Differential processing of “small” and “large” multi-digit numbers

Mariya Lozin and Michal Pinhas

Department of Behavioral Sciences, Ariel University, Ariel, Israel, 40700

Author Notes

This study was carried out by the first author under the supervision of the second author in partial fulfillment of the PhD requirements at Ariel University.

We thank Shira Chana Bienstock for valuable feedback and scientific editing, Ami Feder and Estie Arram Feder for their helpful comments on an earlier draft of this manuscript, and Yair Graithzer, Malachi Zivan, Sivan Cohen, Efrat Korakin, Harel Pshitizky, and Mevaseret Halevy for assisting with the data collection.

This research was supported by the Israel Science Foundation (grant no. 1348/18 to MP).

Stimuli and Data are available via the Open Science Framework at https://osf.io/u84tc/

Correspondence concerning this article should be addressed to Mariya Lozin or Michal Pinhas, Department of Psychology, Ariel University, Ariel, Israel, 40700. Email: mariyalo@ariel.ac.il or michalpi@ariel.ac.il

Abstract

Little is known about the mental representation of large multi-digit numbers that are usually beyond our personal experience. The present study explored the processing mechanisms of these numbers in a series of experiments, using the numerical comparison task. Experiment 1 included within and between-scale comparisons of multi-digit numbers varying in their left digits (e.g., 8,000,000), with one group comparing small numbers (tens, hundreds, thousands) and the other large ones (millions, billions, trillions). In Experiment 2, comparisons of small (tens, hundreds) and large (millions, billions) multi-digit numbers that varied in their left and right digits (e.g., 8,000,003) were presented in separate blocks. Experiment 3 presented small and large multi-digit numbers (from tens to trillions) that varied in their left digits in the same block. We found novel compatibility effects between the left digit and scale components, as well as between the left digit, right digit, and scale components, and extended the previously reported unit-decade compatibility effect to larger scales. We also obtained global and scale distance effects for all scales in most conditions. Importantly, both compatibility and distance effects showed context dependency in large, but not small, multi-digit numbers. Overall, these results demonstrate that small and large multi-digit numbers are processed differently. We discuss these differences and propose a processing model that accounts for them.

Keywords: large numbers, multi-digit numbers, compatibility effect, distance effect, numerical comparisons

Differential processing of "small" and "large" multi-digit numbers

Modern society requires daily use of multi-digit numbers (i.e., numbers composed of more than one digit, including fractions, decimal numbers, etc.) such as comparing prices, observing digital time displays, and calculating budgets. However, while much numerical cognition research has been dedicated to understanding how single-digit numbers are processed, studies exploring multi-digit numbers have usually been limited to two- and three-digit numbers, with only a few studies focusing on numbers larger than a million (see Nuerk et al., 2011; Nuerk et al., 2015 for reviews).

There may be inherent qualitative differences between *small* (e.g., decades, hundreds, thousands) and *large* (e.g., millions, billions, trillions) scales of multi-digit numbers – hereafter termed small and large numbers. Although frequency of exposure to different categories of numbers may depend on the specific culture (e.g., Dehaene & Mehler, 1992; Dorogovtsev et al., 2006; Krajcsi et al., 2016; Verguts et al., 2005), exposure to small numbers (tens to thousands) is usually much more common (e.g., within bills, purchases, salaries), whereas large numbers (millions to trillions) are less encountered on a daily basis (e.g., buying an apartment, hearing about governmental budgets in the news). Thus, the latter require a much higher level of abstraction than the former. Accordingly, it is possible that each type of number is processed differently. This notion is supported by the log-to-linear shift hypothesis according to which logarithmic coding (i.e., condensing large numbers into a smaller space) is replaced by linear coding as individuals gain experience with a varied set of numbers (e.g., Hurst et al., 2014; Siegler et al., 2009; Siegler & Booth, 2004; Siegler & Opfer, 2003; Thompson & Opfer, 2010). In addition, research on mathematical learning has shown that adults, even those who teach mathematics, struggle with understanding large numbers (Brass & Harkness, 2017; Gough, 2008;

Kabasakalian, 2007; Kastberg & Walker, 2008). Given that large numbers play important roles in various scientific and geopolitical contexts (e.g., Batt et al., 2008; Dorogovtsev et al., 2006; Dunning, 1997; Landy et al., 2013; Landy et al., 2017; Meffe, 1994; Resnick et al., 2017), understanding how this understudied category of numbers is processed may contribute significant value to many domains.

The symbolic Arabic number system offers a repertoire of numbers that extend beyond direct human experience. This two-dimensional system uses digit shapes from 0 to 9 as the quantity of the base dimension, while digit position represents a power of base ten as the place value dimension (Zhang & Norman, 1995). Hence, “2” refers to two units, but the digit 2 in “2,000,000” means two million. The understanding of both dimensions is crucial for managing multi-digit numbers, as well as creating a structural representation of place value on a few levels. The first level involves perceptual processing of place identification without activation of a numerical value, whereas the second level activates a spatially distributed sequence of digits representing the values associated with these positions (Nuerk et al., 2015). Furthermore, some studies have shown that place value-related information is retrieved automatically (e.g., García-Orza et al., 2017; Kallai & Tzelgov, 2012; Nuerk et al., 2015). Because of the place value structure of the Arabic number system, a number's overall length—namely, its physical size—conveys crucial information about its numerical magnitude. The longer the number is, the larger its value. Hence, detecting the length of a given number is a readily available cue for estimating its numerical value and an integral part of multi-digit number processing (Dixon, 1978; Hinrichs et al., 1982; Huber, Klein, et al., 2014). Moreover, relying on the physical size dimension of a number seems especially relevant for comparisons of numbers from different scales because, at least theoretically, such comparisons can be resolved based on the comparison of the numbers’ lengths while ignoring other factors (e.g., the exact values of the whole numbers, the numerical values of the numbers' constituents). Some models propose that the processing of a number's length occurs during the early visual analysis of multi-digit numbers via a "visual analyzer", and also involves the detection of other syntactical and lexical aspects (e.g., parsing the number into triplets, detecting 0's and their positions; e.g., Cohen & Dehaene, 1991; Dotan & Friedmann, 2018, 2019). Still, the impact of the interplay between physical and numerical size on the processing of multi-digit numbers is not well-understood and understudied.

Accumulating evidence supports the notion that rather small magnitudes (i.e., single-digit numbers) are internally represented along an ordered continuum of magnitudes (e.g., Dehaene, 1992; Dehaene et al., 1990; Gallistel & Gelman, 2000; Verguts & Fias, 2008). This "mental number line" metaphor has been suggested based on several robust effects, including the distance effect (Dehaene et al., 1990; Hinrichs et al., 1981; Moyer & Landauer, 1967) in which reaction times (RTs) are faster when comparing numbers that are further apart than closer together (e.g., 2 to 7 vs. 2 to 3) and the ratio effect of shorter latencies for number comparisons of smaller ratios (1 vs. 10) than larger ratios (5 vs. 10) (Moyer & Landauer, 1967). Besides this sensitivity to the intrapair gap, number comparisons are also affected by their overall magnitudes. This is known as the size effect (Moyer & Landauer, 1967), reflected by faster responses for comparisons of smaller than larger numbers with a fixed intra-pair distance (e.g., 3 to 4 vs. 8 to 9). A few models have been suggested to account for these effects related to coding along the mental number line. For instance, a linear model with scalar variability suggests equally spaced distributions for each numerosity along the line with increasing variability with increasing magnitude (e.g., Gallistel & Gelman, 2000), whereas a logarithmic model suggests a fixed variability for each numerosity distribution but increasing distances between the means of adjacent distributions with increasing magnitude (Dehaene, 1989; Dehaene et al., 1990). Alternatively, faster processing for smaller over larger numbers can also be accounted for by differences in number frequency, as the former are more frequent than the latter (e.g., Dehaene & Mehler, 1992; Dorogovtsev et al., 2006; Krajcsi et al., 2016; Verguts et al., 2005).

In a *holistic model* of number representation, both single- and multi-digit numbers are represented along the same mental continuum of magnitudes (e.g., Brysbaert, 1995; Dehaene et al., 1990; Reynvoet & Brysbaert, 1999). Thus, similar to single digits, multi-digit numbers are processed as integrated entities rather than individual components (e.g., “27” instead of “2 decades and 7 units,” respectively), and comparisons of multi-digit numbers are predicted to be exclusively affected by the overall distance between the numbers. An alternative, *decomposed model* (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005), suggests that each component, or constituent, of a multi-digit number is represented separately and maps onto a different mental number line. Thus, comparing a pair of two-digit numbers involves two separate comparisons (i.e., comparing the two decade digits and comparing the two unit digits); the larger the multi-digit numbers, the more componential comparisons. Componential distance effect findings, such as decade/unit digit distance effects (i.e., faster responses for more distant than closer decade/unit digits; e.g., Moeller et al., 2009; Verguts & De Moor, 2005), support this model.

The main evidence undermining a pure holistic model of number representation comes from the unit-decade compatibility effect, first described by Nuerk, Weger, and Willmes (2001). In a two-digit number comparison task, a given pair of numbers is *compatible* if *both* the decade and unit digits of one number are larger than the other (e.g., 53 vs. 68, as both $5 < 6$ and $3 < 8$), and *incompatible* if the decade/unit digit of only *one* number is smaller/larger than the other (e.g., 59 vs. 74, as 5 < 7 but 9 > 4). The unit-decade compatibility effect consists of slower RTs for incompatible than compatible pairs, reflecting separate but interactive comparisons of the overall difference between the number pair, and the magnitudes of the decade and unit digits (e.g. Nuerk et al., 2001; Nuerk et al., 2004; Nuerk & Willmes, 2005). Similarly, other compatibility effects were documented in positive and negative single-digits with polarity signs (Huber et al., 2015), physical quantities involving numbers and measurement units (Huber et al., 2015), decimal fractions (Varma & Karl, 2013), and monetary price categories (Macizo & Ojedo, 2018). Collectively, these data support the notion of a general model of componential multi-symbol number processing (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005).

Support for one model does not invalidate another, however, as evidenced by *a hybrid model* (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Verguts & De Moor, 2005; Nuerk & Willmes, 2005). Indeed, Nuerk and Willmes (2005) suggested that numbers are represented in both decomposed and holistic manners. Activating the whole magnitude and each digit, both representations can be activated in parallel or inhibit one another, depending on the task (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005). Importantly, the hybrid model can account for inconsistent findings in multi-digit number processing: data from comparing two-digit numbers to a standard number held in memory (e.g., Dehaene et al., 1990; Hinrichs et al., 1981) or presented sequentially (Ganor-Stern et al., 2009; Zhang & Wang, 2005) have been consistent with the holistic view, while comparisons of simultaneously presented multi-digit numbers have been aligned with the decomposed model (e.g., Ganor-Stern et al., 2009; Hinrichs et al., 1982; Poltrock & Schwartz, 1984; Zhang & Wang, 2005; Zhou et al., 2008).

Furthermore, decomposed and hybrid models give rise to an important architectural question concerning sequential (e.g., Korvorst & Damian, 2008; Meyerhoff et al., 2012; Poltrock & Schwartz, 1984) versus parallel (Moeller et al., 2009; Nuerk et al., 2004; Verguts & De Moor, 2005) number processing. Parallel processing assumes that digits in a multi-digit number are processed simultaneously, regardless of their place values. Accordingly, Nuerk and Willmes (2005) postulated that parallel decomposition accounts for the unit-decade compatibility effect. Conversely, sequential processing assumes each digit in a multi-digit number is processed sequentially from left-to-right (Meyerhoff et al., 2012). For example, processing the number 78 will start from the decade digit 7, and then processing of the unit digit 8 will begin. Poltrock and Schwartz (1984) compared four- and six-digit numbers varying in one digit only, located at different place value positions. They observed an increase in RTs as the leftmost unequal digit was located further to the right, consistent with sequential processing. Furthermore, Korvorst and Damian (2008) adopted the unit-decade compatibility effect to three-digit numbers and found hundred-decade and hundred-unit compatibility effects, indicating decomposed processing of units, tens, and hundreds. However, the latter effect was smaller than the former one, suggesting that units may cause less interference than tens because of a left-to-right sequential processing gradient of multi-digit numbers.

Only a few studies tested number processing in very large scales. Rips (2013) asked participants to place numbers from conventional (hundred to septendecillion) and fictitious (e.g., bazillion, kabillion, zillion) scales on a presented number line. He found that while the estimated positions of numbers from smaller conventional scales (hundred to quadrillion) were placed in correct order and distances between them, fitting a logarithmic function, conventional numbers larger than quadrillion (quintillion to septendecillion) were often incorrectly sequenced and placed similarly to fictitious items (associated with large standard errors and compressed scaling). Another study by Landy et al. (2013) asked participants to place large numbers (from 1 thousand to 1 billion) along a line. The results showed that while 1 million's placement on the line was, on average, incorrect, other numbers were placed correctly (i.e., linearly) between 1 thousand and 1 million, and between 1 million and 1 billion. Similarly, Siegler and Booth (2004) showed that participants generate logarithmic patterns with an unfamiliar range of numbers, such as 0 to 10,000,000. In addition, Landy et al. (2017) found that participants make use of numerical categories based on scale words (e.g., “thousands”, “millions”) when placing numbers across a large range, with linear coding within each category.

