Cognition, 129 (3), 512-529. doi:10.1016/j.cognition.2013.07.007

How do we convert a number into a finger trajectory?

Dror Dotan and Stanislas Dehaene

How do we understand two-digit numbers such as 42? Models of multi-digit number comprehension differ widely. Some postulate that the decades and units digits are processed separately and possibly serially. Others hypothesize a holistic process which maps the entire 2-digit string onto a magnitude, represented as a position on a number line. In educated adults, the number line is thought to be linear, but the “number sense” hypothesis proposes that a logarithmic scale underlies our intuitions of number size, and that this compressive representation may still be dormant in the adult brain. We investigated these issues by asking adults to point to the location of two-digit numbers on a number line while their finger location was continuously monitored. Finger trajectories revealed a linear scale, yet with a transient logarithmic effect suggesting the activation of a compressive and holistic quantity representation. Units and decades digits were processed in parallel, without any difference in left-to-right versus right-to-left readers. The late part of the trajectory was influenced by spatial reference points placed at the left end, middle, and right end of the line. Altogether, finger trajectory analysis provides a precise cognitive decomposition of the sequence of stages used in converting a number to a quantity and then a position.

1 Introduction

The invention of multi-digit numbers is a major achievement in our culture. It took mankind centuries to develop the idea that large numbers can be represented with merely 10 symbols by relying on their relative positions. During education, the human brain learns the decimal system and, ultimately, it becomes very intuitive that the digit 4 in 41 stands for four decades, while the digit 4 in 14 stands for four units. But what is it exactly that we understand? How does our brain represent multi-digit quantities, and what are the processes that convert a sequence of digit symbols into this quantity representation? In spite of our growing knowledge of the cognitive and neurological brain mechanisms of numerical cognition, the issue of multi-digit quantities was addressed by relatively few studies, and even fewer have investigated the processes that convert digits into these quantities. The present research explored these issues, and was centered on three major questions: holistic versus decomposed encoding of multi-digit quantities, the use of a logarithmic or a linear quantity scale, and sequential versus parallel processing of the digits in multi-digit numbers. In investigating these questions we aimed not only to describe the various cognitive representations of numbers in educated adults, but also to dissect the successive stages by which multi-digit Arabic numbers are converted into quantities.

1.1 Holistic vs. decomposed quantity representation

One of the main disputes about two-digit quantity representation is between the holistic and decomposed approaches. The holistic approach claims that two-digit numbers are represented as holistic quantities: similarly to single digits, two-digit numerals are recognized as a whole and mapped onto a memorized quantity (Dehaene, Dupoux, & Mehler, 1990; Reynvoet & Brysbaert, 1999). The decomposed approach proposes that when a person deals with symbolic multi-digit numerals, only the quantities associated with the individual digits are activated and manipulated (Nuerk & Willmes, 2005). For example, the decomposed approach postulates that comparing two 2-digit numbers is achieved using two separate comparisons – one of the decade digits and another of the unit digits (Meyerhoff, Moeller, Debus, & Nuerk, 2012; Moeller, Fischer, Nuerk, & Willmes, 2009; Nuerk & Willmes, 2005).

The holistic-decomposed debate often made use of the fact that it takes longer to compare two digits when they are farther apart (Moyer & Landauer, 1967). This distance effect was taken to show that the comparison is performed by converting numbers from the decimal notation to an internal quantity code. The holistic model was supported by the finding of a continuous distance effect even in a comparison task where participants had to compare two-digit number targets to a fixed reference such as 55 (Brysbaert, 1995; Dehaene et al., 1990). Crucially, the unit distance affected the comparison time even when the decade digits were different (e.g., comparing 69 with 55 is faster than comparing 61 with 55), and in certain experimental settings there was no discontinuity at decade boundaries (Dehaene et al., 1990; Dehaene, 1989; Hinrichs, Yurko, & Hu, 1981). To account for this finding, a decomposed model must assume that the unit digits are compared even when they are numerically irrelevant, and that an incompatible unit comparison result interferes with the decade comparison and slows it down. Such an explanation predicts that if the onset of the unit digits is manipulated to be slightly earlier than the decade digit onset, the irrelevant unit comparison should have greater effect and therefore increase its interference in RT. This prediction was refuted, thereby supporting the holistic model (Dehaene et al., 1990).

In a slightly different comparison task, however, in which the subjects have to decide which of two simultaneously presented 2-digit numbers was the larger, the decomposed approach was supported by the discovery that the distance effect is modulated by decade-unit compatibility: for equal overall distance, pairs of two-digit numbers are compared faster when the units comparison result is compatible with the two-digit comparison result (e.g., 32 versus 47, where 2 is smaller than 7) than when the units comparison is incompatible (e.g., 37 versus 52, where 7 is larger than 2). The decomposed model can explain this compatibility effect as an interference from the incompatible unit comparison (Macizo, Herrera, Román, & Martín, 2011; Nuerk, Kaufmann, Zoppoth, & Willmes, 2004; Nuerk, Weger, & Willmes, 2001; Nuerk & Willmes, 2005). The holistic model cannot explain the compatibility effect because such a model considers only the overall distance between the compared numbers. The decomposed model was also supported by a recent study that showed a unit digit quantity effect in two-digit number bisection (Doricchi et al., 2009). However, other studies failed to support the decomposed model because they found no decade-unit compatibility effect, both in number comparison (Ganor-Stern, Pinhas, & Tzelgov, 2009; Zhang & Wang, 2005; Zhou, Chen, Chen, & Dong, 2008) and when using semantic priming paradigms (Reynvoet, Brysbaert, & Fias, 2002; Reynvoet & Brysbaert, 1999).

Holistic and decomposed representations are not necessarily mutually exclusive. Number comparison studies suggest that the decade-unit compatibility effect is found when the numbers are presented simultaneously but not when they are presented sequentially, suggesting that subjects can adopt either a holistic or a decomposed strategy according to task demands (Ganor-Stern et al., 2009; Zhang & Wang, 2005; Zhou et al., 2008; but see Moeller, Nuerk, & Willmes, 2009 for an alternative explanation that conforms to a decomposed approach).

1.2 Compressive versus linear quantity representation

Much evidence shows that the internal quantity representation is tightly related with space, and that quantities are represented along a mental number line: in left-to-right readers at least, the magnitude of numbers influences manual responses made in the right or left side of space (Dehaene, Bossini, & Giraux, 1993; Shaki, Fischer, & Petrusic, 2009), eye gaze direction (Loetscher, Bockisch, Nicholls, & Brugger, 2010; Ruiz Fernández, Rahona, Hervás, Vázquez, & Ulrich, 2011), and the direction to which spatial attention is shifted (Fischer, Castel, Dodd, & Pratt, 2003). Furthermore, magnitude was shown to be encoded not only categorically as “small” or “large”, but in a continuous manner (Ishihara et al., 2006).

A common paradigm to explore the quantity representation consists in analyzing how individuals map numbers to positions on a number line. How subjects map numbers to space is assumed to reflect, at least in part, the structure of the mental number line, and hence of the quantity representation (Barth & Paladino, 2011; Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Booth & Siegler, 2006; Cappelletti, Kopelman, Morton, & Butterworth, 2005; Siegler & Booth, 2004; Siegler & Opfer, 2003; von Aster, 2000; but see Núñez, Cooperrider, & Wassmann, 2012). Number-to-position studies showed that young children initially map quantities using a compressive scale that resembles a log function, but this changes into a linear encoding during the first years of school (Berteletti et al., 2010; Booth & Siegler, 2006; Opfer & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003). The log-to-linear shift was hypothesized to result from education, and indeed compressive encoding was found in uneducated non-western adults but linear encoding was found in American adults (Dehaene, Izard, Spelke, & Pica, 2008). Interestingly, a compressive quantity scale can still be found in educated adults in other tasks that tap an implicit level of representation: inattentive mapping of non-symbolic quantities to position along a line (Anobile, Cicchini, & Burr, 2012), quantity estimation with non-spatial responses (Núñez, Doan, & Nikoulina, 2011), price estimation (Dehaene & Marques, 2002), number bisection (Lourenco & Longo, 2009), and randomness judgment for sequences of numbers (Banks & Coleman, 1981; Viarouge, Hubbard, Dehaene, & Sackur, 2010). Viarouge et al. were even able to define the compressive scale more precisely, because their results fit a log function better than power function.

The existence of a compressive internal number scale is also supported by neuronal recordings in macaque monkeys: neurons tuned to number in parietal and prefrontal cortex exhibit a Gaussian tuning curve only when plotted on a logarithmic scale (Nieder & Miller, 2003). Functional MRI experiments in human adults strongly suggest that such a representation continues to exist in the adult brain, at least for non-symbolic numerosities presented as concrete sets of objects (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

Another kind of evidence comes from studies showing that reaction time is correlated with the logarithm of the target number (Brysbaert, 1995; Dehaene et al., 1990; Dehaene, 1989). Such findings may imply a logarithmic (or compressive) quantity encoding, but they can also be explained if quantity encoding is linear and is increasingly fuzzy as numbers grow larger (the scalar variability model, Cordes, Gelman, Gallistel, & Whalen, 2001; Gallistel & Gelman, 1992).

In summary, evidence exists for both linear and compressive quantity representations. The two representations were found to co-exist even in the same individuals, when tested in different tasks or in different conditions (Anobile et al., 2012; Dehaene et al., 2008; Lourenco & Longo, 2009; Viarouge et al., 2010).

1.3 Parallel versus sequential processing of multi-digit numbers

Another question concerning multi-digit numbers is whether the digits are processed in parallel or sequentially. For words with fewer than 8 letters, expert readers seem to process all of the letters in parallel (Lavidor & Ellis, 2002; Weekes, 1997), and it could be expected that the same would occur with numbers. Indeed, findings from two-digit number comparison suggest that even when effects compatible with a decomposed representation are observed, the quantities of the separate digits are processed in parallel (Moeller, Fischer, et al., 2009). However, longer numbers also involve sequential processing, in number comparison task (Hinrichs, Berie, & Mosell, 1982; Meyerhoff et al., 2012), in recall tasks (Hinrichs & Novick, 1982), and in reading aloud and symbolic comprehension tasks such as same-different judgment and identification of specific sequences of digits (Friedmann, Dotan, & Rahamim, 2010).

1.4 The present research

The present research seeks to understand the process of encoding two-digit Arabic numbers as quantities. To examine this issue we introduce a novel methodology, which is a variation of the number-to-position task. In the traditional number-to-position task, participants are shown numbers and are required to mark, with a pencil, the corresponding position along a number line. By contrast, our participants performed the number-to-position task on an iPad tablet computer, which allowed continuous measurement of the finger trajectory. On each trial, a two-digit number between 0 and 40 was shown on the iPad screen, and the participants dragged their finger from a fixed starting point at the bottom of the screen to a position along a number line that was at the top of the screen (see Figure 1). The experiment software digitized the entire finger trajectory. Finger trajectories are a powerful measure because the finger position at a certain time during the trial tightly tracks the underlying cognitive operations (Santens, Goossens, & Verguts, 2011; Song & Nakayama, 2008a; see also Finkbeiner & Friedman, 2011; Finkbeiner, Song, Nakayama, & Caramazza, 2008; and see Song & Nakayama, 2008a, 2008b, 2009, regarding the use of finger trajectories to analyze non-numeric cognitive processes). Thus, analyzing the finger positions at different times in the trial could reveal how the quantity representation of two-digit numbers evolves over time. This is an advantage over using trial-level measures such as errors, reaction times, or the final position along a number line, because such measures can examine only the quantity representation at the end of the trial, whereas the "number-to-position trajectory" paradigm also allows examining the transient quantity representations.

