UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The mental representation of integers: Further evidence for the negative number line as a reflection of the natural number line

Permalink

https://escholarship.org/uc/item/8629819g

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

Authors

Vest, Nicholas
Alibali, Martha W

Publication Date

2021

The Mental Representation of Integers: Further Evidence for the Negative Number Line as a Reflection of the Natural Number Line

Nicholas A. Vest (navest@wisc.edu)
Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Martha W. Alibali (mwalibali@wisc.edu)
Department of Psychology, 1202 W. Johnson Street
Madison, WI 53706 USA

Abstract

Humans are able to make sense of extraordinarily abstract concepts in mathematics (e.g., negative numbers). What is the underlying representation of these concepts? Integers extend natural numbers by including zero and negative numbers. To study the mental representation of integers, we employed a number comparison task in an online context. We replicated the previously-reported distance effect, in that far comparisons were faster than near comparisons. Namely, we observed reliable distance effects for positive and negative comparisons, and critically, an inverse distance effect for mixed comparisons. We conclude that the mental representation of integers may align with a hypothesis proposing the mental number line for negative numbers mirrors the natural number line. Moreover, we conclude that web-based data collection is a promising tool for future numerical cognition research.

Keywords: numerical cognition; distance effect; mathematics

Introduction

People's initial intuitions about numbers are based, at least in part, on physical experience with countable numbers and observable quantities (e.g., Piaget, 1971). But what about more abstract concepts, for which people do not have physical or perceptual experience? For example, "one" and "two" readily map onto quantities of physical objects (e.g., one jellybean and two jellybeans). But it is much more difficult to display zero, negative one, or negative two jellybeans. How, then, do people acquire an understanding of these abstract concepts, and what are the cognitive underpinnings of quantities such as "negative two"? Through examining behavior via timed trials (such as comparing the magnitudes of pairs of numbers), we can shed light on these questions.

Symbolic magnitude comparison is a commonly studied activity in the field of numerical cognition. Tasks designed to measure magnitude comparison typically involve collecting reaction times and accuracy data from two-choice displays that require binary judgments (e.g., Moyer & Landauer, 1967). Using this type of task, we can examine the behavioral effects manifested in judging the magnitudes of symbolic numbers.

One behavioral effect that has been widely studied is the *distance effect*. The distance effect is the phenomenon that people tend to make number comparisons faster when comparing numbers that are far apart (e.g., 2 vs. 9) than when comparing numbers that are close together (e.g., 2 vs. 4). This effect has been well documented with natural (i.e., positive) numbers (Dehaene et al., 1990; Holloway & Ansari, 2009). However, natural numbers represent only a portion of the symbolic number system.

Integers, which include zero and negative numbers, are an abstract extension of the natural numbers. Zero and negatives are part of a quantitative system that is culturally constructed. Unlike the perceptual sense of magnitude that accompanies natural numbers, negative numbers and zero do not exhibit an obvious mapping to basic perceptual representations. This raises the question of how we represent these abstract quantities. Is the same distance effect observed in people making comparisons with negative numbers?

Previous literature suggests the distance effect may vary for comparisons of different types of numbers (e.g., positive-positive, negative-negative, positive-negative, positive-zero, negative-zero). One hypothesis, termed the *symbol+* hypothesis (Varma et al., 2019), holds that, because zero and negative numbers do not have readily available perceptual mappings, people initially understand them by using symbolic rules (e.g., $-x$ is always less than $+x$ or $+y$). From this perspective, people's initial understanding of negative numbers is in terms of a rule-based conception based on positive number representations (Fischer, 2003). Because people have no need to consult magnitude representations when comparisons involve negative numbers, they should not show distance effects for such comparisons, even comparisons of a positive and a negative number (i.e., mixed comparisons).