Also, several works from the field of mathematics learning revealed difficulties teachers encounter with when dealing with large numbers – beyond a million (Brass & Harkness, 2017; Gough, 2008; Kabasakalian, 2007; Kastberg & Walker, 2008). For instance, Kabasakalian (2007) and Brass and Harkness (2017) provided mathematics teachers with a number line estimation task in which they were asked to place numbers on a presented line flanked from zero to one trillion, and then to verbally explain their placements. In both studies, most teachers mistakenly placed one billion closer to one trillion than to zero. One of the commonly used strategies for this choice, as reported by the math teachers, was using scale number names “millions, billions, trillions” as a natural progression and the prefixes mi-, bi-, and tri- as one, two, and three, respectively (for more strategies, see Brass & Harkness, 2017). Another erroneous strategy was using benchmarks of known place values (e.g., tens, hundreds, thousands, millions, billions, and trillions), placing them in equally spaced intervals on the line. Hence, most errors were based on strategies that assigned ordinal rather than cardinal values to scale numbers.

In summary, our current understanding of multi-digit number processing is mainly based on the study of rather small numbers, some inconsistent findings, and overlapping and/or competing explanatory models. Further direct investigation of large, in comparison to small, multi-digit numbers is necessary to enable a more comprehensive picture of the mental representation of such numbers.

The present study

The present study's primary goal was to explore the processing mechanisms of large numbers by comparing between small (i.e., tens to thousands) and large (i.e., millions to trillions) multi-digit numbers. Specifically, we asked whether large multi-digit numbers are represented and processed in the same way as small multi-digit numbers. For that purpose, in a series of three experiments, participants performed a numerical comparison task in which they were presented with number pairs and were asked to choose the larger number. The numerical comparison task has been commonly used over several decades of research in the field of numerical cognition to capture various behavioral and neural effects, among them the distance and compatibility effects, and to explore the mental operations underlying (multi-digit) number processing (e.g. Ganor-Stern et al., 2009; Moeller et al., 2015; Nuerk et al., 2001; Nuerk & Willmes, 2005; Nuerk et al., 2015;).

In all experiments we evaluated the compatibility between the left digit, right digit, and scale components in both small and large multi-digit numbers in order to replicate the unit-decade compatibility effect found in two-digit numbers (Nuerk et al., 2001), and possibly extend it to larger scales and new components. To our knowledge, compatibility between the scale component and other constituents was not examined previously. Importantly, documenting compatibility effects which involve the scale component is of particular theoretical interest. This is because if found, such effects would indicate that between-scale comparisons do not rely solely on comparing the overall lengths of the numbers, even though processing this information would be sufficient to determine which of the two members of the pair is numerically larger. In Experiments 1 and 2, we further tested for both global and componential (i.e., between left digits, between right digits, and between scales) distance effects. Lastly, Experiment 3 included comparisons of small and large multi-digit numbers with proportional between-pair gaps. In other words, the disparity between number pairs was manipulated in terms of ratio instead of global distance.

We expected the results to be consistent with a hybrid number processing model, reflected by significant compatibility effects between various components of multi-digit numbers (e.g., Huber, Bahnmueller et al., 2015; Huber, Cornelsen et al., 2015; Macizo & Ojedo, 2018; Varma & Karl, 2013), as well as global (or ratio) and componential distance effects (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Verguts & De Moor, 2005; Nuerk & Willmes, 2005). Furthermore, the comparison between various number components in different scales and designs (only scale numbers, including non-scale numbers as well, different proportions of within/between-scale comparisons) would allow mapping the relative importance of each component to multi-digit number processing.

Experiment 1: Compatibility and distance effects in small and large scales

Experiment 1 evaluated compatibility effects in small and large numbers between the left digit and scale components, as well as global and componential distance (i.e., left digit distance, scale distance) effects. We included numbers from six different scales, which varied in their left digits while all other digits were zeros (e.g., 8,000; 5,000,000,000). Three scales of small numbers (tens, hundreds, thousands) appeared in one block, and three scales of large numbers (millions, billions, trillions) appeared in the other block. Four levels of compatibility between the left digit and scale components were tested in each block. In the compatible condition, both the left digit and the scale of one of the numbers in the pair were larger than the same components in the other pair (e.g., 5,000,000,000 vs. 2,000,000). In the incompatible condition, one of the components of the first number (left digit or scale) was larger than the same component of the second number, whereas the other component of the first number was smaller than the same component of the second number (e.g., 2,000,000,000 vs. 5,000,000). In the same left digit condition, the leftmost digits of both pair members were identical and the scales varied (e.g., 2,000,000 vs. 2,000,000,000). Finally, in the same scale condition, the scales of both pair members were identical and their left digits varied (e.g., 2,000,000 vs. 5,000,000).

This design allowed us to compare compatible and incompatible trials, and thus extend the unit-decade compatibility effect (e.g., Nuerk et al., 2001, 2004; Nuerk & Willmes, 2005) to new components (left digit and scale) in small and large numbers. Based on prior research, we predicted that the scale component would be processed in an ordinal, but not cardinal, manner, and therefore, the gap between adjacent small (e.g., tens and hundreds) scales and adjacent large (e.g., millions and billions) scales would be treated similarly. Thus, we expected a similar left digit scale compatibility effect for both small and large numbers (i.e., that would not be further modulated by the numerical range). Additionally, comparing compatible and same left digit trials enabled measurement of the gain in redundant activation of two compatible components (i.e., left digit and scale) versus pairs that varied only in scale (e.g., Biederman & Checkosky, 1970; Miller, 1982). Further, contrasting between-scale comparisons (i.e., compatible, incompatible, and same left digit conditions) and within-scale comparisons (i.e., same scale condition) allowed an extended exploration of the decade crossing effect of poorer performance for crossed-decade comparisons than within-decade comparisons, reported for two-digit numbers (Nuerk et al., 2011). Hence, in this experiment, it was tested as a scale crossing effect in variate scales. Both redundancy gain and scale crossing effects were expected to reveal similar effects sizes in the small and large numerical ranges if the numbers scales' would be processed in an ordinal manner.

In contrast, we assumed that the obtained distance effects would be based on both the cardinal and ordinal properties of the whole numbers and their constituents, and therefore, we expected that both global and left digit distance effects would differ between small and large numbers. In particular, small numbers are more familiar and comprised of fewer components (i.e., shorter) that are physically closer, whereas large numbers are less frequently experienced and contain more constituents, which turns them into more complex stimuli. Hence, we expected these properties to contribute to the processing of small, but not large, numbers as a whole. In accordance, we further predicted a stronger global distance effect and a weaker left digit distance effect for small than for large numbers. As for the scale distance, we expected it to be similar for both small and large numbers, based on presumed ordinal processing of the numbers' scales. Furthermore, consistent with the size effect, we expected faster RTs in small numerical range comparisons than in large (e.g., Krajcsi et al., 2016; Moyer & Landauer, 1967; Pinhas et al., 2010). Together, these predictions were expected to provide evidence supporting a hybrid model of number representation (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Nuerk & Willmes, 2005; Verguts & De Moor, 2005) for both small and large numerical ranges.

Method

Participants

Forty undergraduate psychology students (age range = 18-37 years; average age = 22.39 years; 5 males; 5 left-handed; native language: 37 Hebrew, 2 Arabic, 1 Russian) from Ariel University participated in the experiment for course credit. Participants were randomly assigned into one of two groups: small and large numbers. All participants had normal or corrected-to-normal vision, and had not been diagnosed with dyscalculia, dyslexia, or attention deficit hyperactivity disorder. Study approval was obtained from the Ariel University Institutional Review Board and all students provided written informed consent prior to participation.

Apparatus and stimuli

Stimuli were presented using E-Prime 2.0 software (released candidate 2.0.10.252; Psychological Software Tools Inc., Pittsburgh, PA) on a 22-inch monitor with a 1920 × 1080 pixel resolution. Participants responded by pressing the 'A' or 'L' keys of a standard QWERTY keyboard with their left and right index fingers, respectively.

Stimuli were generated by varying the left digit between “2,” “5,” or “8” (all other digits of the numbers were zeros) and the scale between tens, hundreds, and thousands in the small numbers group, and millions, billions, and trillions in the large numbers group (for stimuli examples, see Figure 1). Each number was paired with all other numbers and appeared once on each side of the screen, creating 144 pairs in total per group. Four types of pair comparisons were created based on the compatibility between the left digits and scales of the number pair: (a) left-digit-scale compatible, (b) left-digit-scale incompatible, (c) same left digit, and (d) same scale. Each pair appeared 9 times, creating a total of 1,296 trials (i.e., 72 pairs × 2 left/right side × 9 repetitions = 1,296) in each number group.

Numbers were presented on both sides of the screen in Ariel font, colored and framed in black on a silver background. The average distance between the centers of the two numbers was 14.9º, assuming a viewing distance of ~ 60 cm from the computer screen. Heights of all stimuli were 0.95º. The widths of numerical stimuli were approximately 1.34º for tens, 2.01º for hundreds, 2.58º for thousands, 5.72º for millions, 8.1º for billions, and 10.47º for trillions.

Procedure

Each participant sat on a height-adjustable chair in front of a monitor with index fingers on the keyboard response keys. Each trial started with a fixation cross appearing at the screen's center for 500 ms, followed by a number pair remaining visible until response, followed by a 500 ms interval of a blank screen before the next trial. Trials were randomly ordered. Participants were instructed to choose the stimulus that represented the larger numerical value by pressing the "A"/"L" key if the largest number appeared on the left/right side, respectively. They were asked to respond as quickly and as accurately as possible. Task instructions were presented in Hebrew (right-to-left reading direction), while the stimuli included only Arabic numbers (left-to-right reading direction).

The experiment began with a training session of 10 trials that included feedback at the end of each trial. During the experiment phase, there was a self-paced rest break after each of 16 successive blocks (81 trials per block). Feedback on mean accuracy and RT of correct responses was provided at the end of each block.

Results

In all experiments, for each participant, we excluded RTs shorter or longer than the range of individual means RT ± 2 *SD* (less than 1% of the data per experiment). Mean RTs of correct responses (~94% of the data) were first submitted to a two-way repeated-measures analysis of variance (ANOVA) with numerical range (small, large) as a between-participants variable and left-digit-scale compatibility (left-digit-scale compatible, left-digit-scale incompatible, same left digit, same scale) as a within-participants variable. The analysis revealed a significant main effect of numerical range, $F(1, 38) = 163.46$, $MSE = 263{,}919$, $p < .001$, $\eta^2_p = .81$, demonstrating faster responses for small (545 ms) compared to large (1065 ms) numbers. The effect of left-digit-scale compatibility, $F(3, 114) = 275.07$, $MSE = 1{,}035$, $p < .001$, $\eta^2_p = .88$, was also significant. Planned contrasts revealed significantly shorter RTs for left-digit-scale compatible (745 ms) than incompatible (786 ms) trials, $F(1, 38) = 90.17$, $MSE = 729$, $p < .001$, $\eta^2_p = .70$, as well as significantly shorter RTs for compatible than same left digit (759 ms) trials, $F(1, 38) = 21.06$, $MSE = 371$, $p < .001$, $\eta^2_p = .36$. An additional contrast indicated significantly slower RTs for same scale trials (929 ms) (i.e., within-scale comparisons) compared to all other compatibility conditions (i.e., between-scale comparisons), $F(1, 38) = 319.11$, $MSE = 30{,}844$, $p < .001$, $\eta^2_p = .89$. In addition, the two-way interaction of Numerical Range × Left-Digit-Scale Compatibility did not reach significance, $F(3, 114) = 1.98$, $MSE = 1{,}035$, $p = .120$.

Next, distance effects were evaluated separately for within- and between-scale comparisons. We used a log scale to estimate global distance, given that logarithmic coding is usually found for larger, unfamiliar number ranges (e.g., Dehaene et al., 2008; Rips, 2013; Siegler & Opfer, 2003; Thompson & Opfer, 2010). A Linear Mixed Model (LMM) analysis was conducted on mean RT with the factors of global intrapair distance (27 log-transformed distance levels), left digit distance (3, 6), and scale distance (one scale, two scales) for the between-scale comparisons analysis, and the factors of global intrapair distance (6 log-transformed distance levels) and left digit distance (3, 6) for the within-scale comparisons analysis. Importantly, no high correlations or collinearity was found between the various distance factors in each of the analyses. As seen in Table 1, LMM analyses revealed significant global distance effects for both within- and between-scale comparisons for the large, but not for the small, numerical range, demonstrating decreased RT with the increase in global intrapair distance. The left digit distance effect was significant only for within-scale comparisons in the small numerical range. Also, responses were significantly slower for distance of one (i.e., tens vs. hundreds, hundreds vs. thousands in small range; millions vs. billions, billions vs. trillions in large range) compared to two (i.e., tens vs. thousands in small range; millions vs. trillions in large range) scales in both small and large numerical ranges.

Lastly, in order to explore potential differences in the time course of the compatibility and distance effects as the experiment progressed, we conducted additional analyses dividing the RT distribution of each of the effects into equal bins. The results did not reveal evidence for a systematic influence of the bin factor on the compatibility and distance effects in either of the numerical ranges. The full description and results of these analyses are provided in the supplemental materials.

