Research Reports

Encoding of Numerical Information in Memory: Magnitude or Nominal?

Kimmo Eriksson*[a], Marcus Lindskog[b]

[a] School of Education, Culture and Communication, Mälardalen University, Västerås, Sweden. [b] Department of Psychology, Uppsala University, Uppsala, Sweden.

Abstract

In studies of long-term memory of multi-digit numbers the leading digit tends to be recalled correctly more often than less significant digits, which has been interpreted as evidence for an analog magnitude encoding of the numbers. However, upon closer examination of data from one of these studies we found that the distribution of recall errors does not fit a model based on analog encoding. Rather, the data suggested an alternative hypothesis that each digit of a number is encoded separately in long-term memory, and that encoding of one or more digits sometimes fails due to insufficient attention in which case they are simply guessed when recall is requested, with no regard for the presented value. To test this hypothesis of nominal encoding with value-independent mistakes, we conducted two studies with a total of 1,080 adults who were asked to recall a single piece of numerical information that had been presented in a story they had read earlier. The information was a three-digit number, manipulated between subjects with respect to its value (between 193 and 975), format (Arabic digits or words), and what it counted (baseball caps or grains of sand). Results were consistent with our hypothesis. Further, the leading digit was recalled correctly more often than less significant digits when the number was presented in Arabic digits but not when the number was presented in words; our interpretation of this finding is that the latter format does not focus readers' attention on the leading digit.

Keywords: numerical recall, numerical representations, positional numeral system, attention, guessing

People encounter numerical information in a multitude of situations in their day-to-day life. Numbers come in several formats, they may come across as large or small depending on the context, and the information they contain may be used in many ways. When numerical information is encountered, it is potentially stored in long-term memory and can later be recalled. For instance, a person reading in the morning paper about a shipwreck disaster that cost 376 lives may at a lunchtime conversation attempt to recall how many lives were lost. Here, we are interested in how accurate such recalled numbers tend to be and what this may tell us about how numerical information is encoded in long-term memory.

Memory for numbers was studied already by Thorndike (1910), when he investigated the relation between memory for numbers and memory for words. The topic has since received surprisingly little attention. In particular, only a few studies have addressed how numbers are encoded into long-term memory. Brainerd and Gordon (1994) suggested that numerical information could be stored in memory either verbatim or in a gist format. The verbatim information takes the form of the exact number itself (e.g., Farmer Brown has 9 cows)

while the gist of a number is defined relationally (e.g., Farmer Brown has more cows than dogs). Here we are interested only in the verbatim type of memory; our studies therefore use stimulus material that presents only a single number and no relational information.

Hinrichs and Novick (1982) proposed that the verbatim information about a number could be encoded in one of two forms, magnitude encoding or nominal encoding. In the former, numbers would be stored as approximate analog magnitudes. In the latter they would be stored digit-by-digit, like arbitrary labels without any numerical magnitude information. The idea that numbers may be encoded in memory by means of analog magnitude encoding is derived from investigations into how numbers are represented when people encounter and actively manipulate them. Moyer and Landauer (1967), for example, demonstrated that single digit comparisons were subject to a Numerical Distance Effect (NDE): When participants are asked to say which of two digits represents the greater number, responses tend to be faster when the numerical distance between the digits is large (e.g., 1 and 9) than when it is small (e.g., 4 and 5). Subsequent studies have also demonstrated the SNARC effect (Dehaene, Bossini, & Giraux, 1993), where people respond faster with the left hand to small numbers and faster with the right hand to large numbers. These findings support the notion that number symbols may be represented as magnitudes on a mental number line that runs from left to right (see also Hubbard, Piazza, Pinel, & Dehaene, 2005; Izard & Dehaene, 2008; Nuerk, Wood, & Willmes, 2005; Wood, Willmes, Nuerk, & Fischer, 2008).

Although the abovementioned NDE and SNARC studies do not address whether an analog magnitude representation is used in long-term memory, a couple of other studies do. Hinrichs and Novick (1982) found that the leading digit of a multi-digit number tended to better remembered than the following digits. This finding, which we shall refer to as the *leading digit correctness phenomenon*, was obtained also by Thompson and Siegler (2010).

The Leading Digit Correctness Phenomenon Is Consistent With Analog Encoding

Modelling Analog Encoding

The basic idea of the analog encoding hypothesis is that an encountered number is encoded as a magnitude with a certain degree of noise, which becomes evident when the number is recalled and reported. Moreover, the degree of noise is thought to be the main source of recall errors. A straightforward model is that a given presented number *X* would be recalled as

$X' = X + e$

where the error term *e* is drawn from a normal distribution with mean zero and standard deviation *σ,* representing the degree of noise. For a given presented number *X* and the recalled number *X'* we shall refer to the recall error as measured by the difference *X' – X* as the *residual.* Thus, the parameter *σ* can be empirically estimated as the standard deviation of residuals.

Modelling Analog Encoding With Log-Compression

While the model above is perhaps the most straightforward way to describe the analog encoding hypothesis it has been suggested that analog magnitude representations might not be linear but rather log-compressed (e.g., Dehaene, Izard, Spelke, & Pica, 2008). We therefore also make an alternative specification of the model, based on the common idea of a log-compressed representation. In this case the model would instead be

log X' = log X + $e_{\log}$

where $e_{\log}$ is the error term in the log-compressed representation. We would assume $e_{\log}$ to be drawn from a normal distribution with mean zero and some standard deviation $\sigma_{\log}$, which would be estimated by the standard deviation of the observed distribution of *log-residuals* (i.e., $e_{\log}$ = log X' – log X). This can be rewritten as X' = $X \times \exp(e_{\log})$.

Note that the literature on analog encoding also includes hybrid models (e.g., Nuerk et al., 2001) in which each digit is assumed to be encoded separately as an analog magnitude. The same mathematical equations as above would then apply, only X would refer to the value of a single digit instead of the number.

The Leading Digit Correctness Phenomenon

As argued by both Hinrichs and Novick (1982) and Thompson and Siegler (2010), the leading digit correctness phenomenon is consistent with numbers being encoded as analog magnitudes. Namely, the inherent imprecision of analog encoding would not affect the leading digit as often as it would affect less significant digits. For instance, if the presented number is 376 it takes a recall error of at least 24 to alter the hundreds digit, whereas it only takes a recall error of 4 and 1 to alter the tens digit and the ones digit, respectively.

An Alternative Hypothesis: Nominal Encoding With Value-Independent Mistakes

We shall now present an alternative hypothesis according to which the leading digit correctness phenomenon would be consistent also with nominal encoding. To begin with, we shall follow McCloskey (1992) in proposing that numbers are encoded digit-by-digit using a place-value representation (rather than as a single analog magnitude). Use of a place-value representation has been supported in studies using a number comparison task with multi-digit numbers (Poltrock & Schwartz, 1984). A strong indication that digits in different positions are processed separately is that number comparisons tend to be slower when separate digit comparisons are incompatible (e.g., 47 vs. 62, where 4 < 6 but 7 > 2) than when they are compatible (e.g., 42 vs. 57, where 4 < 5 and 2 < 7). This finding has been obtained both for two-digit numbers (Nuerk, Weger, & Willmes, 2001; Verguts & De Moor, 2005) and three-digit numbers (Korvorst & Damian, 2008).