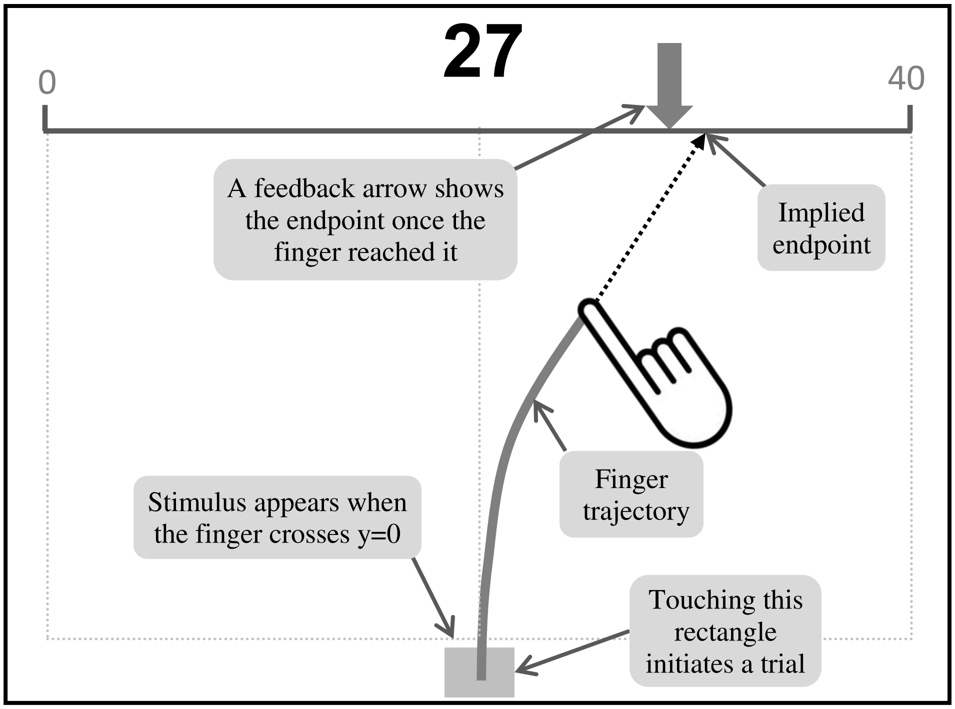

Figure 1. Task and screen layout. Participants were asked to point to the correct location for a 2-digit number on a horizontal line that extended from 0 to 40. On each trial, they first placed their finger on a bottom rectangle. The target appeared when they started moving their finger upward. The entire trajectory was digitized, and the measures were converted into instantaneous estimates of finger coordinates and implied endpoint.

Our study evaluated a broad array of distinct theoretical models that aim to explain how the number-to-position task is performed and what kind of quantity representation is used in this task. The different models assume either holistic or decomposed representations, linear or logarithmic quantity scales, similar or different processing of one-digit and two-digit numbers, and the last model also assumes a spatial strategy to aim the finger to the desired position. Each of these models predicts a certain spatial mapping of numbers to positions as well as a certain temporal pattern of finger trajectories. By analyzing various parts of the finger trajectories, we can reject some models and probe which of the models best fits the observed trajectories.

The predictions of the models are illustrated in Figure 2 as simulations of the predicted trajectories. These simulations are admittedly over-simplified and are provided only for visualization purposes. They ignore several parameters such as the fact that the finger changes its direction gradually and not abruptly, the fact that the finger velocity is not constant, and the existence of noise. Their purpose is only to convey graphically the variables that are supposed to influence finger trajectories.

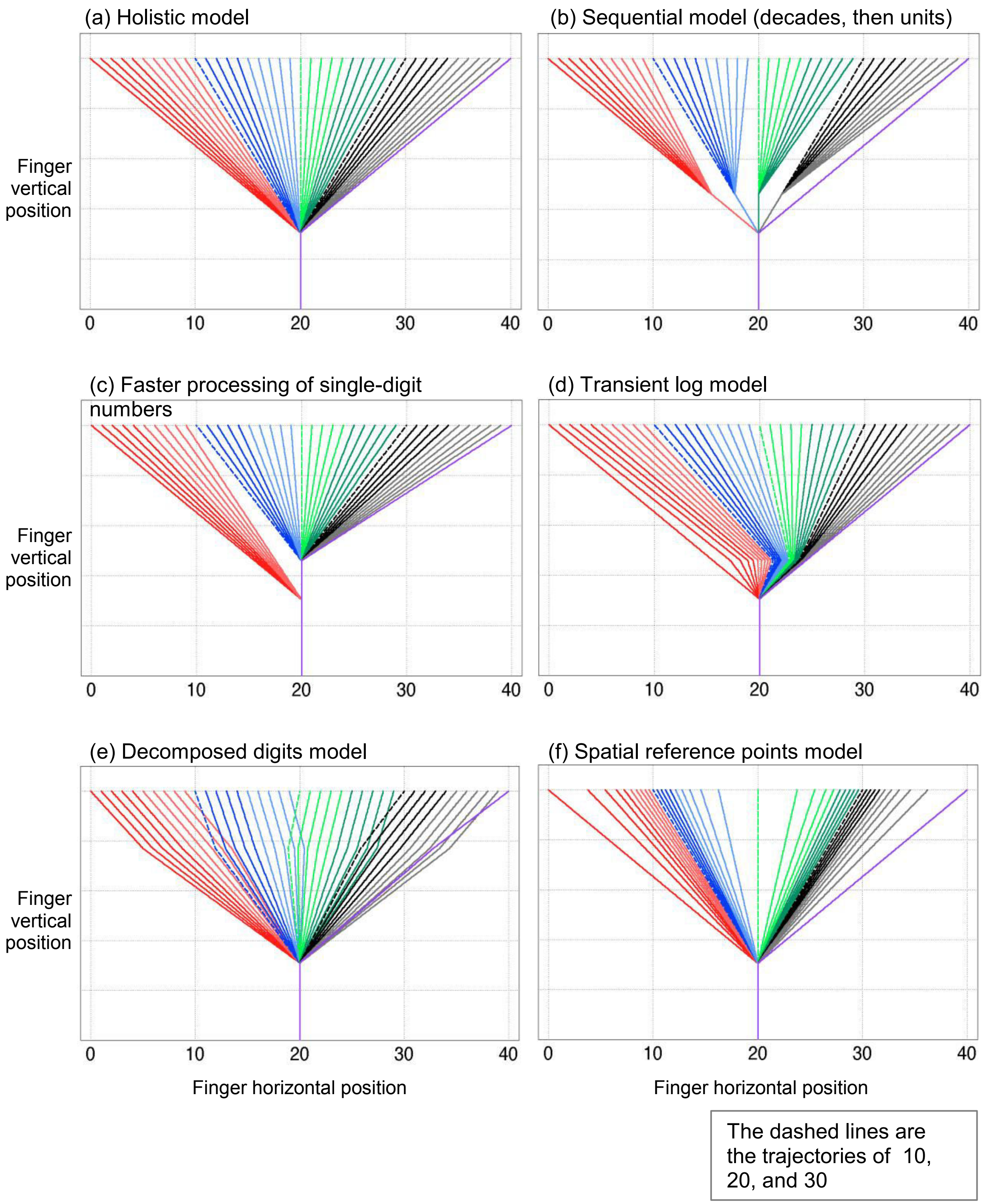

Figure 2. Six idealized models for spatial trajectories. Models are explained in detail in the main text.

All six simulations in Figure 2 assumed, for illustration only, an overall movement time (from starting point to the number line) of 1300 ms and a constant finger velocity. They also assumed that all trajectories begin with an exact upwards movement of the finger and deviate at a certain time. The models differ from each other with respect to the direction each trajectory takes once they branch apart.

a) The *holistic model* (Figure 2a) assumes a holistic quantity representation mapped to a linearly-organized number line. The simulation assumes that at 400 ms the trajectories branch towards the target position along a linearly-organized number line. Note that although this model is called "holistic", the same shape of trajectories is also predicted by a decomposed model that assumes that the unit and decade digits are processed in parallel and affect the finger position in exactly 1:10 ratio.

b) The *sequential model* (Figure 2b) assumes that the quantity representation is decomposed, and that the decade digit is processed earlier than the unit digit. The simulation assumes that at t=400 ms the finger starts moving towards the position of the relevant whole decade (because the unit digit information is still unavailable at this time). At t=600 ms the unit digit was processed too and again the trajectories branch apart, this time towards the correct target position along a linearly-organized number line.

c) The next model (Figure 2c) assumes a linear organization of the numbers along the number line, with faster processing of single-digit numbers than of two-digit numbers. The simulation assumed that the single-digit trajectories branch apart at t=400 ms, and at this time the finger starts moving towards the target position along a linearly-organized number line. The two-digit trajectories branch at t=600 ms, and in this case too, the finger starts moving towards a linearly-organized number line.

d) The *transient log* model (Figure 2d) assumes that a holistic, logarithmic quantity representation is first constructed, and this representation is then overridden by a linear representation. The simulation assumed that at t=400 ms the finger starts moving towards the target position along a logarithmically-organized number line, and at t=600 ms the finger direction is re-adjusted to aim towards a linearly-organized number line.

e) The *decomposed digits* model (Figure 2e) assumes that on top of the two-digit quantity, the quantities of each of the digits would also affect the finger position. One possible mechanism which may create such an effect is the existence of an intermediate stage in which the two digits of the target are not yet fully assigned to their respective unit and decades locations (Friedmann et al., 2010; Greenwald, Abrams, Naccache, & Dehaene, 2003). For a transient period, the two digits would therefore be floating and potentially submitted to illusory conjunctions (Treisman & Schmidt, 1982): the unit digit might be partially bound to the decade location, or vice-versa. The resulting quantity representation will be a linear combination of the two-digit quantity with the single-digit quantities. Our simulation represents the single-digit quantities using a "decomposed-digit factor" which is defined as the average of the two digits, linearly rescaled to the range between -2.5 and 2.5. The simulation assumes that at t=400 ms the finger starts moving towards the position of the target number plus the "decomposed-digit factor", and at t=1000 ms the finger direction is re-adjusted towards the exact linear position of the target number.

f) The *spatial reference points* model (Figure 2f) is specifically concerned with the process of translating a quantity into a spatial position on the visually presented number line. Inspired by previous work on the role of reference points in proportionality judgments (Hollands & Dyre, 2000; Spence, 1990), it assumes that the target's position on the number line is estimated with respect to three reference points: the two ends of the number line (0 and 40) and its middle (20). The position to which a number is mapped is assumed to be obtained by comparing the relative proportions of the estimated distances to the nearest two reference points (e.g., 7 is positioned by comparing its distances from 0 and from 20). Crucially, these estimated distances are scaled by a compressive function, giving rise to a non-linear bias term which pushes the participant's responses away from the reference points. Previous findings supported this notion (Barth & Paladino, 2011; Sullivan, Juhasz, Slattery, & Barth, 2011). To account for their findings, Barth and Paladino used a power function with the exponent as a free parameter, but in the present study, to avoid over-fitting (Opfer, Siegler, & Young, 2011) we used a log-based function log(*d*+1), where *d* is the linear distance. Thus, the exact position of 27 is calculated by the proportion between its estimated distances from 20 and 40, i.e., between log(7+1) and log(40-27+1), using the formula $\frac{\log(7+1)}{\log(7+1)+\log(13+1)} = .4407$. This proportion is then rescaled within the interval 0-20 to obtain a location on the complete range 0-40, i.e.: 20 + 20 * .4407 = 28.814 (very similar results were obtained with a power function with exponent 0.5, as proposed by Krueger (2010). In fact, the correlation between the log-based and the power-based functions over the integers between 0 and 40 is r = .9994).