Discussion

The findings of Experiment 1 revealed a novel left-digit-scale compatibility effect, demonstrating faster responses for compatible than incompatible trials. As we expected, the effect was similar for both small and large numbers. These findings extend the previously reported unit-decade compatibility effect (e.g., Nuerk et al., 2001, 2004) in two respects. First, compatibility effects are now documented for the first time for large numbers (beyond millions). Second, the compatibility effect was demonstrated between the left digit and scale components.

Inconsistent with our prediction based on the decade crossing effect (Nuerk et al., 2011) of poorer performance for crossed-scale comparisons, responses for within-scale comparisons (i.e., same scale trials) were slower than all types of between-scale comparisons. In our study, within-scale comparisons were based only on comparison of the numbers' left digits given that all other digits were zeros. Thus, the higher proportion of between-scale comparisons might have steered participants' attention to the more relevant component: the scale (Moeller, Nuerk et al., 2009; Nuerk et al., 2011), which in same-scale trials cannot guide a decision. Also, the apparent physical length differences between two numbers from different scales, which do not exist in within-scale comparisons, presumably contributed to the shortening of RTs in between-scale comparisons (e.g., Dixon, 1978; Hinrichs et al., 1982). Previous research has consistently shown that the physical size of a number is processed automatically, even when not required by the task, and this can contribute to, or interrupt, task performance (e.g., Feder et al., 2021; Fitousi, 2014; Henik & Tzelgov, 1982). In addition, as expected, a redundancy gain was obtained for between-scale comparisons with two compatible components (left digit and scale) in contrast to those that varied in scale. Crucially, both the compatibility effect and the redundancy gain reinforce the idea that the physical size/length of the numbers was not the only information processed in between-scale comparisons, even though relying solely on this kind of information was sufficient to reach a comparative decision in these cases.

Importantly, the main difference between small and large numbers in Experiment 1 manifested in various distance effects, further modulated by the type of comparison. Inconsistent with our predictions, the left digit distance effect was evident only for within-scale comparisons of small numbers, and global distance effects were obtained for large numbers in both within- and between-scale comparisons. The left digit distance effect obtained for within-scale comparisons of small numbers was significant when global distance was controlled for, indicating that the distance between the left digits of the numbers contributed more to the decision than the global distance. On the other hand, the global distance effect obtained for within-scale comparisons of large numbers was significant when left digit distance was controlled for, indicating that processing the distance between the left digits of the numbers was not as effective in comparisons of more complicated numbers. As for between-scale comparisons of large numbers, global distance was significant when both the left digit and scale distances were controlled for (with the latter also being significant), indicating that not only the scales, but also the whole numbers, were considerable in the comparative process. Moreover, as predicted, scale distance effects were apparent for between-scale comparisons of both small and large numbers, and emerged as the most substantial contributors in the analyses, reflecting the prominence of the scale component (Moeller, Nuerk et al., 2009; Nuerk et al., 2011). The dominance of this component seem to relate to the fact that it is easily detected by noticing the differences in the physical sizes of the numbers (e.g., Dixon, 1978; Feder et al., 2021; Fitousi, 2014; Henik & Tzelgov, 1982; Hinrichs et al., 1982; Huber, Klein, et al., 2014), and its processing demonstrated greater attention to the ordinal (and not cardinal) properties of the compared numbers (Brass & Harkness, 2017; Kabasakalian, 2007).

Collectively, and as expected, the findings of Experiment 1 demonstrate that the numbers' task relevant and irrelevant constituents were processed in several ways, negating the notion of a pure holistic (e.g., Brysbaert, 1995; Dehaene et al., 1990; Reynvoet & Brysbaert, 1999) or pure decomposed (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005) representation, supporting a hybrid representation of multi-digit numbers (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Verguts & De Moor, 2005).

Lastly, it should be noted that our findings are based on a sample of non-expert participants (i.e., undergraduate psychology students) who are probably not well-experienced or familiar with the kinds of large numbers under investigation. Hence, our results may not be generalizable to a population that is more experienced with large numbers. Furthermore, almost all participants were native Hebrew speakers (37/40), a language that is characterized by right-to-left reading direction for words, which is incongruent with the left-to-right reading direction of Arabic numbers. Reading direction was previously found as a possible contributor to the effect size of the unit-decade compatibility effect (Moeller et al., 2015). Although the scale component examined here differs from the unit and decade (i.e., the number's left-digit) digit components because it is not associated with a given place or direction in the multi-digit number string, it may still be prudent to explore whether the left-digit-scale compatibility effect is enlarged among native speakers of a left-to-right reading direction language.

Experiment 2: Testing non-scale numbers

The stimuli in Experiment 1 were comprised of only one non-zero digit (i.e., multiples of tens, e.g., 20, 5,000,000). Such numbers may be processed differently due to their uniqueness and role as anchors in our number system (e.g., "decade number effect"; Nuerk et al., 2011). Accordingly, in Experiment 2 we examined the processing of small and large scales within-participants (but between-blocks), while number pairs varied in three components: the left digit, the right digit (other digits of numbers larger than tens were zeros), and the scale. In order to constrict the length of the experiment, we included two small scales (tens and hundreds) and two large scales (millions and billions).

The design of Experiment 2 allowed us to examine two types of compatibility effects: *scale-left-right digit compatibility effect* (i.e., compatibility between three different constituents, compatibility between the scale and one of the non-zero digits with an incompatible relation to the other non-zero digit, and scale incompatibility with both of the non-zero digits) for between-scales comparisons and *left-right digit compatibility effect* for within-scale comparisons. The latter left-right digit compatibility effect replicates the unit-decade compatibility effect found for two-digit numbers for larger scales. Half of the trials in the experiment were between-scale comparisons and the other half were within-scale comparisons.

Our hypothesis predicted that compatible trials in both types of within- and between-scale comparisons would be responded to faster than incompatible trials (Korvorst & Damian, 2008; Nuerk et al., 2001). Providing additional compatibility conditions (detailed below) would help us better understand the influence of the three competitive components on multi-digit number processing by comparing the redundancy gains and the prominence hierarchy between the scale, left, and right digit components. This would be accomplished by comparing compatibility effects and redundancy gains of all components in both small and large number ranges. Hinging on the findings of Experiment 1, the scale component was expected to primarily dominate between-scale comparisons because it conveyed the most important mathematical information—a crude evaluation of the whole number values, instantiated by the differences in their lengths—compared to the left and right digits of the number. Then, in both between- and within-scale comparisons, and based on same idea of mathematical relevance, the left digit was expected to dominate the comparisons more than the right digit, which was expected to have the least amount of influence on comparative decision-making (Moeller et al., 2009; Nuerk et al., 2011). Finally, consistent with the size effect (e.g., Krajcsi et al., 2016; Moyer & Landauer, 1967; Pinhas et al., 2010), small numbers were expected to be responded to faster than large numbers.

Method

Participants

Twenty undergraduate psychology students (age range = 20-25 years; average age = 21.9 years old; 1 male; all right-handed; all native Hebrew speakers) from Ariel University participated in the experiment for course credit.

Apparatus and stimuli

The set of stimuli was generated by varying the left and right digits of the number (other digits of numbers larger than tens were zeros). Four different scales were used, creating two levels of numerical range: "tens and hundreds" and "millions and billions." Selection of numbers to the stimuli set was driven by several considerations: (a) left and right digits of the number were not identical or sequential; (2) each of the two non-zero digits within the pair was unique (e.g., 104 vs. 58); and (3) for each number, one of the digits (left or right) was odd and the other was even. Based on these considerations, twelve couplings of left-right digits were created, and divided into within-scale comparisons (e.g., 23 vs. 58) and between-scales comparisons (e.g., 203 vs. 58). This created four *scale-left-right digit compatibility* conditions for between-scale comparisons: (a) left-right-digit-scale compatible (i.e., when the scale was compatible with both the left and right digits, e.g., 41 vs. 609), (b) left-digit-scale compatible (i.e., when the left digit is compatible with the scale and incompatible with the right digit (e.g., 49 vs. 601), (c) right-digit-scale compatible (i.e., when right digit are compatible with the scale and incompatible with left digit (e.g., 61 vs. 409), and (d) both digit-scale incompatible (i.e., when both non-zero digits are incompatible with the scale, e.g., 69 vs. 401) (for stimuli examples, see Figure 2A). There were four *left-right-digit compatibility* conditions for within-scale comparisons: (a) left-right digit compatible (e.g., 14 vs. 58), (b) left-right digit incompatible (e.g., 49 vs. 61), (c) same left digit (e.g., 61 vs. 69), and (d) same right digit (e.g., 14 vs. 54) (for stimuli examples, see Figure 2B).

Between-scale number pairs with the same left or right digits were treated as fillers and excluded from all analyses (e.g., 69 vs. 409, 69 vs. 601).

In each one of the two blocked numerical ranges (i.e., tens and hundreds, millions and billions), all left-right digit couplings included four within-scale compatibility conditions, four between-scale compatibility conditions, and four between-scale filler couplings. This resulted in 768 trials in total for the experiment (4 within-scale compatibility couplings × 2 numerical ranges × 2 left/right side × 16 repetitions = 256) + (4 between-scale compatibility couplings × 2 numerical ranges × 2 left/right side × 16 repetitions = 256) + (4 filler couplings × 2 numerical ranges × 2 left/right side × 16 repetitions = 256).

Procedure

Each numerical range block started with a short training session of 10 trials. During the experimental phase, there were 8 successive sub-blocks (64 trials per sub-block). The order of numerical range blocks was counterbalanced across participants.

All other details of the method were identical to those described in Experiment 1.

Results

First, mean RTs of correct responses (~94% of the data) were divided to within- and between-scale comparisons. As expected, responses to the former were significantly slower than those to the latter comparisons (815 vs. 589 ms, respectively), $t(19) = 17.83$, $p < .001$.

Second, we evaluated compatibility effects for between-scale comparisons. A two-way repeated-measures ANOVA was conducted on mean RTs with numerical range (tens to hundreds, millions to billions) and left-right-digit-scale compatibility (left-right-digit-scale compatible, left-digit-scale compatible, right-digit-scale compatible, both digit-scale incompatible) as within-participants variables. A main effect of numerical range was non-significant, $F(1, 19) = 1.26$, $MSE = 14{,}619$, $p = .275$. The main effect of left-right-digit-scale compatibility was significant, $F(3, 57) = 55.55$, $MSE = 895$, $p < .001$, $\eta^2_p = .75$. Planned comparisons contrasted responses of both digit-scale incompatible trials (628 ms) with each one of the three compatible conditions and revealed significantly slower responses for the former in all cases: left-right-digit-scale compatible (542 ms; $F(1, 19) = 89.85$, $MSE = 3{,}282$, $p < .001$, $\eta^2_p = .83$), left-digit-scale compatible (590 ms; $F(1, 19) = 47.19$, $MSE = 1{,}255$, $p < .001$, $\eta^2_p = .71$) and right-digit-scale compatible (580 ms; $F(1, 19) = 50.46$, $MSE = 1{,}803$, $p < .001$, $\eta^2_p = .73$) trials, respectively.

Importantly, the two-way interaction of Numerical Range × Left-Right-Digit-Scale Compatibility was significant, $F(3, 57) = 6.84$, $MSE = 1{,}122$, $p = .001$, $\eta^2_p = .27$ (see Figure 3). Planned follow-up analyses included the testing of three interaction contrasts, each comparing whether the RT difference between two specific levels of the left-right-digit-scale compatibility factor (i.e., left-right-digit-scale compatible vs. both digit-scale incompatible; left-digit-scale compatible vs. both digit-scale incompatible; right-digit-scale compatible vs. both digit-scale incompatible) significantly differed between the two numerical ranges (i.e., tens to hundreds vs. millions to billions). Two of these interaction contrasts did not reveal significant differences in the size of the compatibility effects between the two numerical ranges: the left-right-digit-scale compatibility effect (left-right-digit-scale compatible vs. both digit-scale incompatible trials, $F(1, 19) = 3.83$, $MSE = 2{,}756$, $p = .065$) and the right-digit-scale compatibility effect (right-digit-scale compatible vs. both digit-scale incompatible, $F(1, 19) = 0.57$, $MSE = 2{,}173$, $p = .460$). However, the third interaction contrast revealed a significantly smaller left-digit-scale compatibility effect (left-digit-scale compatible vs. both digit-scale incompatible) in the small (tens to thousands), than in the large (millions to billions), range, $F(1, 19) = 5.44$, $MSE = 2{,}756$,

p = .031, η^2_p = .22. Because this interaction contrast was significant, we continued to examine whether the left-digit-scale compatibility effect was significant at each of the numerical ranges by computing simple contrasts. We found a significant left-digit-scale compatibility effect in the small, $F(1, 19) = 4.11$, $MSE = 1{,}780$, $p = .057$, $\eta^2_p = .18$), and large, $F(1, 19) = 29.99$, $MSE = 2{,}231$, $p < .001$, $\eta^2_p = .61$, numerical ranges.