We further follow Hinrichs and Novick (1982) in proposing that place-value encoding is nominal, that is, digits are stored as labels rather than as magnitudes. There is then no inherent imprecision that can account for recall errors. Instead we propose that errors in recall arise because of mistakes in encoding and/or recall. Mistakes in encoding could be that one or more digits are misread or misheard, but also that one or more digits fail to be encoded due to insufficient attention (e.g., Naveh-Benjamin, Guez, & Marom, 2003). Mistakes in recall could be that different encoded pieces of numerical information are mixed up, or simply that people guess at the digits they are unable to recall. These possible mistakes in encoding and recall all have in common that the value reported for a mistaken digit is independent of the value of the presented digit. Mistaken digits should therefore be uncorrelated with presented digits. This property stands these mistakes apart from the imprecise encoding of an analog representation, for which an imprecisely recalled digit should typically be close to the presented digit. We shall refer to our hypothesis as *nominal encoding with value-independent mistakes*.

Before proceeding we remark that there is another kind of "mistake" in nominal encoding that would in fact yield a relation to the value of the presented digit, namely if the person rounds the presented number before encoding it. However, analysis of the data from Thompson and Siegler (2010) reveal that only about 5% of

recall errors were consistent with rounding of the presented number. We therefore choose not to further elaborate on the possibility of rounding, as it could at best explain only a small minority of observed recall errors.

The Leading Digit Correctness Phenomenon

Of the possible mistakes listed above, recall failure *due to lack of attention* is especially interesting because it could lead to the leading digit correctness phenomenon. To see this, note that the most information about the magnitude of a number is conveyed by the place and value of the leading digit. People who are aware of this property of the positional numeral system should give particular attention to the leading digit. As a consequence, the leading digit should particularly often be correctly recalled.

Modelling Nominal Encoding With Value-Independent Mistakes

We shall now present a formal model of nominal encoding with value-independent mistakes. In contrast to the analog encoding model, we must now describe errors in recall of a given digit position rather than an entire number. We shall use letter i to denote the value of the presented digit (in the position of interest) and letter j to denote the value of the recalled digit (in the same position). We shall assume that for each value of j from 0 to 9 there is a fixed probability q_j that a presented digit is mistakenly recalled as value j. This means that the conditional probability of observing recalled digit j when having presented digit i is assumed to be given by

P($j \mid i$) = q_j + (1 − q)×I$_{j=i}$

where $q = q_0+q_1+...+q_9$ is the probability of a mistake being made at all and I$_{j=i}$ is an indicator variable equal to 1 if $j=i$ and 0 otherwise. The second term says simply that in case no mistake is made in the encoding-recall process the recalled digit will equal the presented digit. The fact that the right-hand side is otherwise independent of the value of the presented number (i) means that mistakes are value-independent.

Fitting this model to a set of recall data amounts to estimating the values of the parameters q_0, q_1,..., q_9. The model implies that the number of digits that are recalled as having value j while presented as any other value should be binomially distributed with parameter q_j. We therefore obtain the maximum-likelihood estimate of q_j as the corresponding observed proportion. Letting f_{ij} denote the observed absolute frequency of digits presented as value i and recalled as value j, the maximum-likelihood estimate is

$$q_j \approx \sum_{i \neq j} f_{ij} \Big/ \sum_{i \neq j} \sum_{k} f_{ik}$$

Distinguishing Between the Hypotheses

Above we have presented two competing hypotheses based on analog and nominal encoding, respectively. These hypotheses differ in what is the primary object of recall. According to the analog encoding hypothesis, the primary object of recall is the magnitude of the number, whereas according to our nominal encoding hypothesis it is separate digits that are primarily recalled. Thus the two hypotheses address two different kinds of data: residuals for the entire number and residuals for each digit, respectively. This means that we cannot make a direct comparison of how well the two hypotheses fit a dataset. The closest comparison we can make is between what the analog encoding hypothesis predicts about the distribution of residuals for entire numbers and what our nominal encoding hypothesis predicts about the distribution of residuals for *the leading digit*, as

the leading digit contains most of the information about the magnitude of the number. Note that the approach of analyzing residuals in order to draw conclusions about the underlying memory representation is similar to that used to investigate what specific strategies and representations that underlie the generation of random numbers (see e.g., Scott, Barnard, & May, 2001).

We use letter "A" to label predictions from the analog encoding hypothesis; the label "Alog" for predictions from the log-compressed version of the analog encoding hypothesis; and label "N" for predictions from the nominal encoding with value-independent mistakes hypothesis.

Predictions About How Residuals Should Depend on the Presented Number

The hypotheses make distinct predictions about how residuals should vary with the presented number or digit. For the analog encoding hypotheses, the predictions are very simple and follow directly from the form of the equations that define the models of analog encoding:

Prediction A-1. According to the model of analog encoding, residuals should be independent of the presented number.

Prediction Alog-1. According to the model of analog encoding with log-compression, log-residuals should be independent of the presented number.

We shall now derive what the nominal encoding hypothesis predicts about how residuals in the leading digit should depend on what leading digit was presented. Let i denote the value of the presented leading digit. As q_j is the probability that i is mistakenly recalled as j, the conditional expected residual E[residual | i] can be written as the following weighted sum of possible recall errors:

E[residual | i] = $(0{-}i)q_0 + (1{-}i)q_1 + ... + (9{-}i)q_9 = (0q_0 + 1q_1 + ... + 9q_9) - iq$

Note that only the last term depends on which digit value i was presented. Thus we have the following prediction.

Prediction N-1. Residuals for the leading digit are expected to decrease linearly with the value of the presented digit, with a slope of $-q = -(q_0+q_1+...+q_9)$.

Predictions About the Relation Between the Standard Deviation of Residuals and Other Parameters

The hypotheses differ in their conception of the relation between the origins of correct and incorrect recall. Whereas our nominal encoding hypothesis attributes incorrect recall to mistakes, the analog encoding hypothesis attributes both correct and incorrect recall to the same imprecise encoding such that the correctness of a particular instance of recall is determined by whether the imprecision turned out to be negligible or substantial in that instance. This difference between the hypotheses has important consequences for their ability to account for data on correct and incorrect recall. Whereas the nominal encoding hypothesis is consistent with any ratio between correct and incorrect recalls, the analog encoding hypothesis predicts that the proportion of correct recalls should be consistent with the overall degree of imprecision in recall exhibited in the data. For instance, a study of children's recall of three-digit numbers (Thompson & Siegler, 2010, Study 3) reported that the hundreds digit (i.e. the leading digit) was recalled incorrectly in 72% of all instances, indicating that the overall degree of imprecision must have been very large, and hence exactly recalled numbers should have been very uncommon. We shall now derive the precise predictions from the analog models for the proportion of correct recalls.