Another hypothesis, termed the *analog+* hypothesis (Varma et al., 2019), proposes that, with time and experience, people transform their initial representations of integers by extending the mental number line for natural numbers "to the left" of zero, to include negative integers, consequently reconstructing their *magnitude* representation of natural numbers to include negatives (Fischer, 2003). The analog+ hypothesis predicts that, because positives and negatives make up the same mental number line, all comparisons of integers will show distance effects. Studies examining the symbol+ and analog+ hypotheses via timed magnitude comparisons have yielded mixed findings (Tzelgov et al., 2009; Krajcsi & Igács, 2010).

Varma and Schwartz (2011) proposed a third representation of integers, termed the *analog-x* hypothesis. This hypothesis holds that the magnitude representation of number is restructured by the integer symbol system. Importantly, this hypothesis suggests that integers are understood as points on a mental number line, and that the negative number line is a *reflection* of the natural number line, similar to a reflection in a mirror. Importantly, the analog-*x* model predicts that mixed comparisons (i.e., comparisons of a positive and a negative number) will show an *inverse* distance effect: magnitude comparisons will be slower for pairs of a positive and negative number that are further apart. Due to the psychophysical scaling of magnitude representations of the mental number line for natural numbers and its reflection for negative numbers (see Varma et al., 2019), the difference between the (perceived) distances from 0 of the positive and the negative numbers is greater for mixed pairs that are close together (e.g., -2, 1) than for pairs that are farther apart (e.g., -2, 6), yielding an inverse distance effect. Indeed, Varma and Schwartz (2011) found the predicted interaction between comparison type and distance, such that participants demonstrated an inverse distance effect for mixed comparisons.

In this work, we sought to replicate the findings reported by Varma and Schwartz (2011), and to extend these findings from an in-person data collection setting to an online setting. In recent years, a plethora of tools have been developed to implement timed behavioral studies online such as jsPsych, PsychoPy, and Gorilla (de Leeuw, 2015; Peirce et al., 2019; Anwyl-Irvine et al., 2020). Research examining the efficacy of these resources has been promising. In a recent survey of over 200,000 participants, Anwyl-Irvine and colleagues (2020) found that modern web-platforms provide reasonable accuracy and precision for display duration and manual response time. It is not surprising then, that when Huber and colleagues (2017) studied a large sample of adults comparing two-digit positive numbers in an online setting, they replicated effects found in lab, including the distance effect.

In the current study, we sought to examine the numerical distance effect in an online setting with positive and negative integers. We used Varma and Schwartz's (2011) stimuli and materials, and we extended their prior research to examine behavioral response times collected via an online platform.

We seek to address these questions: (1) Can we collect precise reaction times for numerical comparison tasks through an online research tool? Assuming we can, (2) Do participants display distance effects for magnitude comparisons that involve positive integers, negative integers, and zero? (3) Do the distance effects observed from data collected online replicate Varma and Schwartz's (2011) pattern of findings, therefore providing support for the analog-*x* hypothesis? We hypothesize that data collected online will indeed display behavioral effects similar to those reported by Varma and Schwartz (2011). However, we further hypothesize that reaction times collected online will display lags similar to those found in previous online studies (see Bridges et al., 2020).

Method

Participants

Fifty-five participants (64% female) completed an online survey via Qualtrics that lasted approximately 30 minutes. 67% of the sample self-identified as White, 13% as Asian, 5% as Hispanic or Latino, 4% as African American, 2% as Hawaiian/Pacific Islander, , 4% as some other racial or ethnic identity, and 5% as more than one race or ethnicity. Participants attended a university in the Midwestern United States and were recruited from an undergraduate participant pool for an introductory psychology course. Each participant received extra credit in their Introduction to Psychology course for participating.

Stimuli and Design

Materials, stimuli, and the preregistration are publicly available through the Open Science Framework at https://osf.io/km3tz/?view_only=57e22f1b612644f594ad5c961643c136.