The six models are not necessarily mutually exclusive. It is possible that different quantity representations dominate different parts of the trajectory. In fact, two of the models make this assumption explicitly: both the transient log and the decomposed digits models assume an intermediate representation that is then overridden by a linear representation. Our methodology allows for this possibility because we analyze the finger position in several time points along the trajectory. Another possibility is that two quantity representations co-exist simultaneously. This should result in a finger position that is some weighted average of the two quantity representations. Our methodology allows for this possibility by using regression analyses in which predictors from several theoretical models are put into a single regression model.

2 Method

2.1 Materials and Procedure

The experiment was performed on an iPad tablet computer[1]. Numbers between 0 and 40 were presented on screen, and the participants were required to point with their index finger at the corresponding position along an unmarked number line. Each target number was presented 10 times, so there were 410 trials, presented in random order. The experiment was administered in a quiet room in a single session, with a one-minute break every five minutes. The iPad was placed in its cradle (inclined towards the participant, about 15º from the table surface) in front of the participant's dominant hand.

Each trial began with a black screen with a horizontal number line at the top, marked with the labels 0 and 40 in its ends (see Figure 1). The number line remained on the screen throughout the experiment session. A trial started when the participant touched a rectangular area at the middle bottom part of the screen, thereby causing a fixation cross to appear above the middle of the number line. The participant then started sliding her finger forward (upward) towards the number line, and when the finger was 70 pixels (13.5 mm) away from the bottom of the screen, the target number between 0 and 40 replaced the fixation cross. The participant then moved her finger continuously up to the number line, to a position that corresponds with the target number. As the finger crossed the number line the target number disappeared, an acknowledgement "click" sound was played, and a green feedback arrow showed the participant where her finger actually hit the number line. The arrow did not show how accurate the response was – its purpose was only to help the participant improve the finger's motor aiming. The participant could then initiate the next trial whenever she wanted to, which was usually immediately.

A trial was considered as the time between the target onset and the finger crossing the number line. There were several rules that had to be kept within these time boundaries, and breaking any of them caused the trial to fail:

1. Only one finger could touch the screen, and it could not be lifted from the screen in mid-trial.
2. The finger could not move backward (downward).
3. To initiate a trial, the finger had to move upwards (and not sideways) from its starting point. The purpose of this rule was to prevent strong deviations in the initial direction of the finger trajectories.
4. To ensure that the experiment provided continuous trajectory information, a minimal finger velocity limit of 6 mm/sec was enforced, except for a grace period of the first 300 ms of each trial. The finger average velocity was also restricted: the finger had to complete 1/3 of the vertical distance to the number line within one second from the target onset, and reach the number line within two seconds from the target onset. The average-velocity restriction

[1] Details about our application, including a demo, can be found in http://www.trajtracker.com.

was interpolated linearly (e.g., the finger had to reach 1/6 of the vertical distance to the number line within half a second).

When a trial failed due to violation of one of these restrictions, the target number disappeared and an error message with feedback sound was presented. The target number of the failed trial was presented again later during the experiment, so each participant could complete 410 successful trials. Five participants, who had more than 25 failed trials (5.7%), were excluded.

2.1.1 Technical Specifications

The experiment used an Apple iPad-1 device and the software was written in Objective-C. The iPad screen size is 197x148mm, its resolution is 1024x768 pixels, and it reports the finger coordinates with the same resolution. The device was placed in landscape orientation. The screen background was black throughout the experiment. The number line was white, 844 pixels long (162mm), two pixels wide, and was located 80 pixels below the top of the screen, centered. The target number was shown in Arial bold white font, centered above the number line, and the digits were 10 mm high. The numbers 0 and 40 at the ends of the number line were shown in light grey Helvetica font and were 5 mm high (so they were a little less salient than the target number). The height and width of the fixation cross was 7.7 mm. The feedback arrow was green, 7.7 mm high, and was pointing downwards with its tip touching the number line.

The rectangular area that the participant touched to initiate a trial was dark grey and its size was 60x40 pixels, in landscape orientation. The target onset was triggered when the finger reached a distance of 70 pixels from the bottom of the screen. The vertical distance between the target onset point and the number line was 618 pixels (119mm).

2.2 Training

The experiment began with a short training that was done in four stages, each stage introducing some of the experiment rules. The first stage of training resembled the experiment procedure described above, with two differences: no minimal velocity was enforced, and no target number appeared. Instead of the target number, a downward-pointing red arrow appeared somewhere above the number line, and the participant was instructed to aim her finger “towards the red arrow”. The second training stage was the same, but it also enforced minimal finger velocity. The minimal velocity was visualized as an upward-moving horizontal line, and the participants were instructed to maintain their finger above the line. In the third training stage, minimal velocity was still enforced but the guiding horizontal line was not shown. The last training stage was identical to the experiment procedure, i.e., the targets were no longer red arrows but numbers between 0 and 40. The participants were shown the positions of 0, 20, and 40 but not of other numbers. In each training stage the experimenter first demonstrated what should be done, and the participant then performed a few training trials.

2.3 Participants

21 healthy adults participated voluntarily in the experiment. Ten of them were native speakers of Hebrew, which is read from right to left (RTL), and the rest were left-to-right (LTR)

readers – 9 French, one Italian, and one Thai. Numbers in Hebrew are read from left to right, like in English, and are printed using the same characters 0-9. All participants were right-handed, and their mean age was 35;5 (SD=10;7). There was no significant age difference between the LTR and RTL groups ($t_{(19)} = .66$, two-tailed $p = .52$).

2.4 Data Encoding

Several measures were calculated per trial. The *trial endpoint* is the position in which the finger crossed the number line, encoded using the number line's scale (0-40; endpoints were out of this range if the participant pointed outside of the number line). The *endpoint bias* is the difference between endpoint and the target number, with positive values indicating rightward bias. The *endpoint error* is the absolute value of endpoint bias. Finally, a trial *movement time* is the time elapsed from the target onset until the finger crossed the number line.

The finger trajectory throughout the trial was recorded as a sequence of x,y coordinates with timestamps attached to each. The finger position was sampled at the highest possible rate provided by the iPad (M=16 ms between subsequent samples, SD=1 ms) and then transformed into a fixed sampling rate of 100Hz using cubic spline interpolation. The fixed sampling rate allows comparing the finger coordinates in identical post-target-onset times between different trajectories.

2.5 Data Cleanup

Failed trials and trials with outlier endpoints were excluded from all analyses. A trial was considered as failed if one of the experiment's restrictions were violated, e.g., if the finger was lifted from the screen or was moved too slowly, or if the movement time was less than 200 ms. An outlier endpoint was defined with respect to endpoints of the 10 trials with the same target number, as an endpoint that exceeded the 25^{th} or 75^{th} percentile by more than 1.5 times the inter-quartile range.

2.6 Statistical Analysis

Each of the six models was analyzed using a two-stage analysis. The first stage was a set of regression analyses, one per participant. The dependent variable in these regressions was the finger x coordinate, and the predictors depended on the theoretical model being assessed. For example, the transient log model assumes that the logarithmic and linear representations co-exist, so it was assessed by a regression with two predictors – the target number (between 0 and 40) and its logarithm.

Each of these regression analyses was carried out per time point, in 50 ms intervals. This allowed examining how the quantity representation evolves over time. For example, the logarithmic model predicts that the log predictor will be strong in the regressions done for early time points, but that the linear predictor will be strong in late time points.

The second stage checked which of the predictors showed a consistent pattern over all participants – namely, whether the b values, which were obtained for a specific predictor in a specific time point for each of the participants, were significantly different from zero. This was assessed using repeated measures ANOVA which compared the b values with zero values (this was a within-subject factor), with a between-subject factor of language group, and the participant as the random factor. Non-significant b values were also included in this analysis. These ANOVAs were run per predictor and per time point, in 50 ms intervals. One-tailed p values were used when the mean b value was positive (which indicates a predicted result), and two-tailed p values were used when mean b was negative.

3 Results

3.1 General performance

The average movement time (from target onset until the finger reached the number line) was 1.11 seconds (SD = .14). The mean rate of endpoint outliers was 4.9% (SD = 1.8%). Excluding outliers, the mean endpoint error was 5.9 mm (SD = 1.59 mm), i.e., 1.45 numerical units on the 162 mm long, 0-40 number line. The mean endpoint bias was -.45 numerical units (SD = .34), i.e., a small leftward bias. The average rate of failed trials was 2.7% (SD = 1.4%). The total of 238 failed trials had the following failure reasons: minimal velocity violation (44.5%), multiple fingers touched the screen (20.2%), finger backward movement (10%), finger lifted from screen (9.7%), trials shorter than 200 ms (8.8%), and starting the trajectory sideways rather than upwards (6.7%).

The LTR group was slightly quicker than the RTL group: the mean movement time was 1.06 seconds in the LTR group vs. 1.18 seconds in the RTL group (unpaired $t_{(19)} = 2.15$, $p = .04$). The LTR group was also less accurate, with a mean endpoint error of 1.61 vs. 1.27 in the RTL group ($t_{(19)} = 2.13$, $p = .05$). There were no significant differences between the two groups with respect to the rates of outliers and failed trials ($|t_{(19)}| < 1.25$, $p > .22$).

Figure 3a shows a typical example for the shape of trajectories in the experiment (one line per trial). It shows the raw trajectories of one of the participants when responding to the target numbers 1, 12, 29, and 37. Figure 3b shows the trajectories, averaged over all trials and participants. For each target number, a median trajectory was calculated per participant by re-sampling the raw trajectories into equally spaced 201 time points and finding the median coordinate per time point (Santens et al., 2011; Song & Nakayama, 2008a). These median trajectories were then averaged across participants.