Assuming that the context of the recall task requires the response to be an integer, correct recall is equivalent to $|X' – X| < 0.5$. Denoting the cumulative distribution function of the standard normal distribution by Φ we then obtain the following prediction (note that σ is the standard deviation of the distribution of residuals).

Prediction A-2. The probability of correct recall should be $P_{corr} = \Phi(0.5/\sigma) - \Phi(-0.5/\sigma)$.

To obtain the corresponding prediction from the log-compressed model of analog encoding, we rewrite the condition for correct recall, $|X' – X| < 0.5$, as $|\exp(e_{log}) – 1| < 0.5/X$, which in turn can be rewritten as $\log(1 – 0.5/X) < e_{log} < \log(1 + 0.5/X)$. This yields the following prediction.

Prediction Alog-2. The probability of correct recall of a given number X should be $P_{log-corr}(X) = \Phi(\log(1 + 0.5/X)/\sigma_{log}) – \Phi(\log(1 – 0.5/X)/\sigma_{log})$. Note that, in contrast to prediction A-2, the probability of correct recall here depends on the value of the presented number.

As we discussed above, the nominal encoding hypothesis requires no corresponding relation between the standard deviation of residuals and the proportion of correct recall. Instead, the assumption of value-independence of mistakes requires the patterns of residuals to be described by the ten parameters q_0 through q_9, representing frequencies of certain mistakes in recall, and the standard deviation of residuals should satisfy a certain relation to these parameters. Recall that E[residual | i] was calculated in connection with prediction N-1. Similarly, we have E[residual^2 | i] = $(0–i)^2q_0$ + $(1–i)^2q_1$ + ... + $(9–i)^2q_9$. Let f_i denote the observed absolute frequency of digits presented as i. We then have

$$\mathrm{E}[\text{residual}] = \sum_i f_i \mathrm{E}[\text{residual} \mid i] \ / \sum_i f_i$$

and

$$\mathrm{E}[\text{residual}^2] = \sum_i f_i \mathrm{E}[\text{residual}^2 \mid i] \ / \sum_i f_i$$

Prediction N-2. By definition, the standard deviation of the residuals of the leading digits is the square root of (E[residual^2] – E[residual]^2), for which the predicted value under the nominal encoding hypothesis can be calculated from the values of q_0 through q_9 according to the equations above.

Predictions About Correctness of the Second Digit When the Leading Digit Is Recalled Incorrectly

In case the leading digit is incorrectly recalled, the analog encoding hypothesis attributes this to the imprecision being so large that even the leading digit is affected. This implies that the imprecision in recall has completely overwhelmed the magnitude represented by the second digit, so that no signal from the second digit of the presented number would remain in the recalled number.

Prediction A-3. In case the leading digit is incorrectly recalled, correct recall of the second digit should only be at chance level.

In contrast, the nominal encoding hypothesis says that each digit is encoded separately. If a mistake has occurred in the leading digit, the second digit may still be correctly encoded. (Although the mistake in the leading digit is an indication that the participant's attention to the numerical information may have been low, so that the likelihood of mistakes in the second digit is increased.)

Prediction N-3. In case the leading digit is incorrectly recalled, correct recall of the second digit is still expected to be at above chance level (unless the attention level tends to be very low).

Reanalysis of Data From Thompson and Siegler (2010)

Thompson and Siegler (2010, Study 3) gave 127 second-graders magnitude estimation tasks and a numeric recall task. Here we are interested in the recall task. The children listened to a set of short stories involving a total of 18 numbers, 14 of which were three-digit numbers (164, 237, 419, 487, 524, 548, 625, 632, 725, 759, 817, 846, 938, 962). Every story contained three numbers, which were either all in a lower range (up to 237), or all in a medium range (between 419 and 532), or all in a higher range (from 725 and up). After each story, children were asked to recall the numbers that had been presented. Thompson and Siegler graciously shared the raw data on these recalled numbers with us. We conducted an analysis of these data to test the various predictions presented above.

Analysis

The analysis used the data on recall of presented three-digit numbers, yielding 127 × 14 = 1,778 data points, out of which one data point was empty and 19 data points were excluded because the recalled number had more than three digits. The exclusion of those stimuli that were single-digit or double-digit makes for a cleaner analysis of the nominal encoding model, while causing only negligible changes in the results of analysis of the analog encoding model. (The four excluded stimuli yielded disproportionally many correct answers, and the support for our conclusion would be even stronger if recall of these stimuli were included in the analysis.)

The key parameters for the analog encoding models were estimated to $\sigma = 300$ and $\sigma_{log} = 1.08$, respectively. The estimated values of the key parameters for the nominal encoding model were as follows: $q_0 = .07$, $q_1 = .16$, $q_2 = .07$, $q_3 = 08$, $q_4 = .10$, $q_5 = .10$, $q_6 = .08$, $q_7 = .05$, $q_8 = .03$, $q_9 = .04$, with sum $q = .78$.

As reported by Thompson and Siegler (2010), the data exhibited the leading digit correctness phenomenon. The hundreds digit was correctly recalled in 29% of all data points, 95% CI [27%, 31%], compared to just 15% correctness in the tens digit, 95% CI [13%, 17%], and 13% correctness in the ones digit, 95% CI [11%, 14%]. Throughout, confidence intervals are calculated using the bias-corrected and accelerated (BCa) bootstrap method.