Understanding of Zero Participants answered problems at either the beginning or end of the study (counter-balanced order) to assess their knowledge of zero and the additive inverse principle. Assessment items included two questions about zero (e.g., "What is zero?"), three additive inverse problems (solve and explain: $x + -x = 0$), a number line task ("Choose where 3 is on the number line."), a question about the additive identity principle, and a question about the additive inverse principle. These questions addressed a secondary research question that we do not address in this report, so data from this assessment are not considered herein.

Magnitude Comparison Task To investigate the mental representation of integers, participants engaged in a timed magnitude comparison task that was adapted from Varma and Schwartz (2011). Participants were presented number comparisons in PsychoPy (Peirce et al., 2019) hosted online via Pavlovia (Figure 1). The integers were grouped into five sets: small natural numbers {1, 2, 3, 4}, large natural numbers {6, 7, 8, 9}, small negative integers {-1, -2, -3, -4}, large negative integers {-6, -7, -8, -9}, and zero {0}. An integer was considered small if $|x| < 5$ and large if $|x| > 5$.


(a) Fixation cross

(b) Stimulus pair

Figure 1: Example of stimuli displayed on PsychoPy and Pavlovia for magnitude comparison task.

The stimuli were pairs of integers that formed three fully crossed within-subjects factors: comparison type, distance, and predicate. Comparison type had four levels: positive, negative, mixed, and zero. Positive comparisons involved comparing two positive numbers, negative comparisons involved two negative numbers, mixed comparisons involved one positive and negative number, and zero comparisons involved zero and either a positive or negative number. Further, mixed and zero comparisons were sorted into sub-categories: mixed-positive, mixed-negative, zero-positive, and zero-negative. Sub-levels were characterized by the value of the number that had the greater absolute value.

The second factor, comparison distance, had two levels: near and far. Distance was determined by the mathematical distance between the integers being compared ($|x - y|$). Near-distance stimuli had a distance of 2 or 3, and far-distance stimuli had a distance of 7 or 8. For positive, negative, and zero comparisons, near-distance stimuli were created by pairing two small integers or pairing two large integers. Far-distance stimuli were created by pairing a small integer with a large integer. Mixed comparisons were created differently. Near-distance stimuli paired a small positive integer and a small negative integer (e.g., -2 vs. 1). Far-distance stimuli paired a small negative integer and a large positive integer (e.g., -2 vs. 6) or vice versa (e.g., -6 vs. 2).

The five sets of integers and these rules for pairing integers yielded 44 stimuli. There were four blocks, and each pair appeared twice within each block, once in each left-right order (e.g., 1 vs. 9, 9 vs. 1). Thus, there was a total of 88 stimuli per block.

The third factor, comparison predicate, had two levels: choose the *greater* or the *lesser* number. Predicate was varied across blocks so as to minimize participant confusion within a block. The order was consistent across participants (greater, lesser, greater, lesser).

Demographic Questionnaire At the end of the session, participants completed a 12-item demographic questionnaire. The questionnaire requested information about participants' age, gender, race and/or ethnicity, previous mathematics courses taken, and standardized mathematics test scores (SAT or ACT).

Procedure

Participants saw eight assessment items related to zero and the additive inverse principle on Qualtrics at either the beginning or end of the study; data from this assessment are not analyzed here.

For the magnitude comparison task, participants were provided a hyperlink that opened a new tab in their browser. The hyperlink directed them to Pavlovia, which displayed stimuli on the participant's computer in their browser. Using the left or right arrow keys on their keyboard, participants made greater and lesser judgements. Participants first completed practice blocks of 12 greater and 12 lesser judgements sampled from each cell of the design. Next, participants completed four experimental blocks.

On each trial, a fixation cross was displayed for 500 ms, followed by the pair of numbers. The fixation cross and number pairs appeared in white Arial font. On PsychoPy, we specified a scaling factor of 0.1 for letter height (i.e., a vertical scale proportionally applied to the width of each character). The fixation cross and number pairs were displayed in center position (Figure 1). Positive numbers were presented without a "+" sign. Each block lasted approximately 1.5 min, and the overall experiment lasted approximately 30 min.

Following the experimental tasks, participants completed the demographic questionnaire.