Third, we evaluated compatibility effects for within-scale comparisons. A two-way repeated-measures ANOVA was conducted on mean RTs with numerical range (tens to hundreds, millions to billions) and left-right-digit within-scale compatibility (compatible, incompatible, same left digit, same right digit) as within-participants variables. The main effect of numerical range was non-significant, $F(1, 19) = 0.14$, $MSE = 24{,}554$, $p = .715$. The main effect of left-right-digit within-scale compatibility was significant, $F(3, 57) = 60.20$, $MSE = 2{,}180$, $p < .001$, $\eta^2_p = .76$. Responses to compatible trials (688 ms) were compared with each one of the three other compatibility conditions, and revealed significantly faster responses in all cases: incompatible (770 ms; $F(1, 19) = 125.15$, $MSE = 1{,}094$, $p < .001$, $\eta^2_p = .87$), same left digit (810 ms; $F(1, 19) = 111.60$, $MSE = 2{,}667$, $p < .001$, $\eta^2_p = .86$), and same right digit (703 ms; $F(1, 19) = 5.09$, $MSE = 936$, $p = .036$, $\eta^2_p = .21$) trials, respectively.

Additionally, the two-way interaction, Numerical Range × Left-Right-Digit Within-Scale Compatibility was significant, $F(3, 57) = 29.02$, $MSE = 1{,}389$, $p < .001$, $\eta^2_p = .60$ (see Figure 4). Planned follow-up analyses included the testing of three interaction contrasts, each comparing whether the RT difference between two specific levels of the left-right-digit within-scale compatibility factor (compatible, incompatible, same left digit, same right digit) significantly differed between the two numerical ranges (i.e., tens to hundreds vs. millions to billions). The interaction contrast testing the differences between the two numerical ranges for compatible versus same left digit comparisons did not reveal significant differences, $F(1, 19) = 0.01$, MSE = 7,308, $p = .929$. The other remaining interaction contrasts were significant. The second interaction contrast, testing the compatibility effect (compatible vs. incompatible), was significantly larger in the small, than in the large, numerical range, $F(1, 19) = 73.30$, MSE = 2,368, $p < .001$, $\eta^2_p = .79$. Additional computations of the simple contrasts of the compatibility (compatible vs. incompatible) at each numerical range indicated a significant compatibility effect in the small, $F(1, 19) = 191.80$, MSE = 2,303, $p < .001$, $\eta^2_p = .91$, but not in the large, $F(1, 19) = 2.52$, MSE = 2,253, $p = .129$, range. The third interaction contrast, testing the difference between compatible and same right digit trials was significantly larger in the small, than in the large, numerical range, $F(1, 19) = 5.09$, MSE = 1,872, $p = .035$, $\eta^2_p = .21$. Additional computations of the simple contrasts of compatible versus same right digit trials revealed a significant effect in the small, $F(1, 19) = 10.53$, MSE = 1,285, $p = .004$, $\eta^2_p = .36$, but not in the large, $F(1, 19) = 0.21$, MSE = 2,256, $p = .651$, range.

Lastly, distance effects were evaluated separately for within- and between-scale comparisons, and separately for small and large numerical ranges. LMM analyses included three types of distances as independent variables: global (30 log-transformed distances), left digit, and right digit. As seen in the Table 2, LMM analyses revealed significant global distance effects in three conditions (i.e., within- and between-scale comparisons in the small numerical range and within-scale comparisons in the large numerical range), consisting of decreased RTs with increased global intrapair distance. In the same way, the left digit distance effect was significant for the small number range in both within- and between-scale comparisons, and for the large number range only for within-scale comparisons. Also, the right digit distance effect was significant only for within-scale comparisons of the small number range.

Similar to Experiment 1, we conducted additional analyses dividing the RT distributions of the compatibility and distance effects into equal bins. As detailed in the supplemental materials, these analyses did not reveal systematic effects of the bin factor for either the compatibility or distance effects in both numerical ranges.

Discussion

The results of Experiment 2 indicating a left-right-digit compatibility effect replicate the classic unit-decade compatibility effect found for two-digit numbers (Nuerk et al., 2001), as well as the hundred-unit compatibility effect obtained for three-digit numbers (Korvorst & Damian, 2008), and further expand these effects to larger scales (millions, billions). We also reproduced the compatibility effect found in Experiment 1 between the left digit and scale component with the inclusion of the right digit component, providing evidence for a novel three-component (left digit, right digit, and scale) compatibility effect in both small and large numbers.

The novel triple left-right-digit-scale compatibility effect (i.e., left-right-digit-scale compatible versus both digit-scale incompatible; left-digit-scale compatible versus both digit-scale incompatible; right-digit-scale compatible versus both digit-scale incompatible) found for between-scale comparisons was significant across both number ranges. Importantly, however, the RT difference between left-digit-scale compatible versus both digit-scale incompatible trials was significantly smaller in the small (e.g., 41_609 vs. 69_401), than in the large (e.g., 4,001_6,000,009 vs. 6,009_4,000,001), number range. In the left-digit-scale compatible condition, the left digit is compatible with the scale, but the right digit is incompatible with both the scale and the left digit; hence, if the left digit promotes the decision, the right digit interferes and distracts this process. Thus, a smaller difference between left-digit-scale compatible and both digit-scale incompatible trials reflets a smaller impact of the left digit and a greater interference of the right digit. In turn, the processing of small, rather than large, numbers seems to be more sensitive to the distraction of an incompatible right digit probably because smaller numbers are comprised of fewer constituents that are grouped closer together. This interpretation is consistent with a recent study by Schäfer and Frings (2021), who found that increasing the spatial separation between stimulus features (target and non-relevant flanking stimulus) decreases the size of the distraction of non-relevant flanking stimulus (i.e., an interference effect, see also e.g., Eriksen & Eriksen, 1974; Fox, 1998; Hommel, 1995; Miller, 1991).

As for within-scale comparisons, both the compatibility effect and redundancy gain were found for the left and right digits in small numbers. As for large numbers, compatible left-right-digit trials were responded to faster than same left digit trials, but not same right digit and incompatible trials, indicating a redundancy gain only for the left digit. These findings conform to the notion of differential mathematical importance for each of the numbers' constituents when reaching a decision in number comparisons, and demonstrate the priority given to the left digit over right digit (Moeller et al., 2009, Nuerk et al., 2011) in this case. Additionally, the distance effect analyses showed that the right digit component contributed to RT predictions only in within-scale comparisons of the small number range. Global and left digit distance effects emerged for between-scale comparisons of the small number range and within-scale comparisons of both the small and large number ranges. None of the included distance factors were obtained for between-scale comparisons of large numbers. Furthermore, overall RTs to the small and large number ranges were similar, in contrast to what was expected from the size effect (e.g., Krajcsi et al., 2016; Moyer & Landauer, 1967; Pinhas et al., 2010).

In conjunction, these results from both between- and within-scale comparisons seem to reveal a componential hierarchy related to the information about magnitude conveyed by each component. Hence, the scale, which is reflected in differences in the numbers' physical sizes, appears to be the most prominent component, followed by the left digit, and then the right digit. These findings, together with the absence of a right digit distance effect (except for within-scale comparisons in the small number range), possibly imply a strategy that reduces attention to the most irrelevant component in large, but not in small, numbers (Moeller, Nuerk, et al., 2009; Nuerk et al., 2011). On the other hand, the observed comparability effects for between-scale comparisons indicate that although the scale was the most prominent component and its processing was sufficient to meet the task requirements, other irrelevant components were processed as well. Notably, in Experiment 2, only two scales were included for each numerical range, negating the evaluation of scale distance as a factor with the current design.

Collectively, our findings seem consistent with the idea of hybrid processing (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Nuerk & Willmes, 2005; Verguts & De Moor, 2005) for both large and small multi-digit numbers. These numbers seem to be represented in both holistic and decomposed manners, which involve the activation of the whole magnitude of the number, as well as the magnitudes of each of the number's various components, respectively.

Experiment 3: Comparing small and large scales with the same ratio

Experiment 3 examined comparisons of numbers ranging from tens to trillions with proportional between-pair gaps. In other words, we manipulated the disparity between number pairs in terms of ratio instead of global distance in order to contrast comparisons from different scales using the same numerical ratios. Consequently, all scales were included within the same experimental block and each comparison could have been from one of three ranges: one of small numbers (tens to thousands) and two of large numbers (tens of millions to billions, and tens of billions to trillions). All number pairs varied in their scales and left digits (all other digits were zeros), allowing examination of the left-digit-scale compatibility effect in the same manner described in Experiment 1, excluding same-scale comparisons. Based on past research and the results of our previous experiments, we expected similar left-digit-scale compatibility and ratio effects in all numerical ranges. Also, we expected to replicate the findings of a scale distance effect with no left digit distance effect obtained for between-scale comparisons in Experiment 1. Finally, we expected Experiment 3's findings to be consistent with the hybrid model (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Nuerk & Willmes, 2005; Verguts & De Moor, 2005) with longer RTs for larger than smaller number ranges, consistent with the size effect (e.g., Krajcsi et al., 2016; Moyer & Landauer, 1967; Pinhas et al., 2010).

Method

Participants

Twenty undergraduate psychology students (age range = 20-25 years; average age = 21.85 years; 2 male; 2 left-handed; all native Hebrew speakers) from Ariel University participated in the experiment for course credit.

Apparatus and stimuli

Stimuli consisted of numbers from six different scales (tens, hundreds, thousands, millions, billions, trillions) divided into three levels of numerical range: "tens to thousands," "millions to billions," and "billions to trillions." Hence, there was one range of small numbers and two ranges of large numbers. Numbers' left digits varied between "2," "5," and "8," and all other digits were zeros, producing 27 number pairs per range (for stimuli examples, see Figure 5). Based on the intrapair ratio gaps for the "tens to thousands" pairs, pairs in the two other ranges were created (e.g., "20 vs. 5,000" and "20,000,000 vs. 5,000,000,000" having the same intrapair ratio). Each number appeared once on the right and once on the left, and each pair was repeated 8 times, creating a total of 1,296 trials (i.e., 27 pairs × 3 ranges × 2 left/right side × 8 repetitions = 1,296).

Procedure

During the experiment, there were 18 successive sub-blocks of 72 trials each. All other methodology was identical to Experiment 1.

Results

First, mean RTs of correct responses (about 96% of the data) were submitted to a two-way repeated-measures ANOVA with numerical range (tens to thousands, millions to billions, billions to trillions) and left-digit-scale compatibility (left-digit-scale compatible, left-digit-scale incompatible, same left digit) as within-participants variables. A main effect for numerical range, $F(2, 38) = 51.94$, $MSE = 71{,}619$, $p < .001$, $\eta^2_p = .73$, demonstrated faster responses for the "tens to thousands" range (559 ms) than the two other large numbers ranges, $F(1,19) = 55.55$, $MSE = 2{,}410{,}380$, $p < .001$, $\eta^2_p = .75$, with no significant differences between the "millions to billions" (990 ms) and "billions to trillions" (990 ms) ranges, $F(1,19) = 0.000009$, $MSE = 55{,}978$, $p = .998$. In contrast to Experiment 1, no significant main effect was found for left-digit-scale compatibility, $F(2, 38) = 1.4$, $MSE = 1{,}925$, $p =.259$. Importantly, the interaction between numerical range and left-digit-scale compatibility was significant, $F(4, 76) = 2.99$, $MSE = 1{,}373$, $p = .024$, $\eta^2_p = .14$ (see Figure 6). Planned comparisons revealed a significant compatibility effect (compatible vs. incompatible) for the "tens to thousands" range, $F(1, 19) = 54.32$, $MSE = 310$, $p = .001$, $\eta^2_p = .74$, and faster responses for compatible comparisons than same left digit comparisons, $F(1, 19) = 10.26$, $MSE = 139$, $p = .005$, $\eta^2_p = .35$. Identical planned comparisons conducted for the "millions to billions" and "billions to trillions" ranges did not reveal significant results (i.e., no compatibility effect was found for both ranges: $F(1, 19) = 1.25$, $MSE = 1{,}918$, p

= .277; *F*(1, 19) = 2.22, *MSE* = 5,688, *p* = .153; additionally no significant differences were found between compatible and same left digit for both ranges: *F*(1, 19) = 3.37, *MSE* = 2,697, *p* = .082; *F*(1, 19) = 1.46, *MSE* = 8,544, *p* = .242, respectively).

Next, ratio and componential distance effects were evaluated separately for each one of the numerical ranges. LMM analyses on mean RTs included ratio as an independent variable as well as two types of componential distances as independent variables: left digit and scale. As seen in Table 3, LMM analyses revealed a significant ratio effect in the "tens to thousands" range, demonstrating faster responses as the ratio between the compared numbers decreased. No significant ratio effects were obtained for the larger ranges. Additionally, the scale distance effect was significant for all three ranges, with slower responses for one (935 ms) than two scale distances (763 ms). Finally, the left digit distance effect was not significant in all ranges.

Similar to the previous experiments, we conducted additional analyses dividing the RT distributions of the compatibility and distance effects into equal bins. These analyses did not reveal any systematic influence of the bin factor on the compatibility and distance effects in either of the number ranges (see supplemental materials).