Testing predictions about how residuals should depend on the presented number — The left panel of Figure 1 shows mean residuals for each presented number demonstrating a clearly negative slope. Using linear regression this slope was estimated to -0.62, $p < .001$, 95% CI [-0.68, -0.56]. The right panel of Figure 1 shows the same phenomenon for the residuals in the leading digit, with a negative slope estimated to -0.63, $p < .001$, 95% CI [-0.68, -0.58]. A similar result at a different scale was obtained for log-residuals, with a negative slope estimated to -0.0011, $p < .001$, 95% CI [-0.0013, -0.0009]. These results are clearly inconsistent with analog encoding (predictions A-1 and Alog-1). In contrast, our nominal encoding hypothesis did predict a negative slope of residuals in the leading digit (prediction N-1), specifically a slope of $-q$. Recall that we estimated the value of parameter q to .78. Thus the observed slope of -0.63 was slightly less steep than predicted by the nominal encoding hypothesis. As we shall discuss below, this might be due to a feature of the design of the experiment.
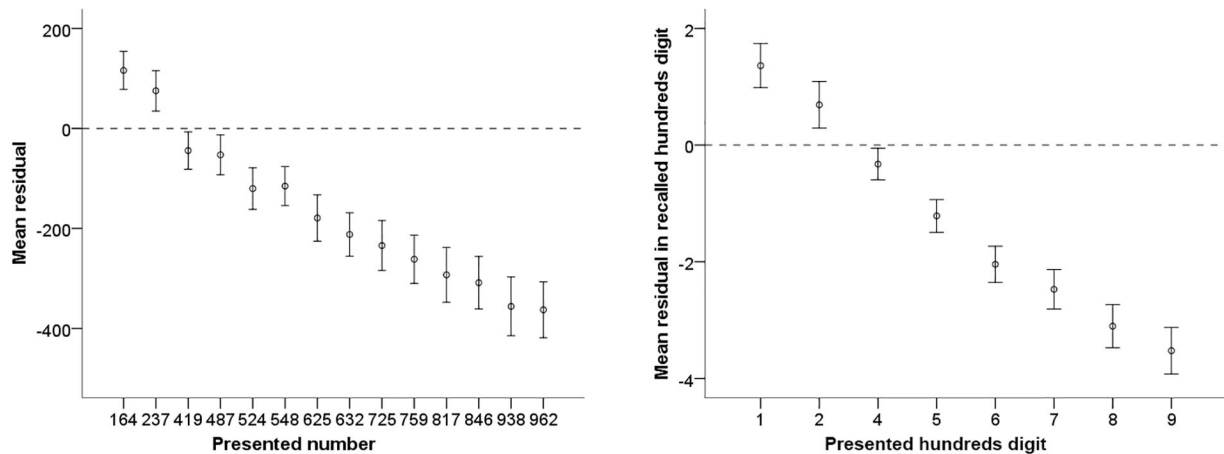
Figure 1. 95% confidence intervals for mean residuals (left) and mean residuals in the leading digit (right). Data from Thompson and Siegler (2010).

Testing predictions about the relation between the standard deviation of residuals and other parameters — Based on the observed standard deviation of residuals (σ = 300), the analog encoding model predicts a probability of correct recall of P_{corr} = 0.13% (prediction A-2). However, the observed proportion of correct responses was twenty times larger, 2.65%, $p < .001$, binomial test. (Although the convention is to simply report "$p < .001$", it is worth noting that the exact value was $p = 10^{-43}$, so it is unthinkable that this large a deviation from the expected proportion of correct responses would occur by chance if the analog encoding model is correct.)

Based on the observed standard deviation of log-residuals (σ_{log} = 1.08), the log-compressed analog encoding model (prediction Alog-2) predicts a range of probabilities of correct recall from 0.23% (for X = 164) down to 0.04% (for X = 962). Even the largest of these probabilities is more than ten times smaller than 2.65%, $p < .001$. Thus, none of the analog encoding models could account for the observed relation between the standard deviation of residuals and the proportion of correct recall.

For the nominal encoding model we instead compared the observed standard deviation of residuals in the leading digit, $\sigma_{leading}$ = 3.00, 95% CI [2.90, 3.10] with the value 3.23 predicted by the frequencies of certain mistakes represented by the estimated values of q_0 through q_9 (prediction N-2). Thus the standard deviation of residuals was slightly smaller than predicted by the nominal encoding hypothesis. Again, this might be due to a feature of the design of the experiment, as discussed below.

**Testing predictions about correctness of the second digit when the leading digit is recalled incorrectly — ** The analog encoding model predicts that correctness of second digits should only be at chance levels in case the leading digit is incorrectly recalled (prediction A-3) whereas the nominal encoding hypothesis predicts correctness as levels above chance unless attention is low (prediction N-3). However, it is difficult to say what exactly would be the chance level, as the possible values for the second digit were not uniformly distributed over the presented numbers (and some digit values are more popular guesses than others). For this reason, we will only carry out a rough analysis. As there are 10 possible digit values, the chance level should be approximately 10%. Out of the 1,254 data points (71%) in which the leading digit was incorrectly recalled, 154 (12%) had the correct value of the second digit. This result seems roughly consistent with analog encoding but

also consistent with nominal encoding under low attention levels. As a large majority of children (71%) made mistakes even on the leading digit, attention levels indeed seem to have been generally low.

Discussion

In this reanalysis of children's recall of three-digit numbers we found that the data tended to support the hypothesis of nominal encoding with value-independent mistakes over the analog encoding hypothesis. Two findings were particularly inconsistent with analog encoding. First, residuals depended strongly on the presented number. Second, the proportion of correct recalls was much too large in relation to the overall low accuracy of recall.

Whereas results were more consistent with nominal encoding, they were not exactly as predicted. This could be due to the recall task being sectioned in several stories, each of which presented several numbers that were always in the same range; either three low numbers, three medium range numbers, or three high numbers. This could cause some mistakes in recall in which one number is mixed up with another number in the same range, which would violate the model assumption of value-independent mistakes.

The Present Study

The reanalyzed dataset was from a study of children who listened to stories and were asked to recall numerical information from these stories. In order to subject the competing hypotheses to more comprehensive testing we conducted two new studies in which adults read a story and were asked to recall numerical information presented in it.

In the data on children's recall, we found that recall may have been influenced by other numbers presented in the same story. To avoid such contamination we ensured that the story in our studies contained only a single piece of numerical information (but lots of other kinds of information). In addition to manipulating the value of the presented number, our studies manipulated two other aspects: the format and the context.

The expected role of format — A reason to manipulate the format is that processing of numbers may depend on whether they are presented in the format of Arabic digits or words (Campbell, 1994; Cohen Kadosh & Walsh, 2009; Gebuis, Kenemans, de Haan, & van der Smagt, 2010). Specifically, we hypothesized that the format might influence the tendency to pay more attention to the leading digit than to less significant digits. Recall that focusing attention on the leading digit makes sense given how the positional numeral system works. We claim that activation of knowledge of the positional numeral system should be more likely when processing the Arabic digits format than the words format. The reason for this claim is that a reader of a multi-digit number given in Arabic digits, say 534, will read it out (silently) as "five hundred and thirty-four" and in order to accomplish this, the reader must necessarily use knowledge of the positional numeral system to map the first digit to hundreds, the second digit to tens and the third digit to ones. In contrast, when reading the words "five hundred and thirty-four" readers need only read each word separately, a task that does not require knowledge of the positional numeral system therefore seems less likely to activate such knowledge. For this reason we expect the words format to lead to less focus on the leading digit, and hence to less recall advantage for the leading digit compared to the less significant digits.

The expected role of context — Previous research investigating numerical judgments more generally has indicated that people's judgments are often influenced by contextual factors such as framing (Tversky &

Kahneman, 1981) and anchoring effects (Tversky & Kahneman, 1974). In the former phenomenon people use the framing of the problem to evaluate the meaning of numbers. They might, for example, find an intervention that saves 200 out of 600 people more appealing than an intervention were 400 out of 600 people will die. In the latter phenomenon people will make estimates of some value by starting from an initial value and then adjusting to get their final answer. According to Tversky and Kahneman (1974) it is reasonable that the starting point may be suggested by the formulation of the problem. More specifically, effects of contextual factors have also been observed in tasks where participants are required to generate random numbers (Scott, Barnard, & May, 2001). Based on our hypothesis that recall errors reflect inattention and guessing rather than imprecision encoding, we should therefore expect errors to be influenced by contextual cues. Specifically, we manipulate whether the number is presented as counting things that usually come in small numbers (baseball caps; Study 1) or in large numbers (grains of sand; Study 2). If some readers have been inattentive to the number of baseball caps or grains of sand mentioned in the story and therefore need to guess, this contextual cue should yield guesses of smaller numbers in Study 1 than in Study 2.