Results

We excluded a total of 12.73% of responses from the analyses due to delays or anticipations, RTs more than ± 3 SDs from the individual's mean, and incorrect responses. Specifically, we trimmed responses outside of the interval of 200-2000 ms (2.62% of the data) and responses more than three SDs from each participant's mean (1.62% of the data). Exclusions were made sequentially in the order indicated. The error rate was low (*M* = 3.00%, *SEM* = 0.79%).

There was also no speed-accuracy tradeoff; the correlation between mean response time and error rate across the 16 cells of the factorial design was low and non-significant, ($r(53) = -0.23$, $p = 0.099$). Therefore, all reported analyses are based on response times on correct trials.

Table 1: Distance effects for current study and Varma and Schwartz (2011) reported in milliseconds.

Distance Effect (Current study)						
	Overall[1]	Far	Near	*t*(54)[2]	*p*[3]	Direction
Positive	726 (32)	705 (31)	747 (32)	-5.645	< 0.001	Standard
Negative	847 (36)	829 (36)	865 (36)	-5.778	< 0.001	Standard
Mixed[4]	699 (31)	705 (31)	693 (31)	2.361	0.018	Inverse
Mixed-positive	705 (32)	706 (32)	703 (32)	0.603	0.547	
Mixed-negative	703 (31)	706 (31)	700 (31)	0.722	0.443	
Zero	703 (30)	697 (30)	708 (29)	-1.339	0.186	
Zero-positive	708 (30)	700 (30)	716 (30)	-1.614	0.113	
Zero-negative	702 (30)	701 (30)	704 (30)	-0.306	0.760	

Distance Effect (Varma & Schwartz, 2011)						
	Overall	Far	Near	*t*(20)	*p*	Direction
Positive	604 (17)	588 (16)	620 (18)	4.324	< 0.001	Standard
Negative	742 (24)	722 (23)	763 (25)	4.464	< 0.001	Standard
Mixed	579 (18)	592 (19)	565 (18)	-4.917	< 0.001	Inverse
Mixed-positive	570 (17)	583 (19)	557 (17)	-2.456	0.023	Inverse
Mixed-negative	585 (19)	600 (20)	570 (19)	-3.826	0.001	Inverse
Zero	608 (17)	608 (17)	607 (18)	-0.072	0.943	
Zero-positive	599 (18)	591 (18)	607 (20)	1.245	0.227	
Zero-negative	614 (18)	623 (19)	606 (18)	-1.752	0.095	(Inverse)

[1] Measures are reported *M* (*SEM*).
[2] Comparisons of near-far.
[3] *P*-values are from paired t-tests.
[4] This set also included one near-distance pair in which both numbers had the same absolute value (-1, 1).

The magnitude effects are of primary interest, as they directly inform the nature of the mental representation of integers. We specifically analyzed whether participants displayed a distance effect for each of the four comparison types (positive, negative, mixed, and zero).

We fit linear mixed-effects models (LMEMs) with reaction time (ms) as the dependent variable and comparison type (positive, negative, mixed, zero), distance (near, far), predicate (greater, lesser), and all possible interactions as predictors. The models included a by-subject random intercept and a by-subject random slope for each predictor (main effects and interactions). Next, we performed pairwise comparisons to test whether the effect of distance differed across comparison types. We used dummy coding to compare the different comparison types. The positive comparison, which is the standard in the literature, was designated the reference group. In our final model, we included these dummy codes, distance, predicate, all possible interactions, a by-subject random intercept, and by-subject random slopes for each predictor. We fit LMEMs using the statistics software RStudio and the R package *lme4* (Bates, Maechler, Bolker, & Walker, 2015). To obtain *p*-values for the fixed effects, we used the Satterthwaite approximation for degrees of freedom available via the R package *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016).

