Published in final edited form as:

Neuropsychologia. 2018 August ; 117: 253–260. doi:10.1016/j.neuropsychologia.2018.05.018.

Your favorite number is special (to you): Evidence for item-level differences in retrieval of information from numerals

Danielle S. Dickson^{a,*} and Kara D. Federmeier^{b,c,d}

^aDepartment of Biology, University of Texas at San Antonio, 1 UTSA Circle Dr., San Antonio, TX 78249, USA

^bDepartment of Psychology, University of Illinois at Urbana-Champaign, 603 E Daniel St., Champaign, IL 61820, USA

^cProgram in Neuroscience, University of Illinois at Urbana-Champaign, 405 N Mathews Ave, Urbana, IL 61801, USA

^dBeckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, 405 N Mathews Ave, Urbana, IL 61801, USA

Abstract

Arabic numerals have come to be used for many purposes beyond representing a particular quantity (e.g., as a label for an athlete on their jersey), but it remains to be determined how this type of meaningfulness is accessed and utilized by readers. Motivated by previous work showing that item-level ratings of personal familiarity can influence traditional indices of memory retrieval, we recorded ERPs while participants read double-digit Arabic numerals (e.g., “65”), presented in a list, and rated whether or not each was familiar/personally meaningful. All numbers repeated after a few intervening trials. The effect of number repetition on the N400 was not impacted by subjective judgments of familiarity, suggesting that all numbers (personally meaningful or not) make initial contact with semantics, facilitating semantic access on second exposure. However, consistent with findings from prior studies of memory for letter strings and visual patterns, there was a late positivity (LPC) on second presentation, selective to numbers rated as familiar. This is the first electrophysiological evidence that readers can use Arabic numerals to guide explicit retrieval of non-numerical information.

Keywords

Numerical cognition; N400; ERPs; Memory; Familiarity

1. Introduction

1.1. General

We encounter Arabic numerals like “81” in many contexts: as a representation of a quantity or numerosity, as a unique identifier (e.g., credit card numbers, confirmation numbers), and

*Corresponding author., danielle.dickson@utsa.edu (D.S. Dickson).

as familiar labels (e.g., sports jerseys) that do not directly involve numerical quantities. Most of the literature on number reading focuses on their use as a symbolic representation of quantity information, with the aim of understanding how people link these symbols to their corresponding numerical meaning. However, the more general use of numerals as items that can contain meaning outside of the context of numerical quantities has been far less explored. In contrast, in the broader text reading literature, there has been great interest in understanding how people comprehend words with multiple possible referents, including ambiguous items like “duck”, which has different meanings across a noun and verb reading, or polysemous items like “university,” which can refer to a concrete place or an abstract entity (e.g., Lee and Federmeier, 2006; Pykkänen et al., 2006). Additionally, some letter strings only hold meaning for a subset of readers (e.g., “VCR” is a meaningful acronym for young adults today, but unlikely to be meaningful to their children) (Laszlo and Federmeier, 2007). Under the assumption that subsets of Arabic numerals can also be uniquely meaningful across individuals, when and how this personalized information is accessible to and appreciated by the reader is not yet understood—indeed, few attempts have even been made to demonstrate that personalized knowledge can play a role at all in number processing (cf., a patient study by Cappelletti et al., 2008).

In order to address how people access meaning and retrieve individualized and specific memories associated with a given numeral, in the present work we ask more directly what happens when people are encouraged to process double-digit Arabic numerals for their broader (potentially non-numerical) meaning, and, in that context, how people’s brain responses to numerals with which they have personalized experience differs from numerals with which they do not claim to have experience. We model our design after successful memory paradigms in the event-related potential (ERP) literature that similarly looked at the effect of familiarity on memory retrieval: critical items are repeated while participants provide ratings of familiarity (reviewed below). We then look at the effect of repetition (i.e., priming) on functionally distinct ERP components, sorted by whether or not a numeral was rated as familiar or not.

1.2. N400 repetition effects

The first measure of interest will be the effect of repetition on the N400, an ERP component linked to initial and fairly automatic semantic memory access (i.e., N400 responses are elicited to potentially meaningful stimuli essentially whenever they are encountered, even when participants are unable to report having perceived the stimuli; Vogel et al., 1998). This classic ERP finding is rather simple: the brain response around 400 ms post-stimulus is less negative in response to items that were previously presented compared to when the same items were initially presented, a phenomenon collectively referred to as the N400 repetition effect (Kutas and Federmeier, 2000). The most straightforward account of N400 repetition effects is that they are due to a reduction in the amount of *new* semantic activity that occurs as a result of prior semantic long-term memory access (Kutas and Federmeier, 2011). That is, the semantic state established with the first presentation of an item persists to some degree over time because semantic features remain active or, on a recent computational model of N400 repetition effects, because the influence of inhibitory connectivity persists, making it harder to reactivate the same semantics (Laszlo and Armstrong, 2014).