Study 1

To ensure an even distribution over different digits, nine three-digit stimulus numbers were constructed such that each of the digits from 1 to 9 appear exactly once in each of the three positions (193, 217, 348, 426, 534, 651, 782, 869, 975). To avoid contamination between stimuli, participants were presented with only one stimulus number each in a between-subjects design. The stimulus number was presented either using digits or words (e.g., 193 vs. one hundred and ninety-three), for a total of 18 versions (nine numbers times two formats). As we presumed that the rate of recall errors might be low when adults are asked to recall a single number we decided to collect a rather large dataset, 540 data points (30 for each of 18 stimuli).

Method

Participants

Participants were 540 American adults (52% male, Mean age = 32), recruited online among users of Amazon Mechanical Turk (mturk.com) for a fee of ¢50 US. Participants were anonymous and gave their informed consent to participate.

Procedure

The study was conducted online and participants were told that the study would examine certain aspects of memory. Participants were asked to read a one-page excerpt adapted from a comic short story by acclaimed writer Connie Willis, and informed that they would later be asked questions about the story. The page contained just one piece of numeric information, namely, the number of baseball caps that a visiting alien had collected. The piece of numeric information varied across eighteen different versions of the text as described above. The material is included in the Appendix.

The story was followed by a filler task (a six-item questionnaire measuring agreement with relativist and objectivist quotations from various scholars; available on request). The filler task took approximately five minutes to complete. Thereafter, participants were asked to recall various pieces of information presented in the story. Specifically, the first question was how many baseball caps the alien had brought (with the additional instruction "If you don't remember exactly, please give your best estimate").

Analysis

We first examined that our data replicated the leading correctness phenomenon. We then examined the predictions from the competing hypotheses, just as we did in our above reanalysis. Analyses were conducted both for the pooled data and separately for the digits and words conditions. However, to avoid an unnecessarily cumbersome results section we report analyses per condition only where this is really relevant.

Results and Discussion

Seven participants were excluded because they did not fill in a recalled number at all. Another two participants were excluded because they filled in negative numbers, another nine because they filled in numbers with more than three digits (i.e., greater than 999). This left 522 participants for further analysis.

The key parameters for the analog encoding models were estimated to $\sigma = 255$ and $\sigma_{log} = 1.34$, respectively. The estimated values of the key parameters for the nominal encoding model were as follows: $q_0 = .17$, $q_1 = .07$, $q_2 = .04$, $q_3 = 02$, $q_4 = .02$, $q_5 = .02$, $q_6 = .02$, $q_7 = .00$, $q_8 = .00$, $q_9 = .01$, with sum $q = .36$.

The Leading Digit Correctness Phenomenon Depending on Format

First we examined whether our data replicated the leading digit correctness phenomenon (Hinrichs & Novick, 1982; Thompson & Siegler, 2010). We split every presented and recalled number into three digits: the hundreds digit, the tens digit, and the ones digit. Overall, the hundreds digit was correctly reported by 65% of participants, 95% CI [60%, 69%]; the tens digit by 54%, 95% CI [50%, 58%]; and the ones digit by 54%, 95% CI [50%, 59%]. Thus, as expected, the leading digit was correct substantially more often than the tens digit, $p < .001$, McNemar's test.

However, separating the digits condition and the words condition revealed an important difference. The digits condition exhibited a very strong leading digit correctness phenomenon, with 75% correct hundreds digits, compared with 56% correct tens digits and 55% correct ones digits. In contrast, the words format did *not* exhibit the leading digit correctness phenomenon; 54% of the hundreds digits were correct, 52% of the tens digits, and 54% of the ones digits. To assess the difference between the two formats in the recall advantage of the leading digit over the tens digit, we focused on those participants who recalled either the hundreds or the tens digit correctly but not both. In the digits condition there were 70 such participants out of which 61 (87%) recalled the hundreds digit correctly. In the words condition there were 89 such participants out of which only 42 (47%) recalled the hundreds digit correctly. This difference was statistically significant, $\chi^2(1, N = 159) = 21.20$, $p < .001$, $\varphi = .37$, a medium effect.

The difference in results between formats was expected from the hypothesized role of Arabic digits in drawing attention to the first digit. It is possible that the numerical information presented in the digits condition was more salient than the corresponding information in the words format. Note, however, that although such difference in salience could explain a difference in the overall proportion of recalled numbers it cannot explain why the leading digits phenomenon occurs in the digits but not the words condition

Testing Predictions About How Residuals Should Depend on the Presented Number

Figure 2 shows mean residuals for each presented number (left panel) and each presented leading digit (right panel), with similar negative slopes. Also log-residuals showed a similar slope (at a different scale). The strong dependence on presented numbers is inconsistent with the analog encoding hypothesis (predictions A-1 and

Alog-1). For the leading digit residuals the slope was estimated to -0.33, $p < .001$, 95% CI [-0.42, -0.24]. This result is consistent with the nominal encoding hypothesis (prediction N-1), which predicted a slope of -0.36.
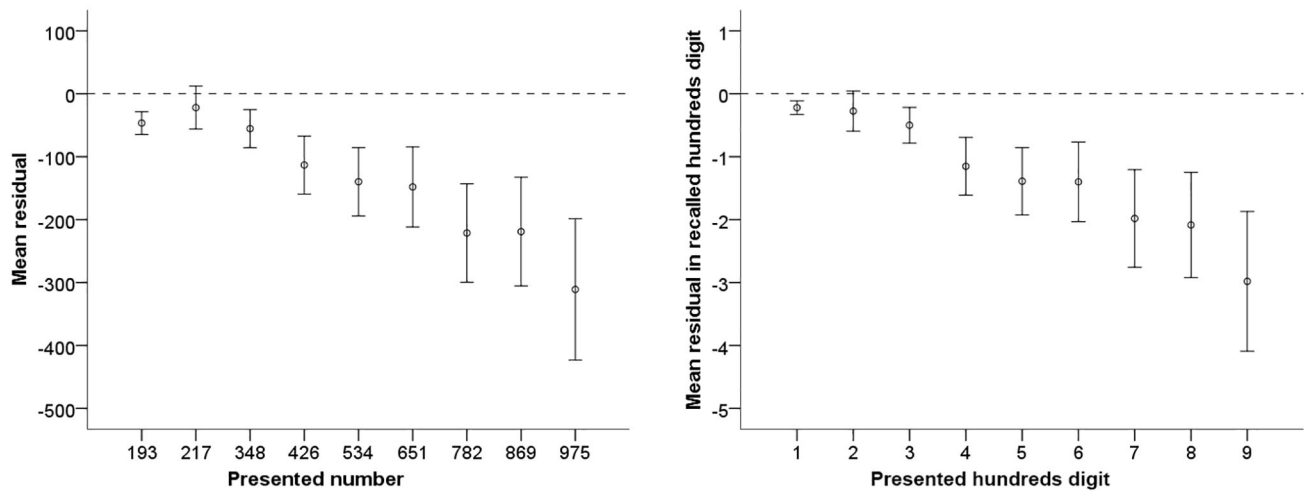

Figure 2. 95% confidence intervals for mean residuals (left) and mean residuals in the leading digit (right) in Study 1.