First, we ran an omnibus test to examine whether it was appropriate to use nonorthogonal contrasts to compare performance on the different types of comparisons (four-level predictor). Indeed, the omnibus test for the main effect of comparison was significant ($F(3, 57.38) = 109.09$, $p < 0.001$). Next, using Fisher's Least Significant Difference (LSD) method, we found that positive comparisons were faster than negative comparisons ($F(1, 52.64) = 215.92$, $p < 0.001$), slower than mixed comparisons ($F(1, 53.78) = 22.34$, $p < 0.001$), and slower than zero comparisons ($F(1, 59.61) = 16.79$, $p < 0.001$). There was also an overall effect of distance ($F(1, 78.78) = 32.44$, $p < 0.001$), with far comparisons ($M = 734$ ms, $SEM = 33$ ms) faster than near comparisons ($M = 754$ ms, $SEM = 34$ ms).

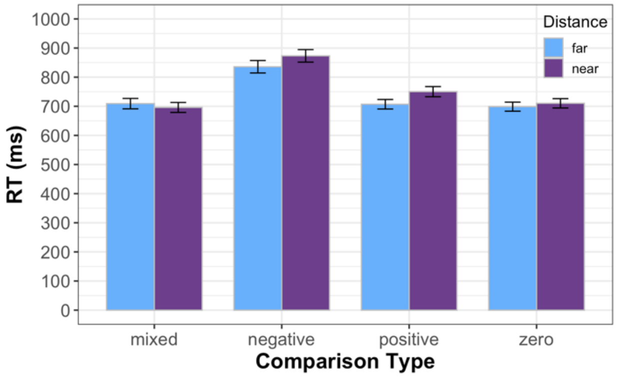

Figure 2: Plot displaying comparison type by distance interaction. Error bars reflect standard error.

Based on Varma and Schwartz (2011), we expected to find a significant interaction between comparison type and distance, and indeed, this interaction was reliable, $F(3, 83.91) = 16.18$, $p < 0.001$. For positive comparisons (the reference group), participants responded faster on far comparisons than on near comparisons. The same was true for negative comparisons, and the size of the effect for negative comparisons did not differ from that for positive comparisons (*F*(1, 70.12) = 0.33, *p* = 0.565). For zero comparisons, the distance effect was similar in direction but was smaller in magnitude than that for positive comparisons (*F*(1, 61.37) = 8.08, *p* = 0.006). Critically, for mixed comparisons, participants displayed an *inverse* distance effect—that is, they responded faster on near comparisons (e.g., -2 vs. 3) than on far comparisons (e.g., -2 vs. 6;); thus, the pattern for mixed comparisons differed significantly from that for positive comparisons, *F*(1, 103.45) = 38.63, *p* < 0.001; see Figure 2). The inverse distance effect observed for mixed comparisons is in line with the analog-*x* hypothesis, which holds that participants mentally represent the negative number line as a *reflection* of the natural number line.

The inverse distance effect that we observed on mixed comparisons was consistent with that observed by Varma and Schwartz (2011). Across all participants, 39 out of 55 participants showed the inverse distance effect for mixed comparisons. For comparison, 42 out of 55 participants showed the standard (i.e., far > near) distance effect for positive comparisons, and 44 out of 55 showed the standard distance effect for negative comparisons (see Table 1 for paired t-tests).

Given the omnibus effects, we next conducted a more fine-grained analysis of mixed and zero comparisons by splitting them into mixed-positive, mixed-negative, zero-positive, and zero-negative comparisons. LMEMs with comparison type, distance, a random intercept, and random slopes revealed no significant interactions (Figures 3 and 4). These findings contrast with those reported by Varma and Schwartz (2011) who found inverse distance effects for both mixed-positive and mixed-negative comparisons (see Table 1). In the current study, participants were slightly faster for each type of comparison, but the difference was non-significant in each case.

Regarding the mixed comparisons, it is worth noting that the stimulus set (which we constructed according to the rules described by Varma and Schwartz, 2011) included one mixed-even pair (-1, 1) which was not included in these sub-analyses. Participants judged this pair substantially faster (*M* = 674.44 ms) than the mixed-positive (*M* = 704.98 ms) and mixed-negative pairs (*M* = 703.42 ms). This near-distance pair contributes to the inverse distance effect observed for mixed comparisons overall.