Discussion

Experiment 3 revealed a significant left-digit-scale compatibility effect, redundancy gain, and ratio effect only for small range (i.e., "tens to thousands"), but not large ranges (i.e., "millions to billions," "billions to trillions"), in contrast to our predictions. Only the scale distance effect was significant in all ranges. The fact that no indication for processing the left digit was obtained in the larger scales may be explained by the lack of within-scale comparisons as fillers in the experiment (Nuerk et al., 2011). This, in turn, possibly led participants to ignore the irrelevant component (left digit) in 100% of the trials, consistent with cognitive control theory, which posits the proportion of fillers in a given experimental design can modulate the effects (Huber, Mann et al., 2014, 2016; Macizo & Herrera, 2013). Moreover, the lack of a left-digit-scale compatibility effect, as well as left digit and global distance effects, for large numbers presumably implies that participants relied on the processing of the scale component alone—instantiated by differences in the physical sizes of the numbers. In contrast, in the smallest "tens to thousands" range, the left-digit-scale compatibility effect was present, indicating automatic processing of the irrelevant left digit component, despite the absence of same-scale fillers. Hence, the left-digit-scale compatibility effect and redundancy gain obtained in small numbers indicate the participants' comparative decisions were not only based on the processing of the physical sizes of the numbers. Thus, these results highlight the difference in processing small and large multi-digit numbers when they appear together in the same block and cannot be solely explained by cognitive control theory (Huber, Mann et al., 2014; Macizo & Herrera, 2011, 2013; Moeller et al., 2013).

Such findings are consistent with the notion of a hybrid model (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005) for small multi-digit numbers, and only decomposed representation of large multi-digit numbers (e.g., García-Orza, & Damas, 2011; Huber et al., 2016; Meyerhoff et al., 2012; Moeller et al., 2011; Nuerk & Willmes, 2005; Verguts & De Moor, 2005).

General Discussion

In this study, we explored the processing mechanisms of large multi-digit number stimuli (from millions to trillions) by comparing them to small multi-digit numbers such as tens, hundreds, and thousands. To do so, we conducted a series of three experiments using the numerical comparison task. In the first experiment, we examined the compatibility between left digit and scale components in small and large numbers (from tens to trillions). The results revealed a novel left-digit-scale compatibility effect. In the second experiment, we included small and large scales, manipulating three components: the left digit, the right digit, and the scale. We reported a novel left-right-digit-scale compatibility effect for small and large numbers. Together with the findings of Experiment 1, the results of Experiment 2 replicate and extend the unit-decade compatibility effect (Nuerk et al., 2001) for large numbers, as well as to new components (i.e., scale). Referring to the unit and decade digits in a two-digit number as the left and right digits of any multi-digit number enabled us to compare between compatibility effects in larger numbers (beyond tens) using the same analysis. In the third experiment, number pairs with proportional intrapair gaps from small and large scales were compared within the same block. The results demonstrated different processing for small and large numbers. Specifically, the findings of Experiment 3 replicated the left-digit-scale compatibility effect, as well as global and scale distance effects for the small scales, whereas only the scale distance effect was observed for the large scales. In addition, RT distribution analyses of the compatibility and distance effects in all experiments revealed an overall stability in the effects sizes as the experiment progressed, inconsistent with either practice or fatigue effects. A summary of the obtained compatibility and distance effects by numerical range and type of comparison per experiment is presented in Tables 4 and 5, respectively.

Are small and large multi-digit numbers treated similarly?

One of the main goals of the current research was to explore whether small and large multi-digit numbers are processed in the same way. Our data from three experiments showed that not all effects were obtained for both small and large scales. In Experiment 1, the difference between ranges is expressed in distance effects, with a significant global distance effect in large, but not in small, number ranges, and with a left digit distance effect only for within-scale comparisons in small number ranges. In Experiment 2, the differences between small and large number ranges were expressed in both within- and between-scale comparisons, for compatibility as well as distance effects. This difference was accompanied by similar latencies for both ranges, and may indicate different underlying processing of the left and right digit components during the same time period (Algom et al., 2015). In Experiment 3, both global distance and left-digit-scale compatibility effects were obtained for the small number range, but not for the two larger ranges.

Only the scale distance effect was similar in all number ranges in Experiments 1 and 3. Moreover, in Experiment 2, overall latencies to small and large between-scale comparisons were similar, inconsistent with the expected size effect (e.g., Krajcsi et al., 2016; Moyer & Landauer, 1967; Pinhas et al., 2010) that predicts longer latencies with increased number magnitudes. Hence, it seems that large scales like millions, billions, and trillions are perceived as being closer to each other than they really are, presumably in line with an internal logarithmic representation (Dehaene, 1989; Dehaene et al., 1990) or even ordinality of scales (or benchmarks of known place value strategy; Brass & Harkness, 2017; Kabasakalian, 2007). In turn, these results imply that large multi-digit numbers are not fully understood, as we do not have an accurate appreciation of their quantity.

Similarly, our findings demonstrate different redundancy gain patterns for small and large numbers. Miller (1982) showed that identical signals presented on different channels contribute to a common activation and lead to faster responses. These results of redundant signal co-activation were related to finding faster responses when stimuli differed in two dimensions than only in one. This pattern occurred even when the decision could have been made based on one dimension alone (Biederman & Checkosky, 1970). We found that compatible trials were faster than same left digit trials for the small number range only in Experiment 3, and faster than same right digit trials for the small number range only in Experiment 2. For large number ranges, no stronger or unique redundancy gain occurred. The redundancy gain obtained when contrasting compatible (i.e., left-digit-scale compatible in Experiment 1, and left-right-digit within-scale compatible in Experiment 2) versus same left digit trials was similar for both small and large numbers.

Can observed differences between small and large numbers be explained by experimental design?

Experiment 1 included small and large scales in different groups, Experiment 2 included small and large scales in different blocks, and Experiment 3 included both scales in the same block. Our experimental designs also differed in their within- versus between-scale comparison proportions (within-scale comparisons were 25%, 50%, and 0% of the trials in Experiment 1-3, respectively). Further, the number of scales presented within blocks varied between experiments (3, 2, and 6 scales for Experiment 1-3, respectively). Based on these experimental design differences, it may be argued that the absence of compatibility or distance/ratio effects for large scales may be accounted for by cognitive control theory (e.g., Huber, Mann et al., 2014; Macizo & Herrera, 2011, 2013; Moeller et al., 2013), according to which diverse strategies are adopted for different task conditions. The cognitive control account postulates that participants dynamically modulate and adjust their behavior during experimental tasks according to specific conditions and requirements (e.g., Botvinik et al., 2001). Cognitive control was reflected in both compatibility (Huber, Mann et al., 2014; Macizo & Herrera, 2011, 2013) and distance (Huber et al., 2016; Moeller et al., 2013) effects. For instance, the unit-decade compatibility effect became larger when the proportion of within-decade filler items increased (Macizo & Herrera, 2011). Also, larger compatibility effects were found for larger proportions of filler (same decade or same unit) to critical items and smaller proportions of incompatible to compatible items (Huber, Mann et al., 2014; Macizo & Herrera, 2013). Similarly, distance effects for hundreds in three-digit number processing were modulated by the number of filler items (Huber et al., 2016). Ergo, these studies suggest that cognitive control is activated more in conflicting and competitive conditions, like incompatible trials, and reconfigures the system to adapt to specific test proportions and conditions.

Based on the proportions of within-scale comparisons, as well as previous findings (e.g., Macizo & Herrera, 2011), we should have expected absence or even reversed compatibility effects in Experiments 1 and 3 due to the small (or lack of) percentage of within-scale fillers. Contrary to these expectations, left-digit-scale compatibility effects emerged for small number ranges in Experiment 3, and for both small and large number ranges in Experiment 1. This implies that participants' attention was drawn to the irrelevant component of the left digit in between-scale comparisons. Moreover, in Experiment 2, which included the largest percent of within-scale trials and neutral fillers, no left-right-digit within-scale compatibility effect was found for large numbers. Thus, when small and large numbers appeared together in the same block (Experiment 3), both compatibility and ratio effects were found for small, but not large, numbers. Such results indicate that ignoring the irrelevant left digit component was not the chosen strategy for any of the number ranges in the same block. They further imply that the processing of the scale component alone (in other words, relying solely on differences in the physical sizes of the numbers) was the chosen strategy when comparing large, but not small, numbers. However, this seemed to be the case only under the unique conditions of Experiment 3.

Taken together, we can conclude that cognitive control cannot serve as a major explanation of the current results given the differences obtained for small and large numbers in all three experiments. Instead, it seems plausible that participants applied different processes or strategies in the same block for different stimuli, depending on the presented number scales. This is in accordance with what Patel and Varma (2018) call *referential processing*, in which task- and stimulus-specific *referents* are recruited to process more abstract mathematical concepts. For instance, natural numbers were recruited as referents in magnitude comparisons of irrational numbers denoted by radical expressions, whereas perfect squares were referents for the same stimuli in number line estimation and arithmetic problem-solving tasks (Patel & Varma, 2018). Similarly, Landy et al. (2013) proposed that scale words serve as referents when performing the number line estimation task across very large ranges. Therefore, our findings appear to demonstrate differential processing for small and large multi-digit numbers, possibly relying on unique componential referents for each of these number ranges.

Which processes underlie the comparisons of large multi-digit numbers?

Our results provide support for both parallel and sequential decomposition of large multi-digit numbers. From one perspective, the compatibility effects obtained in the present study can serve as evidence for parallel automatic processing of numerical constituents (e.g., Moeller, Nuerk et al., 2009; Nuerk et al., 2004; Verguts & De Moor, 2005). Comparisons based on contrasting numbers' left digits produced smaller effects than contrasting scales (instantiated by differences in the numbers' lengths). This result contradicts the idea of left-to-right sequential processing (e.g., Korvorst & Damian, 2008; Meyerhoff et al., 2012; Poltrock & Schwartz, 1984) and confirms the notion of an early visual analysis of the physical parameters of the number's length (e.g., Cohen & Dehaene, 1991; Dotan & Friedmann, 2018, 2019), as a readily available cue for determining its relative value (e.g., Dixon, 1978; Feder et al., 2021; Hinrichs et al., 1982; Huber, Klein, et al., 2014). Notwithstanding, the presented evidence of compatibility effects and redundancy gains in all experiments indicates that the processing of less relevant components, like the left and right digits, was also involved in participants' comparative decisions beyond the processing of the physical size/length of the numbers. Furthermore, the fact that comparisons involving contrasting numbers' left digits produced smaller effects than those contrasting numbers' scales aligns with previous research that postulates asymmetric componential processing of two-digit numbers. For example, Fitousi and Algom (2006) found that the decade digit is ineluctably processed, even when its processing is not relevant for reaching a decision, whereas the unit digit can successfully be ignored.

From another perspective, the effects associated with the left digit component were larger than those associated with the right digit component. Such findings may be explained by the stronger relevance of the left digit for making magnitude-related decisions, but they may also relate to the left-to-right order of the digits. Comparisons based on the right digits were responded to more slowly than all other comparisons, and this pattern was enhanced in the large scales, possibly reflecting sequential processing that is characterized by a further distance from "the starting point" in large numbers.

In addition, over the years, a dominant view in numerical cognition research assumed that number comparisons involve immediate access to the numbers' magnitudes, including placing them along a mental number line, resulting in a distance effect (e.g., Dehaene, 1997, 2001). More recent studies have provided evidence undermining this assumption. The presence of the distance effect seems to depend on the task, as no effect was found under unlimited time conditions using a modified version of the number to position task (Bar et al., 2019) or in an automatic digit-matching task (Goldfarb et al., 2011). Additionally, the distance effect disappears when no order between numerical stimuli is expected (Cohen, 2009). Moreover, the distance effect is also found in comparisons of ordered non-numerical stimuli (e.g., letter names, months of the year, military ranks; Leth-Steensen & Marley, 2000; Sasanguie et al., 2017; Verguts & Van Opstal, 2014), demonstrating that comparisons of any ordered sets may result in the effect. Therefore, it may not serve as an ideal marker for numerical processing.

This evidence, considered together with our findings, results in our suggestion that (at least) two separate independent processes, which run in parallel and race against each other, are automatically activated in number comparisons: (a) a magnitude comparison process that evaluates which number is larger/smaller, and (b) a distance estimation process that evaluates the distance(s) between numbers. It is likely that additional parallel processes may be automatically activated in number comparisons (e.g., detecting an end-value within the pair; e.g., Leth-Steensen & Marley, 2000; Pinhas & Tzelgov, 2012), and that the speed and complexity of these processes vary between stimuli and tasks.

We postulate that in single digit comparisons, both magnitude comparison and distance estimation processes are executed very quickly and it is hard to differentiate between them. Accordingly, past research may not have distinguished between these processes. Rather, in many cases it was assumed that magnitude comparison is performed by assessing the numbers' magnitudes, which are mapped along a mental number line. Such mapping involves an automatic retrieval of the numbers' intrapair distance (e.g., Dehaene, 1992; Dehaene et al., 1990; Gallistel & Gelman, 2000; Moyer & Landauer, 1967; Verguts & Fias, 2008). However, in multi-digit number comparisons, both processes run more slowly given the increased stimuli complexity, and may further involve additional sub-processes. This stimuli complexity can be expressed both in terms of multiple components that constitute each number and in the greater variety of physical characteristics of the stimuli (as occurs in between-scale comparisons). For example, when comparing “25” with “57,” one can compare the numbers’ global magnitudes (i.e., 25 is smaller than 57). Moreover, the decade and unit digit magnitudes can also be compared (i.e., 2 is smaller than 5, and 5 is smaller than 7). Similarly, both global and componential distances can be evaluated (i.e., global distance of 32, decade distance of 30, and unit distance of 2, respectively). The larger the multi-digit numbers are, the more additional sub-processes are involved, resulting in differentiated processing of single digits, and small and large multi-digit numbers.

