensive", and translated to a value from 0-1000.

A "boundary pair" of desks had similar magnitudes but different left digits ($699 and $701) and a matched "non-boundary pair" had the same left digits ($349 and $351).

## Measures

For each participant, we calculated a judgment difference score**:**
***(high boundary response – low boundary response) –***
***(high non-boundary response – low non-boundary response)***

**judgment difference scores > 0 indicates a left digit effect**

## Results

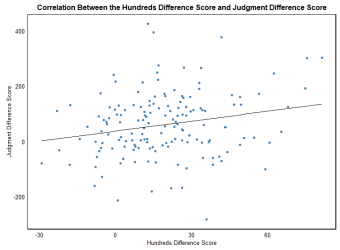
Ratings of high ($M = 724$, $SD = 103$) and low ($M = 664$, $SD = 109$) boundary prices were more different from one another than ratings of (similarly spaced) high ($M = 307$, $SD = 113$) and low ($M = 299$, $SD = 113$) non-boundary prices, consistent with left digit effect.

The judgment difference score was $M = 51$ ($SD = 134$, *range* = -603 – 424) and was significantly different from 0, $t(157) = 4.82$, $SE = 10.66$ $p < .001$, $d = 0.38$, $CI[30.28, 72.40]$), indicating a left digit effect.

## Number Line and Product Judgment Relationship

After removing two outliers with low scores we calculated a correlation between judgment difference score and hundreds difference score. One's judgment difference score was significantly positively correlated with one's hundreds difference score in the first block of Study 1, $r(154) = .21$, $p = .007$.

**Figure 3.** *Scatterplot of relationship between left digit effect measures*




## Conclusions and Future Directions

task reduced the left digit effect and also overall accuracy error.

We also found that the left digit effect in number line estimation was correlated with that of a more complex judgment task.

**Future Questions**
- Are improvements in the left digit effect resulting from the benchmark intervention sustained over time?
- Is the magnitude of one's left digit effect related to various cognitive individual difference measures (e.g., working memory).

## References

(2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development, 89*, 1467-1484.
2. Lai, M., Zax, A., & Barth, H. (2018). Digit identity influences numerical estimation in children and adults. *Developmental Science, 21*, e12657.
3. Kayton, K., Williams, K., Stenbaek, C., Gwiazda, G., Bondhus, C., Green, J., Fischer, G., Barth, H., & Patalano, A. L. (in press). Summary accuracy feedback in number line estimation does not reduce the left digit effect. *Memory & Cognition*.
4. Peeters, D., Verschaffel, L., & Luwel, K. (2017). Benchmark-based strategies in whole number line estimation. *British Journal of Psychology, 108*(4), 668–686.

We thank research students in the Reasoning and Decision Making Lab at Wesleyan University including Charlie Bondhus, Greg Fischer, Jordan Green, Gina Gwiazda, Camila Rodlauer, Claudia Stenbaek, Gillian Weeks, and Christina Xu. We also thank the QAC Summer Apprenticeship 2022. This work was supported by NSF DRL-1920445 and benefited from DRL-1561214 to HB and ALP.