Testing Predictions About the Relation Between the Standard Deviation of Residuals and Other Parameters

Based on the observed standard deviation of residuals ($\sigma = 255$), the analog encoding model predicts a probability of correct recall of $P_{corr} = 0.16\%$ (prediction A-2), significantly and substantially smaller than the actual proportion of correct responses, 37.9%, $p < .001$. (Indeed, the *p*-value is so extremely small that our software could not distinguish it from zero, which means that it is at least smaller than 10^{-300}.)

Similar results were obtained in separate analyses of the digits condition ($\sigma = 209$, $P_{corr} = 0.19\%$, observed proportion correct = 43.2%) and the words condition ($\sigma = 288$, $P_{corr} = 0.14\%$, observed proportion correct = 32.4%). Also the log-compressed model gave similar results ($\sigma_{log} = 1.34$, $P_{corr} < 0.16\%$ according to prediction Alog-2, observed proportion correct = 37.9%). Thus, just as in the data on children's recall, none of the analog encoding models could account for the observed relation between the standard deviation of residuals and the proportion of correct recall.

Next we compared the observed standard deviation in the leading digit, $\sigma_{leading} = 2.50$, 95% CI [2.28, 2.70], with the value 2.61 predicted by the nominal encoding model (prediction N-2). We conclude that the nominal encoding model could indeed account for the observed relation between the standard deviation of residuals and the observed frequencies of mistakes.

Testing Predictions About Correctness of the Second Digit When the Leading Digit Is Recalled Incorrectly

Because the numbers used as stimuli included all digits from 1 to 9 equally often but never included 0, we examined the 136 instances of recall where the hundreds digit was recalled incorrectly and the tens digit was recalled as a non-zero digit. The probability that the recalled tens digit would be correct by chance is then one chance in nine (11%). The actual frequency of correct recall of the tens digit in this subset of the data was 38%, 95% CI [29%, 46%], which is much higher than expected by chance. We then examined if the results differed between the two formats. In the words format, the tens digit was correct in 42 out of 89 instances of recall

(47%). In the digits format the tens digit was correct only in 9 out of 47 instances of recall (19%). Thus, when the hundreds digit was not correctly recalled the frequency of correct recall of the tens digit was higher for numbers presented in words than for numbers presented in digits, $\chi^2(1, N = 136) = 10.32$, $p = .001$, $\varphi = .28$, a medium effect. This finding is consistent with the hypothesis that, compared to the digits format, the words format shifts the focus of attention from the leading digit to being more evenly spread across the digits.

Conclusion

In Study 1 we replicated the leading digit correctness phenomenon and found the data to be inconsistent with the analog encoding hypothesis. Importantly, the data from our adult population had the same features as the data from the child population in Thompson and Siegler (2010), namely, a marked negative slope for the mean residuals and a high proportion of correctly recall numbers in relation to the overall inaccuracy of recall. These features were well predicted by the nominal encoding model. In support of the nominal encoding hypothesis we also found the format – words or Arabic digits – to influence recall in a manner consistent with the digits format focusing attention on the leading digit. Specifically, the leading digit correctness phenomenon essentially disappeared when numbers were presented in the words format. The analog encoding hypothesis seems unable to account for this finding too.

Finally, note that Study 1 used a contextual cue that presumably influenced recall mistakes. More specifically, it is possible that participants anchored their judgments of the number of baseball caps on previous experiences (Tversky & Kahneman, 1974), namely that baseball caps typically come in small numbers. In Study 2 we replicated Study 1 with another contextual cue to test this proposed effect and to ascertain the robustness of our main findings under different contextual cues.

Study 2

We conducted a second study with the aim of replicating the results of the first study with numbers counting something else. Specifically, the stimulus of Study 1 was the number of baseball caps, whereas in Study 2 we changed the story such that the number instead refers to grains of sand. In addition to testing the robustness of results, this manipulation provides a test of the role of contextual cues for numeric recall.

Method

Participants

Participants were 540 American adults (51% male, Mean age = 34), recruited online among users of Amazon Mechanical Turk (mturk.com) for a fee of ¢50 US.

Procedure

The procedure was identical to that of Study 1 except for one important change. The term "baseball caps" was changed to "grains of sand" both in the short story participants were asked to read and in the subsequent questionnaire.

Results and Discussion

Nine participants were excluded because they did not fill in a number at all, two participants were excluded because they filled in negative numbers, and 23 participants were excluded because the recalled number had more than three digits. This left 506 participants. Note the greater frequency of recalled numbers with more than three digits in Study 2 (23 out of 540) than in Study 1 (10 out 540 in Study 1), $\chi^2(1, N = 1080) = 5.28$, $p = .022$, $\varphi = .07$. The effect, albeit small, is in the expected direction; the context of sand grains elicited more guesses of large numbers than did the context of baseball caps.

The key parameters for the analog encoding models were estimated to $\sigma = 202$ and $\sigma_{log} = 0.93$. The estimated values of the key parameters for the nominal encoding model were as follows: $q_0 = .09$, $q_1 = .04$, $q_2 = .03$, $q_3 = 03$, $q_4 = .01$, $q_5 = .02$, $q_6 = .02$, $q_7 = .01$, $q_8 = .00$, $q_9 = .01$, with sum $q = .27$.

The Leading Digit Correctness Phenomenon

Replicating Study 1, the digits condition exhibited a very strong leading digit correctness phenomenon whereas the words condition did not. In the digits condition there was 83% correct hundreds digits, 95% CI [78%, 87%], compared with only 61% correct tens digits, 95% CI [54%, 67%], and only 61% correct ones digits, 95% CI [55%, 67%]. In the words format condition there was just 67% correct hundreds digits, 95% CI [61%, 73%], which was not significantly different from the 61% correct tens digits, 95% CI [54%, 67%], nor significantly different from the 62% correct ones digits, 95% CI [56%, 68%].

We assessed this difference between the two conditions in the same way as in Study 1. Among those participants who recalled either the hundreds or the tens digit correctly but not both, 65 out of 73 (89%) recalled the hundreds digit correctly in the digits condition, compared to only 43 out of 71 (61%) in the words condition. As in Study 1 this difference between the two formats in the recall advantage of the leading digit over the tens digit was statistically significant, $\chi^2(1, N = 144) = 15.57$, $p < .001$, $\varphi = .33$, a medium effect.