Moreover, we conjecture that while magnitude comparison is essential for making a comparative decision and is based on the ordinality of the presented numbers, distance estimation is not essential, and is based on the cardinality of the compared numbers. Information concerning the ordinality of the presented numbers seems to be grasped in some cases from differences in the physical sizes of the numbers. Furthermore, as described earlier, many mathematics teachers do not use linear or even logarithmical knowledge when dealing with tasks that involve large multi-digit numbers, and instead rely on the ordinality of the scales in order to reach a decision (Brass & Harkness, 2017; Kabasakalian, 2007). Consistently, the scale emerged as the most prominent component in guiding the comparative decisions across all experiments in the current research.

We additionally posit that when large multi-digit numbers are compared, although both processes are activated automatically, the magnitude comparison process runs faster than the distance estimation process. Consistent with the notion of a race model (e.g., Bundesen, 1987), the race is over once one of the processes is completed and the required response can be made. Accordingly, and importantly, we suggest that the two processes race against each other according to a parallel minimum-time *stopping rule*, and that each individual process operates based on a sequential (serial) exhaustive stopping rule based on partial parameters of the Double Factorial Paradigm (e.g., Algom et al., 2015; Eidels et al., 2010; Fitousi & Algom, 2018, 2020; Townsend & Nozawa, 1995). A stopping rule refers to the quantum of required information for selecting the response. Two stopping rule alternatives are described in the literature: (a) a minimum-time stopping rule, in which the response can be made after the completion of the first process (of two or more processes), and (b) an exhaustive stopping rule, in which the response can be made after the completion of all processes (Townsend & Ashby, 1983; Townsend & Nozawa, 1995). A parallel minimum-time stopping rule refers to processing information from different channels and with different levels of relevance simultaneously; when the first crucial information for response is obtained, the process will stop. In contrast, a sequential exhaustive stopping rule refers to a process that stops only after all sub-processes are completed in a sequential manner (e.g., Algom et al., 2015; Eidels et al., 2010; Fitousi & Algom, 2018, 2020; Townsend & Nozawa, 1995).

The assumption of racing according to a parallel minimum-time stopping rule (Algom et al., 2015; Fitousi & Algom, 2018, 2020) can explain how one component's processing resulted in one effect, but did not produce another. For instance, between-scale comparisons of large numbers in Experiment 1 resulted in an advantage for compatible trials over incompatible and same left digit trials, showing that both task-relevant and irrelevant components were processed. Notwithstanding, the same comparisons produced global and scale distance effects, but not a left digit distance effect. Thus, it seems that in all three experiments, the magnitude comparison process ran faster than the distance estimation process and relied not only on the processing of scale, but also the left and right digits. Instead, the distance estimation process was slower, and stopped when a decision was obtained based on the magnitude comparison process. Simultaneously, the distance estimation process involved only the processing of scale and global distance.

Likewise, the magnitude comparison and distance estimation processes each appeared to operate using a sequential exhaustive stopping rule (Algom et al., 2015; Fitousi & Algom, 2018, 2020). Although decisions of between-scale comparisons can be based on the scale component alone, the findings of both effects indicated a continued processing of additional information, with an advantage for the magnitude comparison process over the distance estimation process. Inspecting the sub-processes of distance estimation reveals that scale distance, the most relevant component for reaching a decision, resulted in distance effects for all scales and conditions. Moreover, global distance, which contains redundant information but is also relevant for reaching a decision, resulted in distance effects in almost all conditions (except for the large number ranges in Experiment 3). Nevertheless, in Experiment 2, the left digit distance effect was found only in the small scales and within-scale comparisons of large multi-digit numbers, while the right digit distance effect was not found at all. The order of these various distances (i.e., scale, global, left digit, and right digit distances) is based on the degree of their decision making relevance. Assuming the distance estimation process ran more slowly than the magnitude comparison process can explain the stopping-point of the former in different experimental conditions, as a function of completing parallel magnitude comparison processing.

The only exception was Experiment 3, in which presentation of a wide range of scales in the same block led participants to change their strategy for large number comparisons, basing their decisions on scale alone, thereby relying on differences in the physical size of the numbers. The absence of left-digit-scale compatibility, as well as ratio and left digit distance effects, only among the comparisons of large multi-digit numbers further indicates that, in this context, the magnitude comparison process operated according to a minimum-time stopping rule, not an exhaustive stopping rule, as in all other conditions (Algom et al., 2015; Fitousi & Algom, 2018, 2020).

Future directions

Our novel findings of differential processing for small and large numbers lay a foundation for the theoretical postulation that at least two processes underlie multi-digit number processing. Further research can deepen, extend, and substantiate our theoretical postulation in a few ways. One possible line of research could focus on the influence of cognitive control on parallel and sequential strategies of race models, and attempt to validate our proposed explanations. This could be done, for instance, by directly contrasting three blocked ranges of number comparisons: small multi-digit numbers, large multi-digit numbers, and both small and large multi-digit numbers.

Additional future research could extend the current findings by employing other tasks. In the current study, we wished to examine new compatibility effects, as well as various kinds of distance effects, in multi-digit number processing, and therefore, we chose to use the numerical comparison task, commonly used in numerical cognition research for scrutinizing such effects. The use of other tasks may replicate the effects reported here under another paradigm and/or reveal more effects characterizing large number processing. For instance, in the same-different task (e.g., Dehaene & Akhavein, 1995), “different” trials would be equivalent to the “left” and “right” responses in the numerical comparison task, and “same” trials would be considered as a novel control condition. Such a task would allow examination of compatibility and distance effects for “different” trials, and comparison of the contribution of each component of interest in contrast to "same" relevant trials. Another possibility would be use of the number-line estimation task (e.g., Landy et al., 2013, 2017) or the two-numbers-to-two-positions task (Bar et al., 2019) that would enable exploring the dissociation between the magnitude comparison and distance estimation processes that appears to characterize large number processing.

The sample in the present study was limited in two manners: participants' native language right-to-left reading direction and the fact that they were non-experts in large multi-digit numbers. Future studies may wish to examine the influence of reading direction on compatibility effects involving the left digit, right digit, and scale components in large numbers. Moeller et al. (2015) found that the unit-decade compatibility effect interacts with reading directions (left-to-right vs. right-to-left) and number word inversions (non-inverted vs. inverted; e.g., forty-seven vs. seven-and-forty, respectively). As the present data is based on a cohort whose vast majority was native Hebrew speakers (77 out of 80 participants across experiments), it might be enlightening to explore compatibility with varied components of multi-digit numbers in languages with a left-to-right reading direction. Finally, it would be of great interest to examine the processing of large multi-digit numbers among a group of expert participants, namely those who are more familiar with these numbers in their daily life (e.g., given their profession or culture). Among such an expert group, the difference in exposure to small and large multi-digit numbers would be expected to be smaller compared to the non-experts participating in this study, which in turn may result in different processing patterns for large numbers.

Conclusion

The present study expanded the field's knowledge of the processing of large multi-digit numbers, which are less frequently occurring and less directly experienced in daily life. We extended the unit-decade compatibility effect to larger scales (up to trillions), and documented novel compatibility effects with other relevant components, such as scale. Our findings revealed that small, as well as large, multi-digit numbers are processed according to a componential hierarchy in which the scale is rated as the most important component. Moreover, our results provided evidence for both holistic and componential processing, consistent with a hybrid representation of multi-digit numbers, utilizing both parallel and sequential processes. Additionally, our results support the dissociation between magnitude comparison and distance estimation processes in large multi-digit numbers. This dissociation probably also occurs when processing small multi-digit numbers, but given the shorter latencies in comparisons of small scales, it is more difficult to distinguish between them. Finally, our data are consistent with the idea of logarithmic coding for large multi-digit numbers, especially when small and large multi-digit numbers appear next to each other. Under such conditions, it seems that the difference between 2 and 5 million is much vaguer than the difference between 2 and 5 decades, which is processed more deeply.

References

Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of response times: Identification of cognitive mechanisms through mathematical modeling. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology* (pp. 63–98). Oxford University Press.

Bar, H., Fischer, M. H., & Algom, D. (2019). On the linear representation of numbers: Evidence from a new two-numbers-to-two positions task. *Psychological Research, 83*, 48–63. https://doi.org/10.1007/s00426-018-1063-y

Batt, C. A., Waldron, A. M., & Broadwater, N. (2008). Numbers, scale and symbols: The public understanding of nanotechnology. *Journal of Nanoparticle Research, 10*, 1141–1148. https://doi.org/10.1007%2Fs11051-007-9344-1

Biederman, I., & Checkosky, S. F. (1970). Processing redundant information. *Journal of Experimental Psychology*, *83*, 486–490. https://doi.org/10.1037/h0028841

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652. https://doi.org/10.1037/0033-295X.108.3.624

Brass, A., & Harkness, S. S. (2017). Pre-service teachers' conceptions of the magnitude of large numbers. *Investigations in Mathematics Learning*, *9*, 53–68. https://doi.org/10.1103/PhysRevSTPER.8.010111

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

Bundesen, C. (1987). Visual attention: Race models for selection from multielement displays. *Psychological Research*, *49*, 113–121. https://doi.org/10.1007/BF00308676

Cohen, D. J. (2009). Integers do not automatically activate their quantity representation. *Psychonomic bulletin & review*, *16*(2), 332–336. https://doi.org/10.3758/PBR.16.2.332

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception and Psychophysics*, *45*, 557–566. https://doi.org/10.3758/BF03208063

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (1997). *The number sense*. Oxford University Press. https://doi.org/10.2307/2589308

Dehaene, S. (2001). Précis of the number sense. *Mind & language*, *16*, 16–36. https://doi.org/10.1111/1468-0017.00154

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 314–326. https://doi.org/10.1037/0278-7393.21.2.314

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

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, *320*, 1217–1220. https://doi.org/10.1126/science.1156540

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Dixon, P. (1978). Numerical comparison processes. *Memory & Cognition*, *6*, 454–461. https://doi.org/10.3758/BF03197479

Dorogovtsev, S. N., Mendes, J. F. F., & Oliveira, J. G. (2006). Frequency of occurrence of numbers in the World Wide Web. *Physica A: Statistical Mechanics and its Applications, 360*, 548–556. https://doi.org/10.1016/j.physa.2005.06.064

Dotan, D., & Friedmann, N. (2018). A cognitive model for multidigit number reading: Inferences from individuals with selective impairments. *Cortex, 101*, 249–281. https://doi.org/10.1016/j.cortex.2017.10.025

Dotan, D., & Friedmann, N. (2019). Separate mechanisms for number reading and word reading: Evidence from selective impairments. *Cortex*, *114*, 176–192. https://doi.org/10.1016/j.cortex.2018.05.010

Dunning, J. B. (1997). The missing awareness, part 2: teaching students what a billion people looks like. *Conservation Biology*, *11*, 6–10. https://doi.org/10.1046/j.1523-1739.1997.96406.x

Eidels, A., Townsend, J. T., & Algom, D. (2010). Comparing perception of Stroop stimuli in focused versus divided attention paradigms: Evidence for dramatic processing differences. *Cognition*, *114*, 129–150. https://doi.org/10.1016/j.cognition.2009.08.008

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, *16*(650), 143–149. https://doi.org/10.3989/arbor. 2000.i650.965

Feder, A., Lozin, M. & Pinhas, M. (2021). No power: exponential expressions are not processed automatically as such. *Psychological Research*, 85, 2079–2097. https://doi.org/10.1007/s00426-020-01381-6

Fitousi, D. (2014). On the internal representation of numerical magnitude and physical size. *Experimental Psychology*, 61, 149–163. https://doi.org/10.1027/1618-3169/a000235

Fitousi, D., & Algom, D. (2006). Size congruity effects with two-digit numbers: Expanding the number line? *Memory and Cognition, 34*, 445–457. https://doi.org/10.3758/BF03193421

Fitousi, D., & Algom, D. (2018). A system factorial technology analysis of the size congruity effect: Implications for numerical cognition and stochastic modeling. *Journal of Mathematical Psychology*, *84*, 57–73. https://doi.org/10.1016/j.jmp.2018.03.006

Fitousi, D., & Algom, D. (2020). A model for two-digit number processing based on a joint Garner and system factorial technology analysis. *Journal of Experimental Psychology: General*, *149*(4), 676–700. https://doi.org/10.1037/xge0000679

Fox, E. (1998). Perceptual grouping and visual selective attention. *Perception and Psychophysics*, *60*, 1004–1021. https://doi.org/10.3758/BF03211935

Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*, 59–65. http://doi.org/10.1016/S1364-6613(99)01424-2

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

García-Orza, J., & Damas, J. (2011). Sequential processing of two-digit numbers: Evidence of decomposition from a perceptual number matching task. *Zeitschrift fur Psychologie / Journal of Psychology*, *219*, 23–29. https://doi.org/10.1027/2151-2604/a000042

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

Goldfarb, L., Henik, A., Rubinsten, O., Bloch-David, Y., & Gertner, L. (2011). The numerical distance effect is task dependent. *Memory and Cognition*, *39*, 1508–1517. https://doi.org/10.3758/s13421-011-0122-z