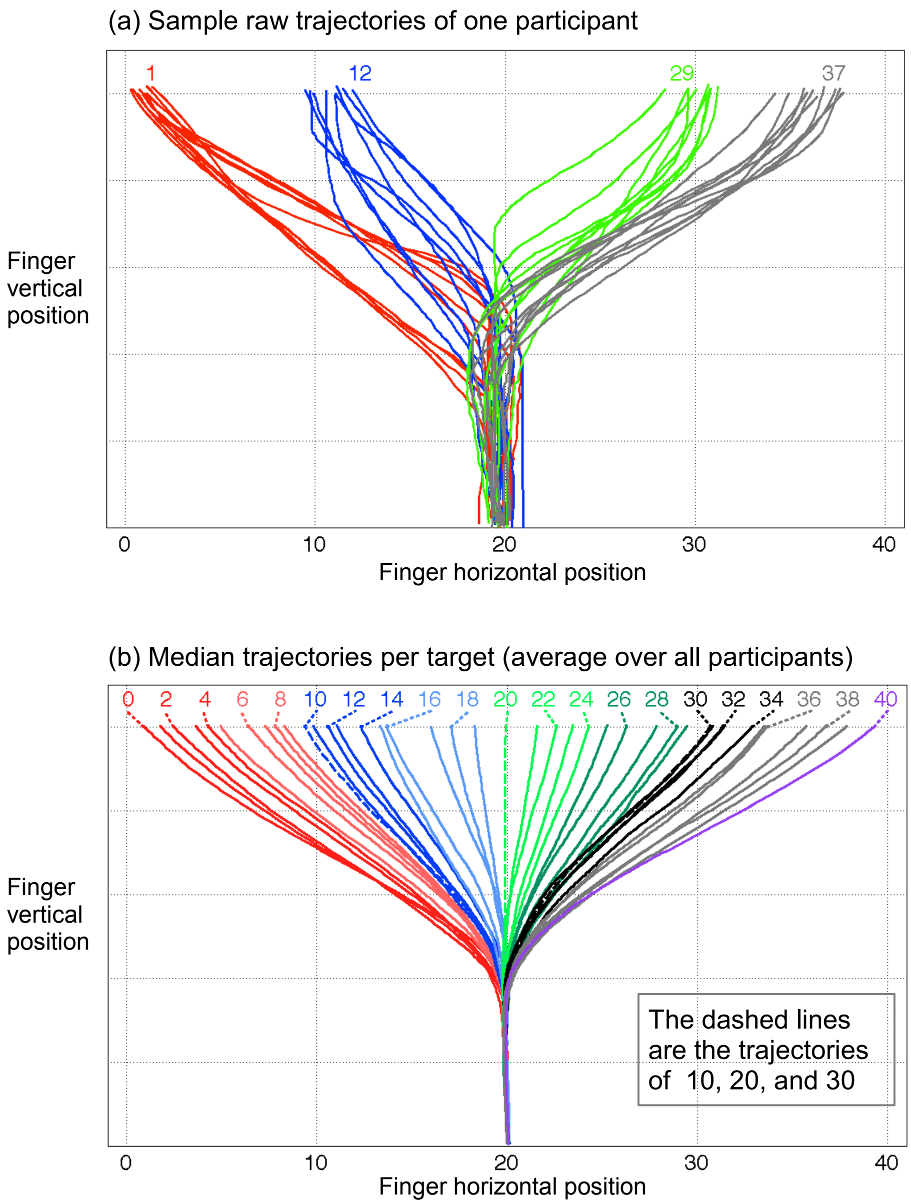

Figure 3. Two depictions of finger trajectories. (a) Spatial depiction of sample trajectories of one participant (finger location on the horizontal and vertical axes) to four distinct target numbers. (b) Spatial depiction of median trajectories for each target number, averaged across all participants.

3.2 Assessment of the models

3.2.1 The holistic model

The holistic model assumes that the task is performed by mapping the two-digit quantity to a linearly organized number line. This model was examined using regression analysis in which the dependent variable was the finger x coordinate, linearly transformed to the 0-40 scale of the number line (this variable is hereby denoted as $X_{0\text{-}40}$). There was a single predictor – the two-digit target number (which will be hereby denoted as $N_{0\text{-}40}$). One regression was run per participant and per time point, in 50 ms intervals.

Figure 4a shows the mean b values of $N_{0\text{-}40}$ over all participants in all time points. The b value (hereby denoted as $b[N_{0\text{-}40}]$) gradually increases as the trajectories branch apart. The mean r^2 value starts at 1% (SD = 1.4%) at t=450 ms and rises up to 97.2% (SD 1.7%) at the end of the trajectories.

A between-participant analysis was done by submitting the $b[N_{0\text{-}40}]$ values of all participants to repeated measures ANOVA (as described in section 2.6) with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor. This showed that $b[N_{0\text{-}40}]$ was significantly larger than zero as early as 450 ms post stimulus onset and in all subsequent time points. There was no significant difference in $b[N_{0\text{-}40}]$ between the left-to-right and the right-to-left readers in any time point ($F_{1,19} < 3.95$, $p > .06$; and in time points earlier than 1400 ms, $F_{1,19} < 3.0$, $p \geq .1$).

These results are in line with the holistic model. The real question, however, is whether the holistic-linear trend would remain strong even when compared with other models. As the next sections will show, the answer to this question is yes.

3.2.2 The transient log model

The transient log model assumes that mapping a number to the number line is governed by a logarithmic quantity representation in the early part of the trajectories, but by a linear-holistic quantity representation in the later parts. This model was examined using regression analysis with $X_{0\text{-}40}$ as the dependent variable, and with two predictors: the target number $N_{0\text{-}40}$, and a logarithmic predictor denoted as $log'(N_{0\text{-}40})$, which is $\log(1+N_{0\text{-}40})$, linearly transformed so that $log'(0) = 0$ and $log'(40) = 40$ (this transformation was used to allow meaningful comparison between the b values of the logarithmic and linear predictors). One regression was run per participant and per time point, in 50 ms intervals. The resulting b values were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor. The transient log model predicts an intermediate stage during which $b[log'(N_{0\text{-}40})]$ will be significantly larger than zero.

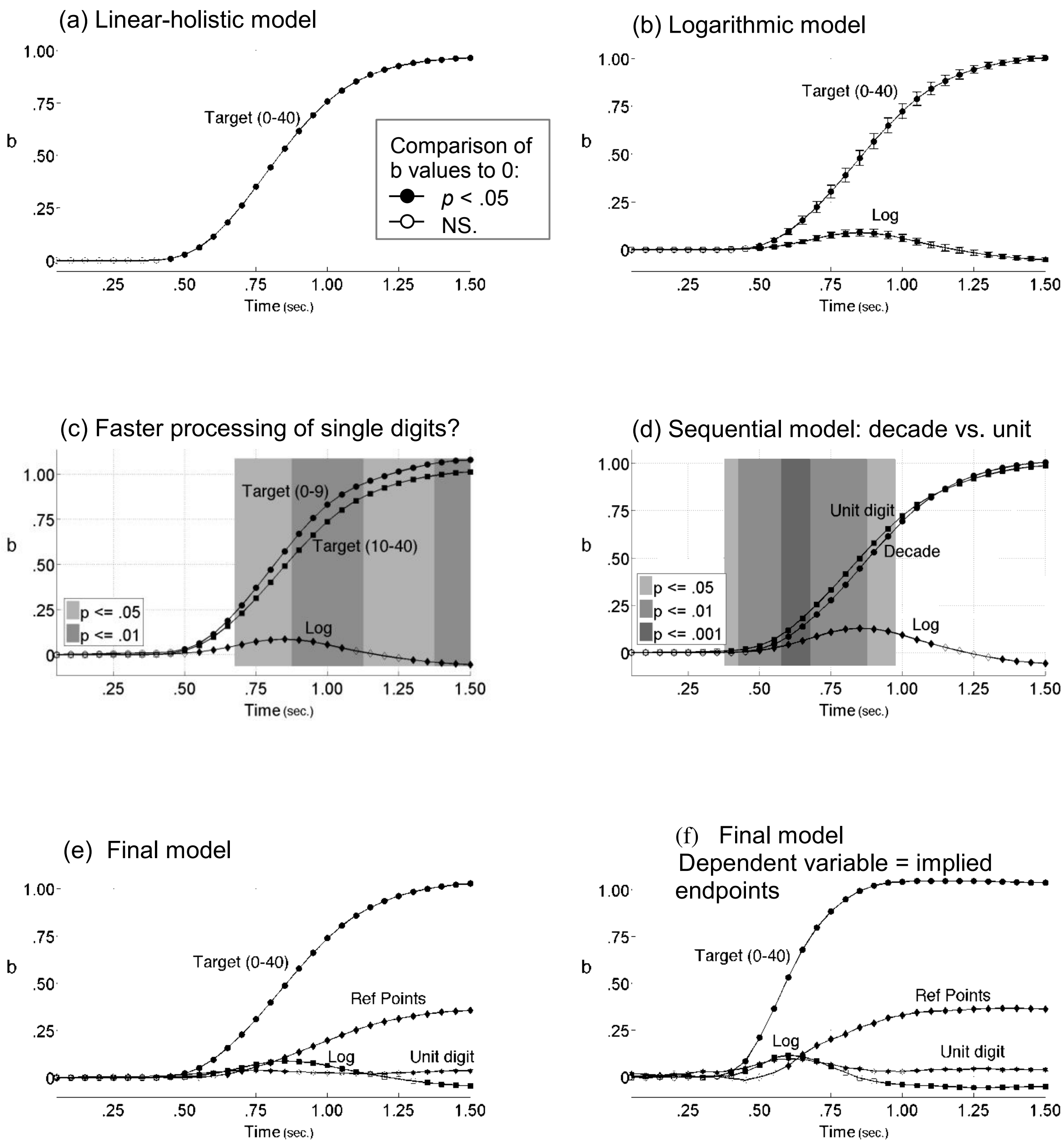

Figure 4. Incremental regression models of the finger trajectories. Each graph results from a multiple linear regression on horizontal finger location ($X_{0\text{-}40}$), performed separately for each subject and each time-point. The regression weights are then averaged over all participants and plotted as a function of time on the x axis. (a) regression with target number; (b) regression with the target and its log, showing a transient logarithmic effect (the error bars show one standard error across subjects); (c) separate assessment of one- and two-digit numbers shows no significant difference between the 0-9 and 10-40 predictors in early time points; (d) regression with distinct regressors for unit digits and decades; (e) final regression model with the target, its unit digit, its log, and spatial reference points; (f) The same regression model as in panel e, applied to the implied endpoint of the trajectory. Note that the effects appear earlier than in panel e: the implied endpoint provides a better reflection of the temporal dynamics of processing, because it directly reflects where the subject is aiming at a given time.

The results confirm this prediction (Figure 4b). $b[log'(N_{0-40})]$ was significantly larger than zero from 500 ms to 1050 ms, which indicates that a logarithmic quantity representation exists during this intermediate time window and then disappears. A linear representation of quantity also exists, as demonstrated by the finding that $b[N_{0-40}]$ was significantly larger than zero in all time points as of 450 ms. Note that at the time points in which the log predictor was significant, the linear predictor was significant too. This shows that the logarithmic quantity representation does not *precede* the linear quantity representation but exists *in parallel* to it. Finally, $b[log'(N_{0-40})]$ was significantly *smaller* than zero at all time points from 1300 ms and onwards. There was no significant difference between the language groups for any of the two predictors ($F_{1,19} < 2.1$, $p > .16$).

The reliable contribution of $log'(N_{0-40})$ could have an alternative explanation: it could be attributed to a logarithmic quantity representation of each of the digits, rather than to a logarithmic representation of the two-digit target number. According to such an explanation, the reason that $log'(N_{0-40})$ is a good predictor is its correlation with the logarithms of the decade and unit digits (indeed, the correlation between $log'(N_{0-40})$ and *log'(decade-digit) + log'(unit-digit)* is high, $r = .95$). To evaluate this alternative explanation, the trajectory data was submitted to regression analysis with X_{0-40} as the dependent variable and with four predictors: N_{0-40}, $log'(N_{0-40})$, *log'(decade-digit)*, and *log'(unit-digit)*. The resulting b values were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor. The alternative explanation predicts that this analysis will show significant contribution of the single-digit predictors, *log'(decade-digit)* and *log'(unit-digit)*, but the results refuted this prediction: *b[log'(decade-digit)]* and *b[log'(unit-digit)]* were not significantly larger than zero in any time point – in fact, they had negative values in all time points later than 650 ms.

Another alternative explanation is that the logarithmic effect results from a spatial or motor bias that is unrelated with the representation of quantities. Such a non-quantity account does not specifically predict a leftward or a rightward bias. To evaluate the specificity of the log function as a predictor, we compared it to a left-to-right mirror of this function. The trajectory data was submitted to regression with X_{0-40} as the dependent variable and with $N_{0\text{-}40}$ and the mirrored log fuction as predictors. This mirror function is denoted as $revlog'(N_{0-40})$ and defined as $revlog'(x) = 40 - log'(40-x)$. The resulting b values were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor. Contrary to the alternative explanation, $revlog'(N_{0-40})$ was not a good predictor of the finger position: the average $b[revlog'(N_{0-40})]$ values were negative in most time points, and the negative values were significantly lower than zero in time points later than 1300 ms. The spatial/motor alternative explanation was also refuted by a control experiment that did not involve numbers but arrows as targets, in which no log effect was found (see below in section 3.6).