Testing Predictions About How Residuals Should Depend on the Presented Number

Replicating the corresponding findings in Study 1, Figure 3 illustrates how mean residuals for each presented number (left panel) and each presented leading digit (right panel) had similar negative slopes; the same went for log-residuals (at a different scale). For the leading digit residuals the slope was estimated to -0.20, $p < .001$, 95% CI [-0.28, -0.13], consistent with the nominal encoding model prediction of a slope of -0.27. Consistent with the contextual cue encouraging mistaken recall of larger numbers in Study 2 than in Study 1, the observed negative slope was somewhat less steep in Study 2 than in Study 1 (-0.20 vs. -0.33). To assess the statistical significance of the difference in slopes we performed an ANOVA on the pooled data from the two studies, with context (baseball caps or sand grains) as a binary factor. The analysis indeed revealed a significant interaction between context and presented leading digit, $F(1, 1024) = 5.91$, $p = .015$, partial eta-squared = .006, a small effect.
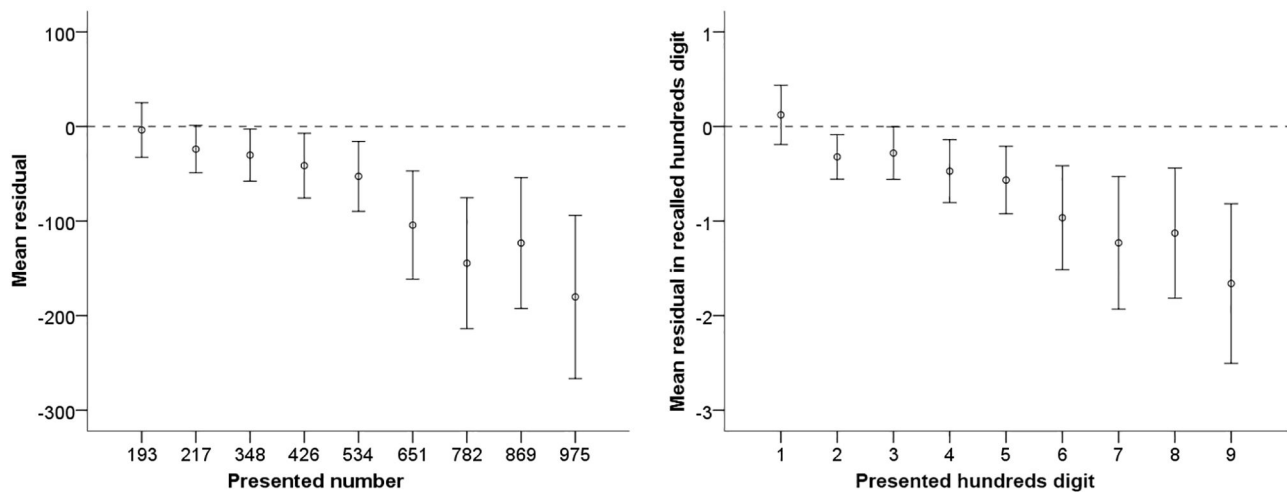
Figure 3. 95% confidence intervals for mean residuals (left) and mean residuals in the leading digit (right) in Study 2.

Testing Predictions About the Relation Between the Standard Deviation of Residuals and Other Parameters

As in Study 1, the proportion of correctly recalled numbers, 46%, was vastly greater than predicted, P_{corr} = 0.20%, by the analog encoding model from the observed standard deviation of residuals, σ = 202. Similar results were obtained in analyses of each condition separately and in analysis based on log-compressed model instead. Also replicating Study 1, the observed standard deviation in the leading digit, $\sigma_{leading}$ = 1.98, 95% CI [1.73, 2.22], was consistent with the value 2.20 predicted by the nominal encoding model.

Testing Predictions About Correctness of the Second Digit When the Leading Digit Is Recalled Incorrectly

As in Study 1 we examined those instances of recall where the hundreds digit was recalled incorrectly and the tens digit was recalled as a non-zero digit (104 data points). Results were similar to the previous study. The frequency of correct recall of the tens digit, given that the hundreds digit was incorrectly recalled, was 38%, 95% CI [28%, 48%], much higher than the chance level of 11%. As in Study 1, this frequency was higher for numbers presented in words (30 out of 68 instances, 44%) than in digits (9 out of 36 instances, 25%), $\chi^2(1, N = 104)$ = 3.67, p = .055, φ = .19, a small-to-medium effect.

Conclusion

Study 2 was designed with two aims in mind. First, we wanted to conduct a replication of Study 1 to evaluate the robustness of our findings. Indeed, we replicated all the main findings. We found strong support for the notion that numbers are stored in memory according to the hypothesis of nominal encoding with value-independent mistakes. Conversely, there is very little support in the data for the analog encoding hypothesis. Our second aim with Study 2 was to evaluate the effect of contextual cues on recall. We found only small effects of contextual cues, but these effects were consistent with the hypothesis that contextual cues may guide guessing when the presented number cannot be accurately recalled.

General Discussion

Numerical information is an important and ubiquitous part of people's day-to-day life. Previous research has primarily focused on the issue of how such information is represented when numbers are mapped from an external representation to an internal meaning (e.g., McCloskey & Macaruso, 1995). Much less attention has been directed towards the interesting topic of how numerical information is represented in long-term memory, when it needs to be retained over an extended period of time.

Here we revisited the leading digit phenomenon, which has previously been taken as an indication that numerical information is stored in memory according to an analog magnitude representation (Hinrichs & Novick, 1982; Thompson & Siegler, 2010). We derived two models of analog encoding, one with linear and one with logarithmic scaling, and evaluated the predictions of these models against a dataset reported by Thompson and Siegler (2010). This analysis revealed that the data was inconsistent with both models of analog encoding. More specifically, neither of the models could account for the proportion of correctly recalled numbers, as implied by the magnitude of the standard deviation of recall errors, or for the observation that the mean residuals varied systematically with the presented number. We also proposed an alternative possibility, the hypothesis of nominal encoding with value-independent mistakes, which made a better account of the data of Thompson and Siegler.

To further investigate this alternative account we conducted two studies that presented participants with a story including a piece of numerical information that varied in numerical value, format (words or digits), and what was counted (baseball caps or sand grains). In both of these studies we found the pattern of results to be consistent with the hypothesis of nominal encoding with value-independent mistakes while the analog model of encoding could not account for the data. We concluded that both the data from Thompson and Siegler (2010) and the data from our two studies supported the notion that numerical long-term memory tends to work by encoding numbers as they are encountered using a place-value representation (McCloskey, 1992), with accuracy of recall determined by attention and guessing. It should be noted that studies supporting analog encoding of numbers have primarily used paradigms where two numbers are to be compared (e.g., Moyer & Landauer, 1967), whereas no such relational judgments were required in our studies. It will be an interesting venue for future research to further investigate the possibility that the signature features of analog encoding become evident only in situations that require relational judgments between numbers.