Gough, J. (2008). A billion is how big? *Mathematics Teaching*, *210*, 12–15. http://www.atm.org.uk/mt/archive/mt210.html

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory and Cognition*, 10, 389–395. https://doi.org/10.3758/BF03202431

Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory and Cognition*, *10*, 487–495. https://doi.org/10.3758/BF03197652

Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 890–901. https://doi.org/10.1037/0096-1523.33.2.483

Hommel, B. (1995). Attentional scanning in the selection of central targets from multi-symbol strings. *Visual Cognition*, *2*, 119–144. https://doi.org/10.1080/13506289508401727

Huber, S., Bahnmueller, J., Klein, E., & Moeller, K. (2015). Testing a model of componential processing of multi-symbol numbers—evidence from measurement units. *Psychonomic Bulletin and Review*, *22*, 1417–1423. https://doi.org/10.3758/s13423-015-0805-8

Huber, S., Cornelsen, S., Moeller, K., & Nuerk, H. C. (2015). Toward a model framework of generalized parallel componential processing of multi-symbol numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 732–745. http://doi.org/10.1037/xlm0000043

Huber, S., Klein, E., Willmes, K., Nuerk, H., & Moeller, K. (2014). Decimal fraction representations are not distinct from natural number representations–evidence from a combined eye-tracking and computational modeling approach. *Frontiers in Human Neuroscience*, *8*, 172. https://doi.org/10.3389/fnhum.2014.00172

Huber, S., Mann, A., Nuerk, H. C., & Moeller, K. (2014). Cognitive control in number magnitude processing: evidence from eye-tracking. *Psychological Research*, *78*, 539–548. http://doi.org/10.1007/s00426-013-0504-x

Huber, S., Nuerk, H., Willmes, K., & Moeller, K. (2016). A general model framework for multisymbol number comparison. *Psychological Review*, *123*, 667–695. https://doi.org/10.1037/rev0000040

Hurst, M., Leigh Monahan, K., Heller, E., & Cordes, S. (2014). 123s and ABC s: developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values. *Developmental science*, *17*, 892–904. https://doi.org/10.1111/desc.12165

Kabasakalian, R. (2007). Language and thought in mathematics staff development: A problem probing protocol. *Teachers College Record*, *109*, 837–876. http://www.tcrecord.org/Content.asp?ContentId=12867

Kallai, A. Y., & Tzelgov, J. (2012). When meaningful components interrupt the processing of the whole: The case of fractions. *Acta Psychologica*, *139*, 358–369. ttps://doi.org/10.1016/j.actpsy.2011.11.009

Kastberg, S., & Walker, V. (2008). Insights into our understandings of large numbers. *Teaching Children Mathematics*, *14*, 530–536.

Korvorst, M., & Damian, M. F. (2008). The differential influence of decades and units on multidigit number comparison. *The Quarterly Journal of Experimental Psychology, 61*, 1250–1264. https://doi.org/10.1080/17470210701503286

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The source of the symbolic numerical distance and size effects. *Frontiers in psychology*, *7*, 1795. https://doi.org/10.3389/fpsyg.2018.00124

Landy, D., Charlesworth, A., & Ottmar, E. (2017). Categories of large numbers in line estimation. *Cognitive Science, 41*, 326–353. https://doi.org/10.1111/cogs.12342

Landy, D., Silbert, N., & Goldin, A. (2013). Estimating large numbers. *Cognitive Science, 37*, 775–799. https://doi.org/10.1111/cogs.12028

Leth-Steensen, C., & Marley, A. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*, 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Lozin, M., & Pinhas, M. (2021, October 7). Differential processing of “small” and “large” multi-digit numbers. Retrieved from https://osf.io/u84tc

Macizo, P., & Herrera, A. (2011). Cognitive control in number processing: Evidence from the unit-decade compatibility effect. *Acta Psychologica*, *136*, 112–118. https://doi.org/10.1016/j.actpsy.2010.10.008

Macizo, P., & Herrera, A. (2013). The processing of Arabic numbers is under cognitive control. *Psychological Research*, *77*, 651–658. https://doi.org/10.1007/s00426-012-0456-6

Macizo, P., & Ojedo, F. (2018). The processing of price comparison is not holistic but componential. *Psychological Research*, *82*, 1144–1157. https://doi.org/10.1007/s00426-017-0884-4

Meffe, G. K. (1994). Human population control: The missing awareness. *Conservation Biology*, *8*, 310–313. https://doi.org/10.1046/j.1523-1739.1996.10030900.x

Meyerhoff, H. S., Moeller, K., Debus, K., & Nuerk, H. C. (2012). Multi-digit number processing beyond the two-digit number range: A combination of sequential and parallel processes. *Acta Psychologica*, *140*, 81–90. https://doi.org/10.1016/j.actpsy.2011.11.005

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*, 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception and Psychophysics*, *49*, 270–288. https://doi.org/10.3758/BF03214311

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

Moeller, K., Fischer, M. H., Nuerk, H. C., & Willmes, K. (2009). Sequential or parallel decomposed processing of two-digit numbers? Evidence from eye-tracking. *The Quarterly Journal of Experimental Psychology*, *62*, 323–334. https://doi.org/10.1080/17470210801946740

Moeller, K., Klein, E., & Nuerk, H.C. (2013). Influences of cognitive control in numerical cognition–adaptation by binding for implicit learning. *Topics in Cognitive Science*, *5*, 335–353. https://doi.org/10.1111/tops.12015

Moeller, K., Nuerk, H. C., & Willmes, K. (2009). Internal number magnitude representation is not holistic, either. *European Journal of Cognitive Psychology*, *21*, 672–685. https://doi.org/10.1080/09541440802311899

Moeller, K., Shaki, S., Göbel, S. M., & Nuerk, H. C. (2015). Language influences number processing–a quadrilingual study. *Cognition*, *136*, 150–155. https://doi.org/10.1016/j.cognition.2014.11.003

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520. https://doi.org/10.1038/2151519a0

Nuerk, H. C., Moeller, K., Klein, E., Willmes, K., & Fischer, M. H. (2011). Extending the mental number line: A review of multi-digit number processing. *Zeitschrift fur Psychologie / Journal of Psychology*, *219*, 3–22. https://doi.org/10.1027/2151-2604/a000040

Nuerk, H. C., Moeller, K., & Willmes, K. (2015). Multi-digit number processing. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (pp. 106–139). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.013.021

Nuerk, H. C., Wager, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), 25–33.

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

Nuerk, H. C., & Willmes, K. (2005). On the magnitude representations of two-digit words. *Psychological Science*, *47*(1), 52–72. https://psycnet.apa.org/record/2005-11470-006

Patel, P., & Varma, S. (2018). How the abstract becomes concrete: Irrational numbers are understood relative to natural numbers and perfect squares. *Cognitive Science*, *42*, 1642–1676. https://doi.org/10.1111/cogs.12619

Pinhas, M., & Tzelgov, J. (2012). Expanding on the mental number line: Zero is perceived as the “smallest”. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 1187–1205. https://doi.org/10.1037/a0027390

Pinhas, M., Tzelgov, J., & Ganor-Stern, D. (2012). Estimating linear effects in ANOVA designs: The easy way. *Behavior Research Methods, 44*(3), 788–794. https://doi.org/10.3758/s13428-011-0172-y

Pinhas, M., Tzelgov, J. & Guata-Yaakobi, I. (2010). Exploring the mental number line via the size congruity effect. *Canadian Journal of Experimental Psychology*, *64*, 221–225. https://doi.org/10.1037/a0020464

Poltrock, S. E., & Schwartz, D. R. (1984). Comparative judgments of multidigit numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 32–45. https://doi.org/10.1037/0278-7393.10.1.32

Resnick, I., Newcombe, N. S., & Shipley, T. F. (2017). Dealing with big numbers: Representation and understanding of magnitudes outside of human experience. *Cognitive Science*, *41*, 1020–1041. https://doi.org/10.1111/cogs.12388

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

Rips, L. J. (2013). How many is a zillion? Sources of number distortion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1257–1264. https://doi.org/10.1037/a0031143

Sasanguie, D., De Smedt, B., & Reynvoet, B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychological Research, 81*, 231–242. https://doi.org/10.1007/s00426-015-0734-1

Schäfer, S., & Frings, C. (2021). Different effects of spatial separation in action and perception. *Psychonomic Bulletin and Review, 28*, 845–852. https://doi.org/10.3758/s13423-020-01867-9

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444. https://doi.org/10.1111/j.1467-8624.2004.00684.x

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

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

Thompson, C. A., & Opfer, J. E. (2010). How 15 hundred is like 15 cherries: Effect of progressive alignment on representational changes in numerical cognition. *Child Development, 81*, 1768–1786. https://doi.org/10.1111/j.1467-8624.2010.01509.x.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge University Press.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology*, *39*, 321–359. https://doi.org/10.1006/jmps.1995.1033

Varma, S., & Karl, S. R. (2013). Understanding decimal proportions: Discrete representations, parallel access, and privileged processing of zero. *Cognitive Psychology*, *66*, 283–301. https://doi.org/10.1016/j.cogpsych.2013.01.002

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

Verguts, T., & Fias, W. (2008). Symbolic and nonsymbolic pathways of number processing. *Philosophical Psychology*, *21*, 539–554. https://doi.org/10.1080/09515080802285545

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic bulletin & review*, *12*, 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A delta-rule model of numerical and non-numerical order processing. *Journal of Experimental Psychology: Human Perception and Performance, 40*, 1092–1102. https://doi.org/10.1037/a0035114

Zhang , J., & Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, *57*, 271–295. https://doi.org/10.1016/0010-0277(95)00674-3

Zhang, J., & Wang, H. (2005). The effect of external representations on numeric tasks. *The Quarterly Journal of Experimental Psychology*, *58A*, 817–838. https://doi.org/10.1080/02724980443000340

Zhou, X., Chen, C., Chen, L., & Dong, Q. (2008). Holistic or compositional representation of two-digit numbers? Evidence from the distance, magnitude, and SNARC effects in a number-matching task. *Cognition*, *106*, 1525–1536. https://doi.org/10.1016/j.cognition.2007.06.003

Table 1

Experiment 1 - Linear Mixed Model Regression Results for RT

Analysis type	Variable	Small numerical range			Large numerical range		
		Mean slope	[95% CI]	*p*	Mean slope	[95% CI]	*p*
Within-scale analysis	Global distance (log)	16.15	[-8.10, 40.39]	.192	-18.50	[-35.61, -1.39]	.034
	Left digit distance	-16.39	[-22.82, -2.97]	.017	-18.32	[-46.31, 9.68]	.200
Between-scale analysis	Global distance (log)	3.19	[-9.27, 15.65]	.616	-44.02	[-56.87, -31.17]	< .000
	Left digit distance	1.25	[-1.56, 4.05]	.383	-1.67	[-8.93, 5.59]	.653
	Scale distance	-68.92	[-83.79, -54.04]	< .000	-25.94	[-38.99, -12.89]	< .000

Note. CI = confidence interval. For response time, a negative mean slope indicates that increasing values of a factor resulted in a slower linear decrease in response time, consistent with a distance effect.

Table 2

Experiment 2 - Linear Mixed Model Regression Results for RT

Analysis type	Variable	Small numerical range			Large numerical range		
		Mean slope	[95% CI]	*p*	Mean slope	[95% CI]	*p*
Within-scale analysis	Global distance (log)	-26.87	[-47.17, -6.57]	.010	-101.00	[-117.84, -84.17]	< .000
	Left digit distance	119.03	[88.06, 149.99]	< .000	-646.24	[-790.47, -502.01]	< .000
	Right digit distance	40.51	[15.10, 65.91]	.005	48.65	[-9.67, 106.97]	.102
Between-scale analysis	Global distance (log)	-107.96	[-129.65, -86.27]	< .000	-5.70	[-40.69, 29.29]	.749
	Left digit distance	26.39	[3.70, 29.16]	.023	5.30	[-36.84, 47.45]	.981
	Right digit distance	-7.27	[-3.69, 89.61]	.385	25.08	[-18.09, 68.25]	.749

Note. CI = confidence interval. For response time, a negative mean slope indicates that increasing values of a factor resulted in a linear decrease in response time, consistent with a distance effect.

Table 3

Experiment 3 - Linear Mixed Model Regression Results for RT

Variable	Tens to thousands			Millions to billions			Billions to trillions		
	Mean slope	[95% CI]	*p*	Mean slope	[95% CI]	*p*	Mean slope	[95% CI]	*p*
Ratio	143.58	[43.45, 243.70]	.005	10.70	[-417.18, 438.58]	.961	-206.85	[-690.65, 276.95]	.401
Left digit distance	32.78	[-12.81, 27.85]	.467	8.72	[-78.16, 95.59]	.844	-35.11	[-133.34, 63.11]	.482
Scale distance	7.52	[10.94, 54.61]	.003	231.38	[138.07, 324.69]	< .000	259.57	[154.07, 365.08]	< .000

Note. CI = confidence interval. For response time, a positive mean slope indicates that increasing values of a factor resulted in a linear increase in response time, consistent with a ratio effect, whereas, a negative mean slope indicates that increasing values of a factor resulted in a linear decrease in response time, consistent with a distance effect.