3.2.3 Faster processing of single digits

This model assumes that numbers are mapped to a linearly-organized number line, but single-digit numbers are processed more quickly than two-digit numbers, and consequently the trajectories of single-digit numbers would branch earlier than the trajectories of two-digit numbers. The model was examined using regression analysis with $X_{0\text{-}40}$ as the dependent variable, with the logarithmic predictor $log'(N_{0\text{-}40})$, and with two additional linear predictors: $N_{0\text{-}9}$, which was the target number for single-digit trajectories and as zero for two-digit trajectories, and $N_{10\text{-}40}$, which was the opposite – the target number for two-digit trajectories and zero for single-digit trajectories. One regression was run per participant and per time point, in 50 ms intervals. The resulting b values were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor.

The "faster processing of single digits" model assumes a time window during which the trajectories of 0-9 have already branched but the trajectories of 10-40 have not yet branched. Consequently, at any time point during this time window (and even at later time points), the trajectories of 0-9 would be farther apart from each other than the trajectories of 10-40 (see Figure 2c), although the trajectories within each of the two groups would be still linearly organized. Thus, the model predicts an intermediate stage during which $b[N_{0\text{-}9}] > b[N_{10\text{-}40}]$.

The results (Figure 4c) did not confirm this prediction. A significant difference was found between $b[N_{0\text{-}9}]$ and $b[N_{10\text{-}40}]$ only in relatively late time point (from 700 ms and onwards, $F_{1,19} > 2.23$, $p < .05$).

The ANOVA showed no effect of language group ($F_{1,19} < 1.2$, $p > .28$).

3.2.4 The sequential model

The sequential model assumes that the quantity representation is decomposed and that the decade digit is processed earlier than the unit digit. The model therefore predicts an early stage during which only the decade digit affects the finger position. This model was examined using a regression analysis with $X_{0\text{-}40}$ as the dependent variable, with the logarithmic predictor $log'(N_{0\text{-}40})$, and with two additional linear predictors: the target number's decade (denoted as D and having the values 0, 10, 20, or 30) and the unit digit (U). The trajectory of target number 40 was excluded from this analysis in order to prevent a possible bias, because there is only one value of U for the decade 40, and the trajectory of 0 was excluded to maintain symmetry (but the results were similar even with the trajectories of 0 and 40 included). One regression was run per participant and per time point, in 50 ms intervals. The resulting b values were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor.

The sequential model predicts a time window during which $b[D] > b[U]$. The results (Figure 4d) did not confirm this prediction. In fact, they were the exact opposite: there was an intermediate time window, from 400 ms to 950 ms, during which $b[D]$ values were significantly *smaller* than $b[U]$ ($F_{1,19} > 4.8$, two-tailed $p < .04$). The language group did not interact with this decade-unit difference ($F_{1,19} < 1.62$, $p > .21$), and the effect was present in both groups (RTL group: from 550 ms to 1000 ms, $F_{1,9} > 2.42$, $p \leq .04$; LTR group: from 400 ms to 700 ms, $F_{1,9} > 5.76$, $p < .04$), indicating that it was not just due to the group of right-to-left readers treating the rightmost digit before the leftmost one. The faster processing of the unit digit was found also when the regression was run without the $log'(N_{0\text{-}40})$ predictor ($b[D] < b[U]$ from 350 ms to 1000 ms, $F_{1,19} > 4.49$, $p < .05$).

These results refute the sequential model. And yet, a parallel model is also not sufficient to fully explain the results. A parallel model assumes similar effects of decades and units on the finger x coordinates, and therefore predicts similar b values to the decade and unit predictors (because the decade predictor is the decade digit multiplied by *10*) – so such a model cannot explain the finding that $b[D] < b[U]$. The models described in the next two sections, however, provide a possible explanation of this finding.

3.2.5 The decomposed digits model

The decomposed digits model assumes that on top of the two-digit quantity, the quantities of each of the digits would also affect the finger position. Such a model would artificially inflate the impact of the unit predictor compared to the decade predictor, because in our regressions the unit predictor U is the unit digit itself, whereas the decade predictor D is the decade digit multiplied by 10. This model can therefore explain the unit-decade difference found in the regression in the previous section. Note, however, that the data allow us to exclude a strictly serial model in which, for a certain period of time, the two digits are freely floating without any binding to position. If that was the case, the $b[U]$ would have transiently been ten times larger than $b[D]$. We did not observe such an extreme effect: at the first point where the two regressors became significant (450 ms), the $b[U]:b[D]$ ratio was only 1.64, and it continuously decreased to reach the average value of 0.95 at the end of trajectories. This finding suggests that a more likely interpretation is that, for a transient period, both the single digits and the two-digit quantity are activated in parallel and contribute to the finger trajectory.

Could we provide a more specific test of this decomposed-digits model? The model predicts that the trajectories of number pairs such as 29 and 31 may be reversed, so that, transiently, 29 would be incorrectly mapped to a position to the right of 31. This is because the large difference between the units (9 versus 1) may override the much smaller difference between the decades (2 versus 3) or the whole-number quantity (29 versus 31). Indeed, there is prior evidence that such decade-unit compatibility effects may confuse two-digit number comparison judgments (Macizo et al., 2011; Meyerhoff et al., 2012; Nuerk et al., 2001; Nuerk & Willmes, 2005).

The general prediction is that, for targets around whole decades, trajectories should tend to be reversed, relative to how they would appear if they were simply based on the two-digit number quantity. This prediction was examined by analyzing the residuals of the log + linear regression (see Section 3.3.2). The residuals (x_{res}) were calculated per participant as the delta between the x value of the per-target median trajectories and the x value predicted by the log + linear regression described above. A median trajectory was created per subject and per target number by calculating the median coordinates for equivalent post-stimulus-onset time points (in 10 ms intervals). Late time points exceeded the movement time of some trajectories; for those, the endpoint was used as the x coordinate. The residuals were calculated with respect to the log + linear regression which was run on the median trajectories. The reason for calculating x_{res} based on median rather than raw trajectories was that this allowed pairing together trajectories from corresponding targets in the within-subject ANOVA hereby described.

These residuals were compared in three separate comparisons, centered on each decade: 8-9 vs. 11-12, 18-19 vs. 21-22, and 28-29 vs. 31-32. Each of these comparisons was done using an ANOVA with x_{res} as the dependent variable, two within-subject factors of size (above or below decade) and distance from decade (1 or 2), a between-subject factor of language, and subjects as the random factor. The decomposed digits model predicts that the trajectories would tend to reverse, namely, that $x_{res}(28,29)$ would be larger than ˆ,$x_{res}(31,32)$, and corresponding differences around the decades 20 and 10.

The results confirmed this prediction: $x_{res}(28,29)$ was significantly larger than $x_{res}(31,32)$ in 550 ms ($F_{1,19} = 3.83$, one-tailed $p = .03$) and in all subsequent time points ($F_{1,19} > 8.01$, one-tailed $p \leq .01$). Similarly, $x_{res}(8,9)$ was larger than $x_{res}(11,12)$ in 550 ms and in all subsequent time points ($F_{1,19} > 4.24$, one-tailed $p \leq .03$). The comparison around 20 resulted in a more complicated pattern: the prediction of the decomposed digits model was confirmed only for an early time window of 50 ms – 550 ms, during which $x_{res}(18,19)$ was significantly larger than $x_{res}(21,22)$ ($F_{1,19} > 3.66$, one-tailed $p < .04$). However, the results were opposite in the late trajectory parts: $x_{res}(18,19)$ was significantly *smaller* than $x_{res}(21,22)$ in 750 ms and in all subsequent time points ($F_{1,19} > 4.45$, two-tailed $p \leq .05$). There was no significant effect of language group in any of these comparisons ($F_{1,19} < 3.9$, two-tailed $p > .06$) except a single time point (200 ms in the comparison around the decade 30; $F_{1,19} = 5.21$, two-tailed $p = .03$) and no interaction between the language group and the above/below decade factor ($F_{1,19} < 3.79$, two-tailed $p > .06$).

In summary, the decomposed digit model accounts for the results in the early time window around 550 ms: the finger trajectories do tend to reverse around all whole decades (though note that this effect was found in the residuals of the linear + log regression; as Figure 3b shows, the effect size was not sufficiently strong to yield a complete reversal of the physical trajectories themselves). However, the pattern in late time windows (from 750 ms) is different: trajectories tend to be reversed around 10 and 30, but 20 has a repulsion effect that pushes the trajectories away (this repulsion effect around 20 is quite visible in Figure 3b). The decomposed digit model cannot explain this pattern, but the next model offers a possible explanation.

3.2.6 The spatial reference points model

The spatial reference points model assumes that in order to determine the position to which a number is mapped, the participant estimates the distances between the number and two out of three reference points: the ends of the number line (0 and 40) and its middle (20). The model further assumes that this distance estimation is logarithmic and not linear. This model predicts that trajectories around 10 and 30 would cluster around the whole decade, and trajectories around 20 will be pushed away from 20 (see Figure 2f). Namely, the model predicts the pattern of results observed for late time windows in the analysis of residuals described in the previous section.

This model was examined using regression analysis with $X_{0\text{-}40}$ as the dependent variable, and with four predictors: the two-digit target $N_{0\text{-}40}$, the logarithmic predictor $log'(N_{0\text{-}40})$, the unit digit U (to account for possible effect of quantities of the decomposed digits), and a spatial-reference-points-based bias *SRP*. The *SRP* predictor was the delta between the target number and the spatial-reference-points-based estimated position, which was defined like the simulation function in section 1.4 (f):

For $0 \le N \le 20$, $SRP(N) = 20 * \frac{\log(N+1)}{\log(N+1) + \log(21-N)} - N$

For $21 \le N \le 40$, $SRP(N) = 20 + 20 * \frac{\log(N-20+1)}{\log(N-20+1) + \log(41-N)} - N$

The b values from this regression were compared with zero using repeated measures ANOVA with b versus zero as a within-subject factor, the language group as a between-subject factor, and the participant as the random factor.

The results (Figure 4e) support the spatial reference points model, as well as the assumption that the SRP effect occurs in late time windows: *b[SRP]* was significantly larger than zero in all time points as of 650 ms ($F_{1,19}$ >4.07, p < .03). As for the other predictors, $b[N_{0\text{-}40}]$ was significantly larger than zero in all time points as of 450 ms ($F_{1,19}$ = 3.25, p = .04 at 400 ms, and $F_{1,19}$ > 14.71, p < .001 thereafter). $b[log'(N_{0\text{-}40})]$ was significantly larger than zero from 550 ms to 1050 ms ($F_{1,19}$ > 3.8, p < .02). *b[U]* was larger than zero from 400 ms and onwards (this effect was marginally significant from 900 ms to 1250 ms, $F_{1,19}$ > 1.88, p < .1, and significant in the other time points, $F_{1,19}$ > 3.16, p < .05). There were no significant differences between the language groups for any of the predictors ($F_{1,19}$ < 2.06, p > .16).

Thus, the line of analyses described above indicated that the position to which a number is mapped along the number line is determined by multiple factors: the linear two-digit quantity representation, a logarithmic quantity representation, the decomposed quantity of the unit digit, and a spatial-reference-point-based bias.