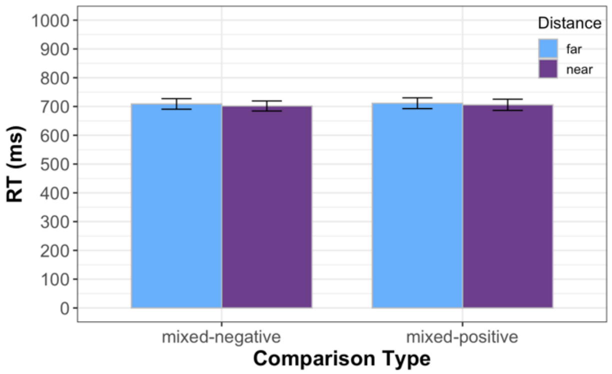

Figure 3: Plot displaying sublevels of mixed comparisons by distance interaction. Error bars reflect standard error.

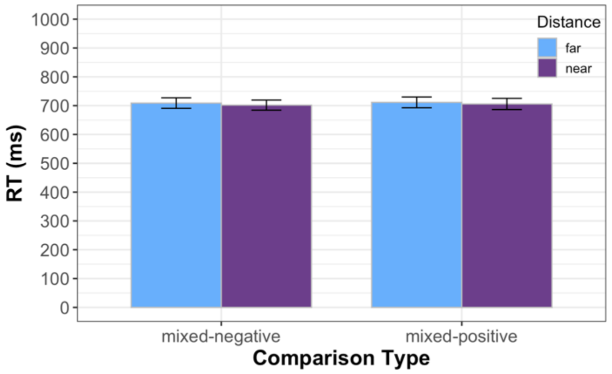

Figure 4: Plot displaying sublevels of zero comparisons by distance interaction. Error bars reflect standard error.

Discussion

In the present study, we investigated people's mental representations of integers in an online context. In a web-based platform for implementing behavioral experiments online participants made judgements of greater or lesser magnitudes for pairs of numbers. In this study, we asked the following questions: (1) Can we collect precise reaction times for numerical comparison tasks through an online research tool? If so, (2) Do participants display distance effects for magnitude comparisons that involve positive integers, negative integers, and zero? (3) Do the distance effects observed from online data replicate Varma and Schwartz's (2011) pattern of findings? We hypothesized that while response times would be lagged, we would detect reliable distance effects similar to those observed by Varma and Schwartz (2011).

Our results revealed that reaction times were indeed slower than those recorded by Varma and Schwartz (2011) in a lab setting. For example, our participants' mean RT was 726 ms for positive comparisons, while the mean RT in Varma and Schwartz's (2011) sample was 604 ms. Some online studies have found a lag of roughly 25-100 ms (de Leeuw & Motz, 2016; Hilbig, 2016), but this did not compare recent online services such as PsychoPy, and this lag varied by browser.

While this lag persists in online studies, several studies, including this one, have successfully replicated well-studied cognitive effects.

In this online study, we found distance effects for positive and negative comparisons, an *inverse* distance effect for mixed comparisons, and no effect for zero comparisons. The distance effects for positive and negative comparisons are consistent with several different hypotheses about the mental representation of integers. However, the inverse distance effect for mixed comparisons aligns with a specific claim about the mental representation of integers, the analog-*x* hypothesis, which proposes a reflection model in which the negative number line is a reflection of the natural number line.

Our findings are similar to that of Varma and Schwartz (2011). Indeed, Varma and Schwartz (2011) found distance effects for positive and negative comparisons, an *inverse* distance effect for mixed comparisons, and no effect for zero comparisons. In order to further investigate the mixed comparisons, Varma and Schwartz (2011) partitioned mixed comparisons into sublevels (mixed-negative and mixed-positive). Among the sublevels, they also found inverse distance effects for mixed-positive and mixed-negative comparisons.