Table 4

Overview of Compatibility Effects by Experiment, Comparison Type, and Numerical Range

Experiment	Analysis type	Effect	Comparison	Numerical range	
				Small	Large
1	Between-scale analysis	Left-digit-scale compatibility	Compatible vs. incompatible	✓[a]	✓[a]
			Compatible vs. same left digit	✓[a]	✓[a]
			Same scale vs. all between-scale conditions	✓[a]	✓[a]
2	Within-scale analysis	Left-right-digit within-scale compatibility	Compatible vs. incompatible	✓	×
			Compatible vs. same left digit	✓	✓
			Compatible vs. same right digit	✓	×
	Between-scale analysis	Left-right-digit-scale compatibility	Left-right-digit-scale compatible vs. both digit-scale incompatible	✓	✓
			Left-digit-scale compatible vs. both digit-scale incompatible	✓[b]	✓[b]
			Right-digit-scale compatible vs. both digit-scale incompatible	✓	✓
3	Between-scale analysis	Left-digit-scale compatibility	Compatible vs. incompatible	✓	×[c]
			Compatible vs. same left digit	✓	×[c]

Note. The ✓ sign represents a significant comparison and the × sign represents a non-significant comparison.

[a] The two-way interaction between the Left-Digit-Scale Compatibility and Numerical Range was not significant in Experiment 1. Thus, the comparisons presented in the table are based on the follow-up analysis conducted on the main effect of left-digit-scale compatibility. [b] The comparison of left-digit-scale compatible vs. both digit-scale incompatible trials was significant in both numerical ranges, but in small numbers the effect was significantly smaller. [c] Both large numerical ranges in Experiment 3 (“millions to billions” and “billions to trillions”) revealed similar non-significant patterns and are therefore referred to as one range in the table.

Table 5

Overview of Distance and Ratio Effects by Experiment, Comparison Type, and Numerical Range

Experiment	Analysis type	Variable	Numerical range	
			Small	Large
1	Within-scale analysis	Global distance (log)	✗	✓
		Left digit distance	✓	✗
	Between-scale analysis	Global distance (log)	✗	✓
		Left digit distance	✗	✗
		Scale distance	✓	✓
2	Within-scale analysis	Global distance (log)	✓	✓
		Left digit distance	✓	✓
		Right digit distance	✓	✗
	Between-scale analysis	Global distance (log)	✓	✗
		Left digit distance	✓	✗
		Right digit distance	✗	✗
3	Between -scale analysis	Ratio	✓	✗[a]
		Left digit distance	✗	✗[a]
		Scale distance	✓	✓[a]

Note. The ✓ sign represents a significant comparison and the ✗ sign represents a non-significant comparison.

[a] Both large numerical ranges in Experiment 3 (“millions to billions” and “billions to trillions”) revealed similar non-significant patterns and are therefore referred to as one range in the table.

Figure 1

Examples of stimuli presented in Experiment 1

Numerical range / Left-digit-scale compatibility	Small		Large	
	Number 1	Number 2	Number 1	Number 2
Compatible	20 20 200	500 5,000 5,000	2,000,000 2,000,000 2,000,000,000	5,000,000,000 5,000,000,000,000 5,000,000,000,000
Incompatible	50 50 500	200 2,000 2,000	5,000,000 5,000,000 5,000,000,000	2,000,000,000 2,000,000,000,000 2,000,000,000,000
Same left digit	20 20 200	200 2,000 2,000	2,000,000 2,000,000 2,000,000,000	2,000,000,000 2,000,000,000,000 2,000,000,000,000
Same scale	20 200 2,000	50 500 5,000	2,000,000 2,000,000,000 2,000,000,000,000	5,000,000 5,000,000,000 5,000,000,000,000

Figure 2

Examples of Between-Scale (A) and Within-Scale (B) Stimuli Presented in Experiment 2.

A: Between-scale stimuli

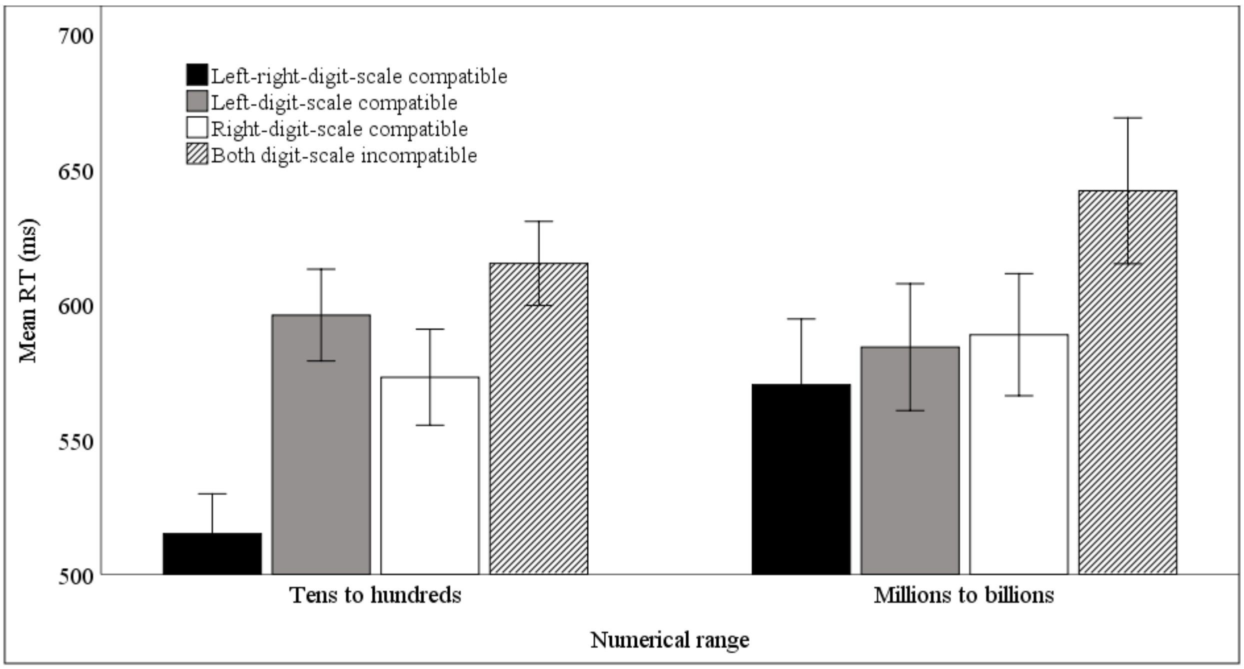
Left-right-digit-scale compatibility \ Numerical range	Small		Large	
	Number 1	Number 2	Number 1	Number 2
Left-right-digit-scale compatible	41	609	4,000,001	6,000,000,009
Left-digit-scale compatible	49	601	4,000,009	6,000,000,001
Right-digit-scale compatible	61	409	6,000,001	4,000,000,009
Left-right-digit-scale incompatible	69	401	6,000,009	4,000,000,001

B: Within-scale stimuli

Left-right-digit compatibility \ Numerical range	Small		Large	
	Number 1	Number 2	Number 1	Number 2
Compatible	41 401	69 609	4,000,001 4,000,000,001	6,000,009 6,000,000,009
Incompatible	49 409	61 601	4,000,000,009 4,000,000,009	6,000,001 6,000,000,001
Same left digit	61 601	49 409	6,000,001 6,000,000,001	4,000,009 4,000,000,009
Same right digit	69 609	41 401	6,000,009 6,000,000,009	4,000,001 4,000,000,001

Figure 3

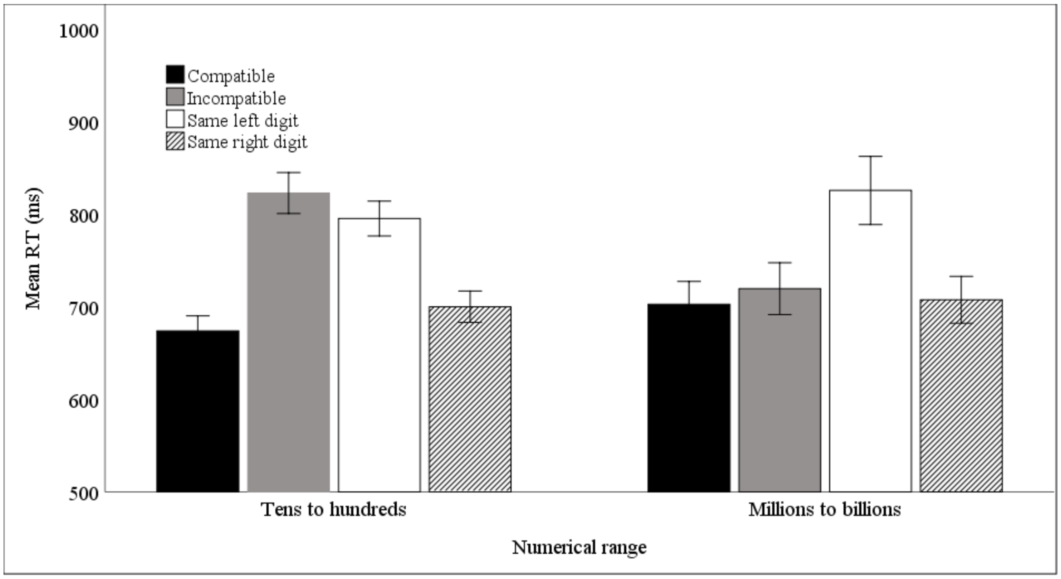
Experiment 2 - Mean RT as a Function of Numerical Range and Left-Right-Digit-Scale Compatibility



Note. Vertical bars denote ± *SE*.

Figure 4

Experiment 2 - Mean RT as a Function of Numerical Range and Left-Right-Digit Within-Scale Compatibility



Note. Vertical bars denote ± *SE*.

Figure 5

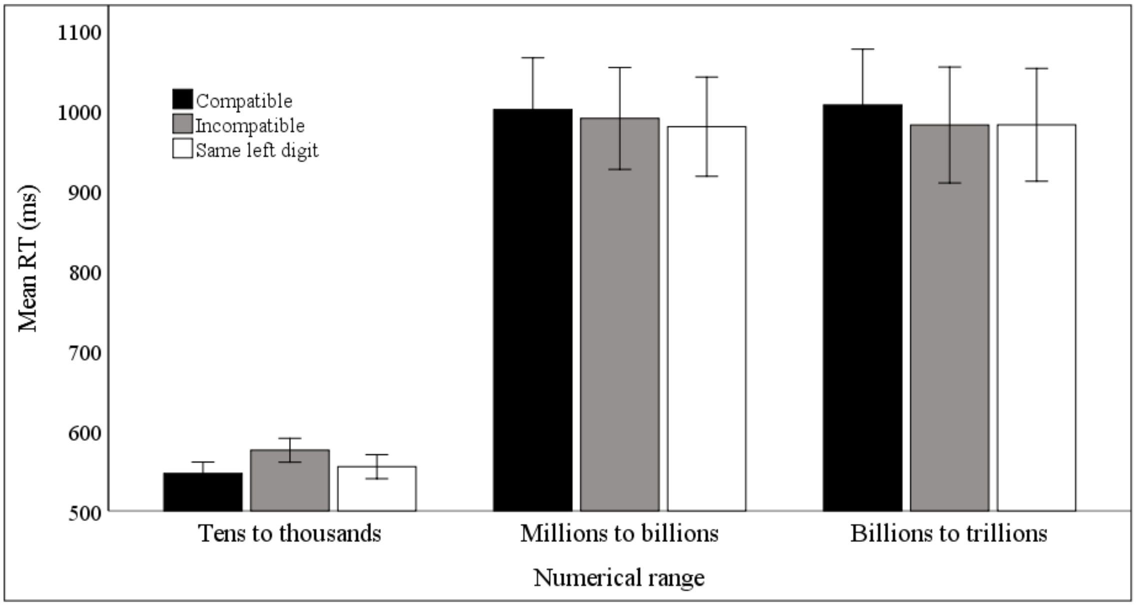
Examples of Stimuli Presented in Experiment 3

Numerical range / Left-digit-scale compatibility	Small		Large			
	Tens to thousands		Millions to billions		Billions to trillions	
	Number 1	Number 2	Number 1	Number 2	Number 1	Number 2
Compatible	20 20 200	500 5,000 5,000	20,000,000 20,000,000 200,000,000,000	500,000,000 5,000,000,000 5,000,000,000	20,000,000,000 20,000,000,000 200,000,000,000	500,000,000,000 5,000,000,000,000 5,000,000,000,000
Incompatible	50 50 500	200 2,000 2,000	50,000,000 50,000,000 500,000,000,000	200,000,000 2,000,000,000 2,000,000,000	50,000,000,000 50,000,000,000 500,000,000,000	200,000,000,000 2,000,000,000,000 2,000,000,000,000
Same left digit	20 20 200	200 2,000 2,000	20,000,000 20,000,000 200,000,000,000	200,000,000 2,000,000,000 2,000,000,000	20,000,000,000 20,000,000,000 200,000,000,000	200,000,000,000 2,000,000,000,000 2,000,000,000,000

.

Figure 6

Experiment 3 - Mean RT as a Function of Numerical Range and Left-Digit-Scale Compatibility



Note. Vertical bars denote ± *SE*.