3.3 Limitations of the spatial reference points model

The final regression model, described in the previous section, showed that the endpoints (and the finger x coordinates in the final trajectory parts) depart from a strictly linear organization along the number line. This bias is captured in the model by the spatial-reference-point function SRP and by the logarithmic predictor (which is significant but with negative values, and therefore cannot reflect logarithmic quantity representation). Although the SRP predictor was significant and the regression r^2 values were high, we believe that the SRP function we used is not the ideal explanation for the endpoint bias. One reason for this belief is the fact that the log predictor was significant with negative values, which is not explained by any theoretical model. Another way to look into this issue is by comparing the actual endpoint biases with the prediction of the SRP function. As Figure 5a shows, the SRP function only partially resembles the observed endpoint biases. Notably, there seems to be an overall leftward bias (mean bias = -.45), which is not predicted by the spatial reference points model. This bias was consistent across participants (comparing the participants' mean endpoint bias versus zero, $t_{(20)}$ = -6.23, two-tailed p < .001).

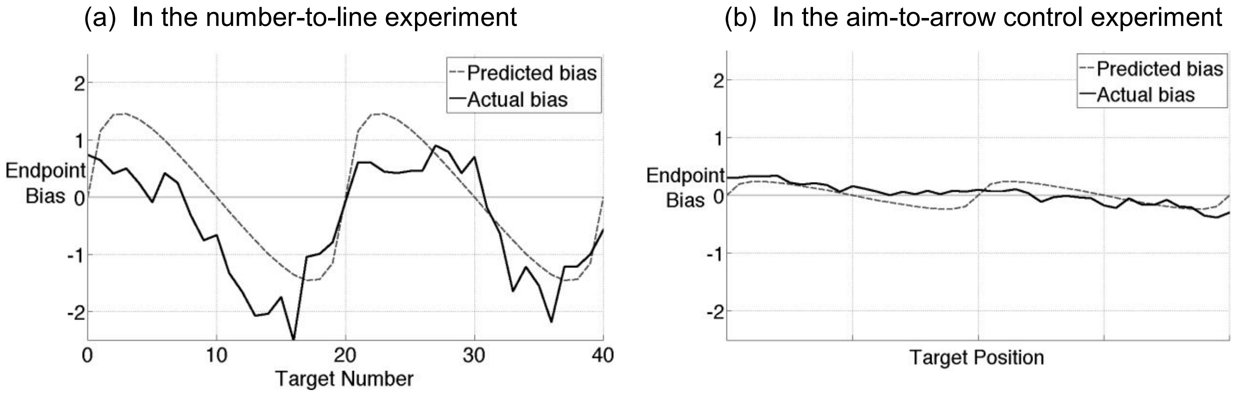

Figure 5. Endpoint biases (averaged over participants) compared to the prediction of the spatial reference points model. The "predicted bias" line shows the prediction of the spatial reference points model, linearly rescaled to fit the actually observed average bias. The y axis specifies the bias using the 0-40 scale.

3.4 Clarifying the time course of access to quantity

Most of the analyses throughout this study were based on the finger x coordinates, and we believe that the results showed it to be a powerful measure of the underlying quantity representation. However, the finger coordinates offer poor temporal granularity. The reason is that the finger position is slow in responding to cognitive changes, because even after the participant changes her cognitive representation of the finger's target position, there are still two things that must happen before the finger coordinate reflects this change: first, the finger must change its direction towards the new target position. Second, once the direction changes, it still takes time for the finger position to change: in essence, finger position is the time integral of direction and therefore smoothes out its fine-grained temporal variations.

We did not find a good way to eliminate the time it takes to change the finger direction, but there is a way to overcome the second factor – the time the finger spends moving in the new direction until its position changes. To overcome this factor, we used the finger's *implied endpoint* at each point along the trajectory rather than its x coordinate. The implied endpoint is the position along the number line that the finger would reach if it keeps moving in its current direction. The current direction (θ_t) was defined as the direction vector between the finger x,y coordinates at times *t - 50* ms and *t*. To prevent invalid values, the implied endpoint was cropped to the range [-2, 42] and was undefined when the finger moved sideways ($|\theta| > 80°$).

The regression of the final model was executed again, and this time the dependent variable was the implied endpoint. The predictors were the same as before: the two-digit target $N_{0\text{-}40}$, the logarithmic predictor $log'(N_{0\text{-}40})$, the unit digit U, and the spatial-reference-points-based bias *SRP*.

The implied-endpoint-based regression (Figure 4f) showed similar trends to those found in the x-coordinate-based regression: the linear factor was significant throughout the trajectory, the log factor in an early time window, and the SRP factor in late time windows. Importantly, the implied-endpoint-based regression indeed revealed earlier effects than the x-coordinate-based regression. $b[N_{0\text{-}40}]$ was significantly larger than zero at all time points beginning 400 ms. $b[log'(N_{0\text{-}40})]$ was significantly larger than zero as early as 450 ms and remained significant until 750 ms. $b[U]$ was significantly larger than zero from 300 ms and onwards, and $b[SRP]$ was significantly larger than zero at all time points beginning 600 ms. Thus, the implied endpoint analysis seems to provide a more accurate picture of how the cognitive quantity representation evolves over time.

The implied endpoint measure has one main disadvantage, which is the reason we did not use it as the main measure in this study, but preferred the finger x coordinate: the implied endpoint is a noisy measure. This happens because the implied endpoint is based on the derivative of position ($\frac{dx}{dy}$, i.e. the direction of the spatial trajectory) and thus is considerably affected even by small changes in its measurement, due to noise or to the iPad's limited resolution. We also investigated another measure that could have provided a more accurate estimate of finger direction and therefore of temporal flow, the time derivative of the finger x coordinate. However, this measure ($\frac{dx}{dt}$) turned out to be less accurate than the implied endpoint – its regression revealed the same four factors, but at later time points than the implied endpoint regression.

3.5 Numerical or spatial effects? Control experiment

Two of the findings described above could have alternative explanations that focus on motor factors rather than numerical processes. One such finding is the bias of endpoints from linear organization: we suggested that this bias is related to the cognitive representation of quantity or position, i.e., it is a bias in the way numbers are mapped to a planned position along the number line. An alternative explanation could be that the bias originates in the processes that guide the finger to this target position. The second finding is the transient log effect: an alternative explanation, mentioned in the end of section 3.3.2, attributes this effect to spatial or motor processes.

To assess these possibilities, we administered a control experiment, in which the target finger position was indexed non-numerically by an arrow. Importantly, no numbers were presented in this control task. If the spatial-reference-points bias originates in a quantity representation, no corresponding bias should be observed in this control task. If, however, the spatial-reference-points bias originates in non-numeric mechanisms, we expect to find a similar bias in the control experiment too. Similarly, if the log effect originates in a spatial/motor process, a similar effect should be observed in the control task too.

Participants

Ten healthy right-handed adults participated voluntarily in this experiment. They were all native Hebrew speakers. Their mean age was 34;3 (SD = 12;8).

Method

The method was similar to the number-to-position experiment, with a single difference: the target stimulus was not a number, but a downward-pointing red arrow placed at a specific position along the top line. The participants were instructed "to move their finger towards the arrow". Thus, this experiment was conducted exactly like the third training stage of the number-to-position experiment (see section 2.2).

Each target arrow could appear in one of 41 positions (corresponding with the positions of the numbers 0-40), and each position was presented four times, i.e., there were 164 trials in the experiment. No trials were defined as outliers because the number of trials per position (4) was insufficient for outlier analysis.

Results

The average movement time was 762 ms (SD = 167 ms). The mean endpoint error (rescaled to 0-40, to allow for comparison with the number-to-position experiment) was .39 (SD = .1). The mean endpoint bias was .02 (SD = .1). The average rate of failed trials was 1.5% (SD = 1.8%). Thus, the aim-to-arrow task was performed faster than the number-to-position task, more accurately, and with fewer errors.

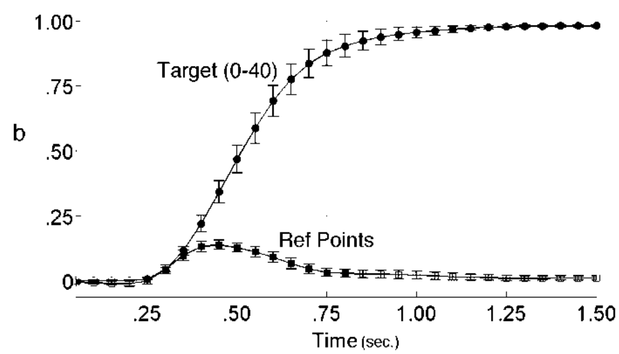

Figure 6. Control experiment, where the subject was asked to point to a flashed arrow: the effect of spatial reference points is present at an earlier moment, suggesting that it arises from a non-numerical level of representation. The error bars show one standard error across participants.

To assess the spatial reference points model, the trajectory data was submitted to regression analysis similar to the regressions reported for the number-to-position experiment. The dependent variable was $X_{0\text{-}40}$ and there were two predictors: the position of the target arrow along the line $N_{0\text{-}40}$, and the spatial-reference-point-based bias function SRP (detailed in section 3.3.6). The regression b values were compared with zero using t-test.

The results (Figure 7) showed that $b[N_{0\text{-}40}]$ was significantly larger than zero in all time points as of 250 ms ($F_{1,9} > 3.12$, $p < .01$), indicating a linear trend that begins even earlier than in the number-to-position experiment. The spatial reference points bias also had a significant effect:

b[SRP] was significantly larger than zero in all time points from 300 ms to 800 ms (much earlier than the value of 700 ms observed in the numerical task). The peak SRP effect size was at 450 ms (*b[SRP]* = .138), and then the SRP effect decreased and in the last part of the trajectories it was very small (*b[SRP]* < .02) and non-significant. Indeed, the organization of endpoints was almost perfectly linear (Figure 5b). This pattern is quite different from the pattern observed in the number-to-position experiment, in which the SRP effect began in later time points, and continuously increased as the fingers approached the number line.

To assess the transient log model, namely, the possibility that some spatial/motor process caused the log effect in the number-to-position task, the trajectory data in the aim-to-arrow task was submitted to a second regression analysis, which was similar to the regression described above, with a single difference – the addition of a third predictor, the logarithmic predictor $log'(N_{0\text{-}40})$. The regression b values were compared with zero using t-test. This analysis showed that the logarithmic predictor had no significant positive effect in any time window. In fact, $b[log'(N_{0\text{-}40})]$ was negative in all time points from 300 ms and onwards, and had significantly negative values from 350 ms until 750 ms ($t_{(9)} < -2.87$, two-tailed $p < .02$). Thus, the logarithmic trend in the number-to-position experiment should not be attributed to spatial or motor factors.

Discussion of the arrows task

The aim-to-arrow task showed that the trajectories deviate from a purely linear organization during an intermediate time window, and that the spatial reference points bias function can account for some of this deviation from linearity.

Why is the SRP bias observed only during an intermediate time window and then disappears? Most likely, as the finger approaches the target arrow, the participant can compare the finger position with the position of the target arrow (which is still visible on screen), and can readjust the finger trajectories to eliminate the bias.

Whether this explanation is correct or not, the results show a spatial reference points bias in a task that does not involve quantity estimation of numbers. It is therefore a plausible assumption that the SRP bias in the number-to-position task, unlike the logarithmic bias, originates, at least in part, in mechanisms unrelated to the quantity representation of numbers.

4 General discussion

This research aimed to clarify the processes involved in converting two-digit Arabic numbers into quantities and then into spatial coordinates. We investigated which cognitive representations are activated during this encoding process, either transiently or not. We used the number-to-position task and tracked the finger trajectories throughout each trial. An analysis of the factors influencing finger movement at various points in time revealed the underlying representations at various stages during a trial. Different predictors were used to assess five different theoretical models of quantity representation, and one model that concerns the way these quantities are mapped to spatial positions.