We acknowledge that our studies, being conducted online, were not as controlled as traditional memory studies. In particular, we had no way of making sure that participants did not take notes. However, for our conclusions this limitation is not very problematic. Cheating by taking notes could not explain the slope of mean residuals (although it could contribute to a high proportion of correct answers). Moreover, we emphasize that we found the same patterns of results in the data from the more tightly controlled study of Thompson and Siegler (2010).

In addition to our main findings we found predicted effects of contextual cues and numerical format. Also these effects are consistent with the hypothesis of nominal encoding with value-independent mistakes. Specifically, contextual cues were expected to affect the guessing part of the recall process, similar to framing and anchoring effects seen in traditional judgment and decision-making tasks (Tversky & Kahneman, 1974, 1981) and random number generating tasks (Scott, Barnard, & May, 2001), whereas format was expected to affect

attention. The effect of format is of particular interest, as there is an ongoing debate about the extent to which the format of numbers affects how they are cognitively processed (e.g., Campbell, 1994; Cohen Kadosh & Walsh, 2009; Gebuis, Kenemans, de Haan, & van der Smagt, 2010). Our results suggest that the digits format tends to focus readers' attention on the leading digit, consistent with its key role in the positional numeral system, whereas numbers presented as words do not. This finding could be examined more in depth in future research, for instance using eye-tracking.

Funding

Competing Interests

The authors declares that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

References

Brainerd, C. J., & Gordon, L. L. (1994). Development of verbatim and gist memory for numbers. *Developmental Psychology, 30*, 163-177. doi:10.1037/0012-1649.30.2.163

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition, 53*, 1-44. doi:10.1016/0010-0277(94)90075-2

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences, 32*, 313-328. doi:10.1017/S0140525X09990938

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

Dehaene, S., Izard, V., Spelke, E. S., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science, 320*(5880), 1217-1220. doi:10.1126/science.1156540

Gebuis, T., Kenemans, J. L., de Haan, E. H. F., & van der Smagt, M. J. (2010). Conflict processing of symbolic and non-symbolic numerosity. *Neuropsychologia, 48*, 394-401. doi:10.1016/j.neuropsychologia.2009.09.027

Hinrichs, J. V., & Novick, L. R. (1982). Memory for numbers: Nominal vs. magnitude information. *Memory & Cognition, 10*, 479-486. doi:10.3758/BF03197651

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews: Neuroscience, 6*, 435-448. doi:10.1038/nrn1684

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*, 1221-1247. doi:10.1016/j.cognition.2007.06.004

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

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition, 44*, 107-157. doi:10.1016/0010-0277(92)90052-J

McCloskey, M., & Macaruso, P. (1995). Representing and using numerical information. *The American Psychologist, 50*, 351-363. doi:10.1037/0003-066X.50.5.351

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

Naveh-Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition, 31*, 1021-1035. doi:10.3758/BF03196123

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*, B25-B33. doi:10.1016/S0010-0277(01)00142-1

Nuerk, H.-C., Wood, G., & Willmes, K. (2005). The universal SNARC effect: The association between number magnitude and space is amodal. *Experimental Psychology, 52*, 187-194. doi:10.1027/1618-3169.52.3.187

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

Scott, S. K., Barnard, P. J., & May, J. (2001). Specifying executive representations and processes in number generation tasks. *The Quarterly Journal of Experimental Psychology: Section A, 54*, 641-664. doi:10.1080/713755994

Thompson, C. A., & Siegler, R. S. (2010). Linear numerical-magnitude representations aid children's memory for numbers. *Psychological Science, 21*, 1274-1281. doi:10.1177/0956797610378309

Thorndike, E. L. (1910). The relation between memory for words and memory for numbers, and the relation between memory over short and memory over long intervals. *The American Journal of Psychology, 21*, 487-488. doi:10.2307/1413353

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124-1131. doi:10.1126/science.185.4157.1124

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453-458. doi:10.1126/science.7455683

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

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, Martin H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science Quarterly, 50*, 489-525.

PsychOpen
publishing psychology

Appendix: Materials Used in Studies 1 and 2

In the text below, XXX stands for a three-digit number that varied between versions within each study. The change from Study 1 to Study 2 is indicated in brackets.

In this study we examine certain aspects of memory. We want you to carefully read the following excerpt adapted from a comic short story by acclaimed writer Connie Willis. Please, do NOT make any notes about the text as that would defeat the purpose of the study. Later in the study we will ask you questions about what you've read.

"You've got to talk to him," Chris said. "I've told him there isn't enough space, but he keeps bringing things home anyway."

"Things?" Stewart said absently. He had his head half-turned as if he were listening to someone out of the holographic image.

"Things. A Buddha, a Persian rug, and XXX baseball caps [grains of sand in Study 2], so far!" Chris shouted at him. "Things I didn't even know they had on Sony. Today he brought home a piano! How did they even get a piano up here with the weight restrictions?"

"What?" Stewart said. The person who had been talking to him moved into the holo-image, focusing as he entered, put a piece of paper in front of Stewart, and then stood there, obviously waiting for some kind of response. "Listen, Chris, darling, can I put you on hold? Or would you rather call me back?"

It had taken her almost an hour to get him in the first place. "I'll hold," she said, and watched the screen grimly as it went back to a two-dimensional wall image on the phone's screen and froze with Stewart still smiling placatingly at her. Chris sighed and leaned back against the piano. There was hardly room to stand in the narrow hall, but she knew that if she wasn't right in view when Stewart came back on the line, he'd use it as an excuse to hang up. He'd been avoiding her for the last few days.

Stewart's image jerked into a nonsmiling one and grew to a full holo-image again. With the piano in here, there wasn't really enough room for the phone. Stewart's desk blurred and dissolved on the keyboard, but Chris wanted Stewart to see how crowded the piano made the hall. "Chris, I really don't have time to worry about a few souvenirs," he said. "We've got real communications difficulties over here with the aliens. The Japanese translation team's been negotiating with them for a space program for over a week, but the Eahrohhs apparently don't understand what it is we want."

"I'm having communications difficulties over here, too," Chris said. "I tell Mr. Ohghhi" She stopped and looked at the alien's name she had written on her hand so she could pronounce it. "Mr. Ohghhifoehnnahigrheeh that there isn't room in my apartment and that he's got to stop buying things, and he seems to understand what I'm saying, but he goes right on buying. I've only got a small apartment, Stewart."

<filler task>

Now try to recall the story excerpt you read earlier.

- How many baseball caps had Mr. Ohghhifoehnnahigrheeh brought to Chris's house? (Enter a number. If you don't remember exactly, please give your best estimate.)

- What was the name of the person Chris talked to?

- What were the Japanese translating team negotiating with the aliens about?

- What did Chris want the other person to see in her apartment?

- When the picture froze, what did it show the person doing?