In contrast to Varma and Schwartz (2011), in the current online study, the inverse distance effects were nonsignificant for mixed-positive and mixed-negative comparisons, considered separately. It is possible that these effects did not emerge in this study due to greater variability of response times due to online data collection. It is also worth noting that the inverse distance effect that we observed for mixed comparisons overall was driven, in part, by participants' fast RTs on the single, near-distance mixed-even pair. Because the far-distance set did not include a mixed-even pair, this represents a possible confound that may cloud the interpretation of the inverse distance effect observed for mixed pairs, overall. In future studies, we suggest that investigators should also include far-distance mixed-even pairs (such as -4, 4) in the stimulus set.

We recognize several limitations to this study. Because this study was delivered without an experimenter present, it is possible that participants were distracted during the experiment. Additionally, collecting data online can be affected by internet connections and browser compatibility. Thus, RT data collected online is likely less precise than RT data collected under more controlled circumstances. To ensure the integrity of the data collected in this study, we adapted inclusion/exclusion criteria identical to those used by Varma and Schwartz (2011) in their lab study. Further, the magnitude comparison task was short (approximately 6 minutes). Kochari (2019) notes, in effort to mitigate issues of distraction in online studies, one can administer shorter experiments. Finally, we may have had inadequate power to detect effects for sublevels of mixed comparisons. In future research, we will increase our sample size in order to more thoroughly investigate these types of comparisons.

The data from the current study suggests that online data collection serves as a promising tool for behavioral research on numerical cognition, and on magnitude comparison in particular. Our study provides evidence that subtle cognitive effects can be detected in RTs collected on a web-based platform, which makes it possible to administer experiments to a large heterogenous sample quickly. Critically, the data from this study replicate Varma and Schwartz's (2011) findings, which align with the hypothesis that the negative number line is a reflection of the natural number line and extend these findings to the online setting.

Acknowledgements

We thank Olympia Mathiaparanam and Holden Manhart for their help in study design and data analysis. This research was supported by a Kellet Midcareer Award from the University of Wisconsin-Madison to Martha Alibali.

References

Anwyl-Irvine, A., Dalmaijer, E., Hodges, N., & Evershed, J. (2020). Online Timing Accuracy and Precision: A comparison of platforms, browsers, and participant's devices. PsyArXiv.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*, 388-407.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*, 1-12.

de Leeuw, J. R., & Motz, B. A. (2016). Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, *48*, 1-12.

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*, 626.

Fischer, M. H. (2003). Cognitive representation of negative numbers. *Psychological Science*, *14*, 278-282.

Hilbig, B. E. (2016). Reaction time effects in lab-versus Web-based research: Experimental evidence. *Behavior Research Methods*, *48*, 1718-1724.

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*, 17-29.

Huber, S., Nuerk, H. C., Reips, U. D., & Soltanlou, M. (2019). Individual differences influence two-digit number processing, but not their analog magnitude processing: a large-scale online study. *Psychological Research*, *83*, 1444-1464.

Kochari, A. R. (2019). Conducting Web-based experiments for numerical cognition research. *Journal of Cognition*, *2*.

Krajcsi, A., & Igács, J. (2010). Processing negative numbers by transforming negatives to positive range and by sign shortcut. *European Journal of Cognitive Psychology*, *22*, 1021-1038.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1-26.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519-1520.

Peirce, J., Gray, J.R., Simpson, S. *et al.* PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods,* 51, 195–203 (2019).

Piaget, J. (1971). The theory of stages in cognitive development.

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

Varma, S., & Schwartz, D. L. (2011). The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts. *Cognition*, *121*, 363-385.

Varma, S., Blair, K. P. & Schwartz, D. L. (2019). Cognitive science foundations of integer understanding and instruction. In A. Norton & M. W. Alibali (Eds.), *Constructing number: Merging perspectives from psychology and mathematics education* (pp. 307-327). Springer.