The findings suggest a multi-stage process which involves both holistic and decomposed quantity representations, with four factors affecting finger movement. These factors are now discussed in turn. The two measures of finger movement – x coordinate and implied endpoint – yielded very similar results, but we focus primarily on the implied endpoint regressions because they provided a more accurate timing of the underlying cognitive processes.

4.1 Linear representation

The strongest predictor of finger movement was the two-digit target number. This linear quantity was a reliable predictor of the implied endpoint at all time points starting at 400 ms following stimulus onset, and until the end of the trial. This finding suggests that a linear representation of the two-digit quantity (either holistic or decomposed) is quickly accessed and dominates the finger movement, as requested by the task. Assuming that it takes approximately 110-120 ms from motor intention to finger movement (Rammsayer & Stahl, 2007; Jaśkowski et al., 2007), our findings suggest that an intention is activated by 280-290 ms. Previous estimates, based on event-related potentials, suggest that digit identification takes places at about 160 ms, and that a quantity representation of single-digit numerals starts activating at 174 ms and is maximally activated approximately 210 ms after target onset (Dehaene, 1996). Based on this earlier study, the series of stages dominating the present task, possible organized in a cascade, are likely to be: identification (~160 ms), quantity (~170-210 ms), representation of the (linear) target location (~290 ms) and first finger deviation towards it (~400 ms). On top of this process there could be additional, faster or more automatic processing routes that process single digits (as is indicated by the finding of an early contribution of the units digit, see section 4.4 below). Such automatic processing is in line with previous studies (Pisella et al., 2000).

4.2 Transient logarithmic representation

The second factor that predicted finger location was the logarithm of the two-digit target number. This factor was a reliable predictor of the implied endpoint from 450 ms until 750 ms post stimulus onset. It indicates that a compressive quantity representation exists during an intermediate time window. We cannot conclude that the quantity representation was strictly logarithmic although the regression predictor was a log function, because several other compressive functions resemble the log function (e.g. a power function with exponent 0.5) and may have accounted for the results just as well.

The finding that the log factor started early and then disappeared suggests that the activation of a compressive representation is automatic rather than the result of conscious reasoning. Indeed, compressive quantity encoding was previously shown in educated adults in several paradigms (Anobile et al., 2012; Dehaene & Marques, 2002; Piazza et al., 2004; Viarouge et al., 2010). In the number-to-position task, however, it was shown only for young children (Berteletti et al., 2010; Booth & Siegler, 2006; Opfer & Siegler, 2007; Siegler & Booth, 2004; Siegler & Opfer, 2003) and for uneducated adults (Dehaene et al., 2008). The current research extends these previous findings and shows that educated adults use a compressive quantity scale even in the context of the number-to-position task. Earlier developmental and anthropological studies suggested that a few years of education suffice to move away from the innate compressive "number sense" that we share with animals (Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel & Gelman, 1992) and develop a linear sense of number (Booth & Siegler, 2006; Dehaene et al., 2008; Siegler & Booth, 2004; Siegler & Opfer, 2003). Nevertheless, the present findings confirm that an intuitive representation of numbers on a logarithmic scale remains dormant even in educated adults (Viarouge et al., 2010). Indeed, in agreement with previous studies, we found that linear and compressive quantity representations co-exist in the same individuals (Anobile et al., 2012; Lourenco & Longo, 2009; Viarouge et al., 2010).

The nature of the quantity scale is a topic of long-lasting debate between two different views. Some researchers showed how speed and accuracy decrease logarithmically as numbers become larger or closer, and suggested that these findings provided evidence for a compressive quantity scale (Brysbaert, 1995; Dehaene et al., 1990). A different interpretation of such findings, however, was offered by the scalar variability model, which proposes that quantities are encoded using a linear scale but the noise surrounding the quantity representation increases with number size (Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Cordes et al., 2001; Gallistel & Gelman, 1992; Whalen, Gallistel, & Gelman, 1999; but see Dehaene, 2001, 2003 for a discussion). Both the logarithmic model and the scalar variability model hold that it is harder to discriminate between large numbers than between small numbers, and therefore the two models are quite hard to separate, as they make very similar predictions when it comes to experimental measures such as reaction time, accuracy, and discriminability of numbers. The present findings, however, seem harder to explain by the scalar variability model. This is because the scalar variability model predicts greater confusability for larger numbers, but what we observed was a systematic *bias* in subjects' motor response. This bias is not easy to explain by the scalar variability model (even if it does not refute this model). The simplest explanation of the logarithmic trend in the number-to-position task is that quantities are indeed encoded using a compressive scale (Siegler & Opfer, 2003).

4.3 Holistic two-digit quantity representation

The finding of a logarithmic contribution to finger position indicates that the quantity representation is not only compressive but also holistic. A decomposed model could have explained the logarithmic factor as an artifact of logarithmic encoding of the single-digit quantities, but this alternative explanation was explicitly tested and ruled out, as we found a better fit of finger position with a log function of the whole 2-digit number. Thus, the results supports a holistic model, in agreement with previous studies (Dehaene et al., 1990; Ganor-Stern et al., 2009; Reynvoet & Brysbaert, 1999; Zhang & Wang, 2005; Zhou et al., 2008).

The results also do not support a sequential model, according to which the decade digit is processed before the unit digit. No time window was found in which the effect of the decade digit on the finger movement was larger than that of the unit digit. The results are therefore in accord with previous studies that showed parallel processing of two-digit numbers (Friedmann et al., 2010; Meyerhoff et al., 2012; Moeller, Fischer, et al., 2009). We also found no evidence that single-digit numbers are processed faster than two-digit numbers, as might be suggested by their simpler notation or higher frequency (Dehaene & Mehler, 1992): trajectories of single digit numbers did not branch apart earlier than trajectories of two-digit numbers.

Although the present study presents strong evidence in favor of a holistic processing of 2-digit numerals, this does not mean that numbers cannot be represented in a decomposed manner. As we reviewed in the introduction, other studies have presented evidence for decomposed processing. Subjects seems to strategically choose to process two-digit quantities holistically or in a decomposed manner, with different contexts facilitating different representations (Ganor-Stern et al., 2009; Greenwald et al., 2003; Reynvoet & Brysbaert, 1999; Zhang & Wang, 2005; Zhou et al., 2008). Some paradigms, such as the number-to-position task and the linear-distribution judgment task used by Viarouge et al. (2010), may encourage estimation and therefore facilitate holistic processing. Conversely, exact processing of several multi-digit stimuli may encourage decomposition strategies. Indeed, decomposed processing was often revealed when subjects had to compare two 2-digit numbers (Meyerhoff et al., 2012; Moeller, Fischer, et al., 2009; Nuerk & Willmes, 2005). A holistic strategy in this task would require encoding two separate 2-digit quantities almost simultaneously, which may be difficult. The number-to-position paradigm is simpler than number comparison because it presents a single target number per trial. In accord with this view, holistic processing was found in other tasks that showed only a single 2-digit number at a time (Dehaene et al., 1990; Ganor-Stern et al., 2009; Reynvoet & Brysbaert, 1999; Zhang & Wang, 2005; Zhou et al., 2008), whereas studies that presented more complicated stimuli – numbers with four or six digits – revealed that the digits can be processed sequentially (Hinrichs et al., 1982; Meyerhoff et al., 2012).

4.4 Effect of unit digit

A third factor influencing finger position was the unit digit, which was a reliable predictor of the implied endpoint from 300 ms post stimulus onset. The unit digit effect is also shown by the finding of a trajectory bias that corresponded with the unit digit: trajectories of target numbers with a small unit digit (1 or 2) were biased to the left compared to trajectories with a large unit digit (8 or 9). Three models can account for these findings: decomposed encoding of single-digit quantities, sequential processing of the 2-digit numbers (first the unit digit and then the decade digit), or transposition of the two digits. All these models focus on decomposed processing of the two digits, and the first model also assumes decomposed single-digit quantities.

The decomposed quantities model assumes that on top of the two-digit quantity, the single-digit quantities are encoded too, and thus the finger is influenced by their mean value. It could be objected that the results showed only an independent contribution of the unit digit and not of the decade digit. Note, however, that with the regression approach, we cannot independently estimate the effects of units u, decades d, and the whole number (=10d+u), as these three variables are linearly dependent. All we can therefore conclude is that the effect of the unit digit is, initially at least, larger than predicted by the equation 10d+u, and this is compatible with an additional contribution of the mean of d and u.

The second model assumes that the two digits are processed sequentially in a reversed order – first the unit digit, then the decade digit. As a result, the unit digit contributes to the quantity before the decade digit does. Thus, for a certain period of time the overall quantity –

whether if holistic or decomposed – over-represents the value of unit digit compared with the decade digit.

The third model that can account for the results is a transposition model. This model assumes a transient stage during which the digits are already identified but are not yet bound to their relative positions (Friedmann et al., 2010), thus creating illusory conjunctions (Treisman & Schmidt, 1982). During this transient stage, both the target quantity (10d+u) and the transposed quantity (10u+d) would be activated (e.g., presenting 28 activates both quantities 28 and 82), either as holistic or decomposed quantities, thus enhancing the overall effect of u on the finger position.

Finally, note that the unit digit effect was relatively small: in the x-coordinate regression, the maximal mean $b[U]$ that was significantly larger than zero was only .039. This peak happened at 750 ms following target onset, and the contributions of the other predictors in that time point were much larger ($b[N_{0\text{-}40}] = .31$, $b[log'(N_{0\text{-}40})] = .07$; the log predictor later reached a peak b value of .088, in 850 ms). Thus, the unit digit effect may indeed originate in decomposed quantity representation, but the more dominant quantity representation in this task is still the holistic one.

4.5 Spatial bias

The last factor to influence finger position was a spatial-reference-point bias function (SRP), which was a reliable predictor of the implied endpoint from 600 ms post stimulus onset and in all later time points – even the endpoints were biased away from a purely linear organization (Figure 7a), in agreement with previous studies (Barth & Paladino, 2011; Sullivan et al., 2011). This factor suggests that the target position in the number-to-position task is obtained using a non-linear estimation of the distances to three fixed reference points: the left end, middle, and right end of the number line.

The SRP bias function was also a significant factor in the aim-to-arrow task, although this task does not involve any numbers or quantities. This finding suggests that the SRP factor originates – at least in part – in a spatial/motor process rather than in the quantity representation or in the process that creates it from the Arabic number. A comparison of the SRP factor between the two tasks is in line with the assumption that this factor reflects a position-estimation error: in the aim-to-arrow task, the estimation error is expected to be larger when the finger is far from the target arrow, and indeed the SRP bias factor was observed early on in the trajectory. Conversely, the number-to-position task never presents the target position, so the position-estimation process continues throughout the trajectory, and correspondingly, the SRP bias was found late in the trajectory and even in its endpoint.

The spatial reference point model was able to account for much of the bias from linear organization in the number-to-position task, and yet the results did not fit the model perfectly. Two major findings indicate that the spatial reference points model should be amended to fully account for the bias observed in the present study: the existence of a global leftward endpoint bias, and the negative contribution of the log predictor to the final regression model (Figure 4e) – a finding that is not explained by any theoretical model yet.

4.6 The successive stages of converting a number to a position

Organizing these factors along a timeline clarifies the process performed in the number-to-position task. When the two-digit target number is presented, the participants first create a transient quantity representation of the unit digit (or, alternatively, a quantity representation of the transposed number, e.g., the quantity 52 upon seeing the target number 25). This representation is activated surprisingly quickly, as it affects the finger direction (implied endpoint) as early as 300 ms after the target number was presented. This finding is however not incompatible with earlier ERP studies, which indicate significant quantity effects as early as 174 ms after target onset (Dehaene, 1996), and with the finding that digit comparison can be performed above chance level 230 ms (Milosavljevic, Madsen, Koch, & Rangel, 2011).

Shortly afterwards, two separate representations of the two-digit quantity are created: a holistic logarithmic representation and a linear representation (either holistic, or decomposed with the unit and decade digits contributing in almost exact 1:10 ratio). The log and linear representations must be active at about 300 ms, since they start affecting the finger direction (implied endpoint) 450 ms and 400 ms after the target onset, correspondingly. The linear representation remains until the end of the trial, but the log representation is transient: 800 ms after the target onset, it no longer affects the implied endpoint.

Finally, as their finger approaches the target line, the participants start adopting a spatial strategy of transforming the two-digit quantity representation into a precise location on the number line. This strategy, which has a measurable effect 600 ms after the target onset, relies on three reference points (the left end, middle, and right end of the line), and results in a bias that pushes the finger trajectories away from these reference points.

4.7 Conclusion

The data presented here suggest that the number-to-position task involves a series of stages and underlying quantity representations: linear and compressive, holistic and decomposed representations, all play a role in this task, together with a spatial strategy for finger aiming. Analyzing finger trajectories provides a powerful method to assess cognitive processes, and specifically number representations, in real time. This measure is sensitive enough to be used not only in choice tasks (Song & Nakayama, 2009) but also as a continuous index of mental representations, especially when used with a device that allows for easy and natural hand movements, such as a tablet computer. In the future, similar paradigms could be used to further elaborate more aspects of quantity representation, to assess other mental computations, and to provide an easy diagnostic tool of the normal and pathological development of numerical abilities (Booth & Siegler, 2006).

Acknowledgements

This research was supported by INSERM, CEA, Collège de France, and a grant from the Bettencourt-Schueller Foundation.

References

Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition*, *122*(3), 454–9. doi:10.1016/j.cognition.2011.11.006

Banks, W. P., & Coleman, M. J. (1981). Two subjective scales of number. *Perception & Psychophysics*, *29*(2), 95–105. doi:10.3758/BF03207272

Barth, H. C., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science*, *14*(1), 125–35. doi:10.1111/j.1467-7687.2010.00962.x

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, *46*(2), 545–51. doi:10.1037/a0017887

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189–201. doi:10.1037/0012-1649.41.6.189

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, *12*(3), 238–243. doi:10.1111/1467-9280.00342

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

Cappelletti, M., Kopelman, M. D., Morton, J., & Butterworth, B. (2005). Dissociations in numerical abilities revealed by progressive cognitive decline in a patient with semantic dementia. *Cognitive Neuropsychology*, *22*(7), 771–93. doi:10.1080/02643290442000293

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, *8*(4), 698–707. doi:10.3758/BF03196206

Dehaene, S. (1989). The psychophysics of numerical comparison: a reexamination of apparently incompatible data. *Perception & Psychophysics*, *45*(6), 557–66.

Dehaene, S. (1996). The Organization of Brain Activations in Number Comparison: Event-Related Potentials and the Additive-Factors Method. *Journal of Cognitive Neuroscience*, *8*(1), 47–68. doi:10.1162/jocn.1996.8.1.47

Dehaene, S. (2001). Subtracting pigeons: Logarithmic or linear? *Psychological Science*, *12*(3), 244–246. doi:10.1111/1467-9280.00343

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, *7*(4), 145–147. doi:10.1016/S1364-6613(03)00055-X

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396. doi:10.1037/0096-3445.122.3.371

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355–361.

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.

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. doi:10.1126/science.1156540

Dehaene, S., & Marques, J. F. (2002). Cognitive euroscience: Scalar variability in price estimation and the cognitive consequences of switching to the euro. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, *55*(3), 705–731. doi:10.1080/02724980244000044

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

Doricchi, F., Merola, S., Aiello, M., Guariglia, P., Bruschini, M., Gevers, W., … Tomaiuolo, F. (2009). Spatial orienting biases in the decimal numeral system. *Current Biology*, *19*(8), 682–687. doi:10.1016/j.cub.2009.02.059

Finkbeiner, M., & Friedman, J. (2011). The flexibility of nonconsciously deployed cognitive processes: Evidence from masked congruence priming. *PloS one*, *6*(2), e17095. doi:10.1371/journal.pone.0017095

Finkbeiner, M., Song, J. H., Nakayama, K., & Caramazza, A. (2008). Engaging the motor system with masked orthographic primes: A kinematic analysis. *Visual Cognition*, *16*(1), 11–22. doi:10.1080/13506280701203838

Fischer, M. H., Castel, A. D., Dodd, M. D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*(6), 555–556. doi:10.1038/nn1066

Friedmann, N., Dotan, D., & Rahamim, E. (2010). Is the visual analyzer orthographic-specific? Reading words and numbers in letter position dyslexia. *Cortex*, *46*(8), 982–1004. doi:10.1016/j.cortex.2009.08.007

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*(1-2), 43–74. doi:10.1016/0010-0277(92)90050-R

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*(3), 444–452. doi:10.1080/17470210802391631

Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 235–247. doi:10.1037/0278-7393.29.2.235

Hinrichs, J. V, & Novick, L. R. (1982). Memory for numbers: Nominal vs. magnitude information. *Memory & Cognition*, *10*(5), 479–486.

Hinrichs, J. V, Yurko, D. S., & Hu, J. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(4), 890–901. doi:10.1037/0096-1523.7.4.890

Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory and Cognition*, *10*(5), 487–495. doi:10.3758/BF03197652

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

Ishihara, M., Jacquin-Courtois, S., Flory, V., Salemme, R., Imanaka, K., & Rossetti, Y. (2006). Interaction between space and number representations during motor preparation in manual aiming. *Neuropsychologia*, *44*(7), 1009–16. doi:10.1016/j.neuropsychologia.2005.11.008

Krueger, L. E. (2010). Reconciling Fechner and Stevens: Toward a unified psychophysical law. *Behavioral and Brain Sciences*, *12*(02), 251–267. doi:10.1017/S0140525X0004855X

Lavidor, M., & Ellis, A. W. (2002). Word length and orthographic neighborhood size effects in the left and right cerebral hemispheres. *Brain and Language*, *80*(1), 45–62. doi:10.1006/brln.2001.2583

Loetscher, T., Bockisch, C. J., Nicholls, M. E. R., & Brugger, P. (2010). Eye position predicts what number you have in mind. *Current Biology*, *20*(6). doi:10.1016/j.cub.2010.01.015

Lourenco, S. F., & Longo, M. R. (2009). Multiple spatial representations of number: Evidence for co-existing compressive and linear scales. *Experimental Brain Research*, *193*(1), 151–156. doi:10.1007/s00221-008-1698-9

Macizo, P., Herrera, A., Román, P., & Martín, M. C. (2011). The processing of two-digit numbers in bilinguals. *British Journal of Psychology*, *102*(3), 464–477. doi:10.1111/j.2044-8295.2010.02005.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*(1), 81–90. doi:10.1016/j.actpsy.2011.11.005

Milosavljevic, M., Madsen, E., Koch, C., & Rangel, A. (2011). Fast saccades toward numbers: simple number comparisons can be made in as little as 230 ms. *Journal of Vision*, *11*(4). doi:10.1167/11.4.4

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*(2), 323–334. doi:10.1080/17470210801946740

Moeller, K., Nuerk, H. C., & Willmes, K. (2009). Internal number magnitude representation is not holistic, either. *European Journal of Cognitive Psychology*, *21*(5), 672–685. doi:10.1080/09541440802311899

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. doi:10.1038/2151519a0

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron*, *37*(1), 149–157.

Nuerk, H. C., Kaufmann, L., Zoppoth, S., & Willmes, K. (2004). On the development of the mental number line: More, less, or never holistic with increasing age? *Developmental Psychology*, *40*(6), 1199–1211. doi:10.1037/0012-1649.40.6.1199

Nuerk, H. C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, *82*(1), B25–33.

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

Núñez, R., Cooperrider, K., & Wassmann, J. (2012). Number concepts without number lines in an indigenous group of Papua New Guinea. *PloS One*, *7*(4), e35662. doi:10.1371/journal.pone.0035662

Núñez, R., Doan, D., & Nikoulina, A. (2011). Squeezing, striking, and vocalizing: Is number representation fundamentally spatial? *Cognition*, *120*(2), 225–235. doi:10.1016/j.cognition.2011.05.001

Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*(3), 169–95. doi:10.1016/j.cogpsych.2006.09.002

Opfer, J. E., Siegler, R. S., & Young, C. J. (2011). The powers of noise-fitting: Reply to Barth and Paladino. *Developmental Science*, *14*(5), 1194–204; discussion 1205–6. doi:10.1111/j.1467-7687.2011.01070.x

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, *44*(3), 547–55. doi:10.1016/j.neuron.2004.10.014

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*(2), 191–201.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, *55*(4), 1127–39. doi:10.1080/02724980244000116

Ruiz Fernández, S., Rahona, J. J., Hervás, G., Vázquez, C., & Ulrich, R. (2011). Number magnitude determines gaze direction: Spatial-numerical association in a free-choice task. *Cortex*, *47*(5), 617–620. doi:10.1016/j.cortex.2010.10.006

Santens, S., Goossens, S., & Verguts, T. (2011). Distance in motion: response trajectories reveal the dynamics of number comparison. *PloS One*, *6*(9), Article ID e25429. doi:10.1371/journal.pone.0025429

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review*, *16*(2), 328–331. doi:10.3758/PBR.16.2.328

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*(2), 428–44. doi: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*(3), 237–250. doi:10.1111/1467-9280.02438

Song, J. H., & Nakayama, K. (2008a). Numeric comparison in a visually-guided manual reaching task. *Cognition*, *106*(2), 994–1003. doi:10.1016/j.cognition.2007.03.014

Song, J. H., & Nakayama, K. (2008b). Target selection in visual search as revealed by movement trajectories. *Vision Research*, *48*(7), 853–61. doi:10.1016/j.visres.2007.12.015

Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks. *Trends in Cognitive Sciences*, *13*(8), 360–6. doi:10.1016/j.tics.2009.04.009

Spence, I. (1990). Visual psychophysics of simple graphical elements. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 683–92.

Sullivan, J. L., Juhasz, B. J., Slattery, T. J., & Barth, H. C. (2011). Adults' number-line estimation strategies: Evidence from eye movements. *Psychonomic Bulletin & Review*, *18*(3), 557–63. doi:10.3758/s13423-011-0081-1

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141. doi:10.1016/0010-0285(82)90006-8

Viarouge, A., Hubbard, E. M., Dehaene, S., & Sackur, J. (2010). Number line compression and the illusory perception of random numbers. *Experimental Psychology*, *57*(6), 446–454. doi:10.1027/1618-3169/a000055

Von Aster, M. (2000). Developmental cognitive neuropsychology of number processing and calculation: Varieties of developmental dyscalculia. *European Child & Adolescent Psychiatry*, *9*, 41–57. doi:10.1007/s007870070008

Weekes, B. S. (1997). Differential Effects of Number of Letters on Word and Nonword Naming Latency. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, *50*(2), 439–456. doi:10.1080/713755710

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, *10*(2), 130–137. doi:10.1111/1467-9280.00120

Zhang, J., & Wang, H. (2005). The effect of external representations on numeric tasks. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, *58*(5), 817–838. doi: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*(3), 1525–1536. doi:10.1016/j.cognition.2007.06.003