In general, it is the case that more conceptually rich stimuli (e.g., words) elicit larger N400 repetition effects than do more conceptually impoverished stimuli (e.g., randomly connected lines) (Nagy and Rugg, 1989; Rugg et al., 1988; Rugg and Nagy, 1989; Rugg, 1990; Van Petten and Senkfor, 1996). In particular, such effects are small or not detectable for stimuli like unfamiliar geometric shapes as well as for unpronounceable strings of letters (e.g., “TXM”) under many task contexts (Rugg and Nagy, 1987; Voss and Paller, 2009b; Voss et al., 2010b). It is possible that Arabic numerals may similarly elicit very small repetition effects. Importantly, however, even among classes of stimuli that have generally been associated with lower levels of rich conceptual information, individual ratings of item-level meaningfulness have been shown to influence the size of N400 repetition effects.

For example, if participants are simply asked to memorize abstract line drawings and shapes, or to judge them for a low-level visual feature, there is no evidence of semantic processing (i.e., no N400 repetition effects; Van Petten and Senkfor, 1996; Voss et al., 2010b; Voss and Paller, 2009b). However, sometimes these shapes can appear to be familiar (e.g., what appears to be a random “squiggle” to one person might appear to be a sketch of a jumping deer to another person), and this link to semantics has been shown to bolster the repetition effect. When individuals are encouraged to think about the shapes’ possible meaningfulness by explicitly rating their familiarity, N400 repetition effects are obtained, selectively for the shapes that individuals rated as familiar (Voss et al., 2010b). Notably, when people are asked to rate letter strings for their familiarity, N400 repetition effects are significant both for meaningful acronyms like “DVD” as well as for letter strings rated as meaningless (Laszlo et al., 2012; cf., Voss et al., 2010a, with a longer retention interval and more intervening items, which led to only wordforms rated as familiar eliciting facilitation on the N400). In general, this work with semantic processing of letter strings and geometric shapes suggests a natural readiness for conceptual engagement with *any* class of stimuli—and the strength of this engagement can sometimes depend on readers’ subjective judgment of their meaningfulness.

Given that numerals rated as familiar might be able to make a more resilient connection with semantic representations than numerals rated as unfamiliar, it is possible that the effect of repetition might differ across numerals as a function of a person’s familiarity with them. At the same time, numerals are a common form of text, whose N400 repetition effects can sometimes be unaffected by semantic ratings. The present study constitutes the first attempt to characterize the N400 response to numerals as a class in an implicit priming context, and, as such, it will be broadly informative about the nature of semantic access for numerals.

1.3. LPC effects as a function of familiarity

Although N400 effects are perhaps the most robust ERP pattern associated with the repetition of meaningful items, there are other ERP responses that may also be affected by our manipulation. These additional effects offer the opportunity to further assess how numbers may be treated similarly or differently from other potentially meaningful stimuli. Specifically, in repetition paradigms, beyond effects on the N400, there are also frequent reports of differences on late positivities (typically an effect on what is conventionally referred to as the “late positive complex” or LPC), which are instead thought to reflect more

explicit and conscious recollective processing of stimuli (e.g., Paller and Kutas, 1992; Paller et al., 1995; Voss and Paller, 2009a). In these cases, items that are rated familiar (or, in traditional memory paradigms, rated as having been recalled with confidence rather than more passively recognized) elicit more positive-going slow waves following the N400 period compared to less familiar items. The strongest evidence for a distinction between N400 and LPC repetition effects comes from patients with hippocampal damage, who have severe impairments in the ability to form and explicitly recall new memories of items and events, although they often retain the ability to learn subconsciously (e.g., there is evidence for implicit priming in amnesic patients) (Warrington and Weiskrantz, 1970, 1982; for review, see Schacter, 1987). In these hippocampal patients, repetition effects can be obtained on the N400, but LPC effects are eliminated (Olichney et al., 2000; Duzel et al., 2001). Also consistent with the interpretation of LPC effects as being dependent on intact hippocampal function and associated with declarative memory, larger LPC effects are typically associated with better behavioral performance (e.g., more successful recollection of the source of a given memory) in recognition tasks (Curran and Cleary, 2003; Duzel et al., 1997; Woodruff et al., 2006; Woroch and Gonsalves, 2010).

Interestingly, in the study showing that N400 repetition effects were observed only for shapes that had been judged to be meaningful, an LPC was also elicited selectively for shapes that were rated as meaningful (Voss, Schendan, & Paller, 2010). In the present experiment, participants are rating their own subjective familiarity with each numeral, and it seems likely that, in order to make these judgments about the personal relevance of a numeral due to prior experiences, recollective processes necessarily occur that should elicit an LPC.

On the basis of the prior literature, then, if LPC effects for numerals are obtained, we expect to see enhanced LPC responses (more positivity, beginning around 500 ms and continuing for a few hundred milliseconds) for repeated items (which are similar to “old” items in explicit memory paradigms). A key novel question is whether this potential LPC enhancement will also be selective (or at least enhanced) to numerals that were rated as subjectively familiar. Such a finding would be consistent with an interpretation that our long-term memories for the subset of Arabic numerals that are uniquely meaningful to us can be deliberately retrieved and impact our subsequent reading of them.

An additional possibility is that there could be a dissociation between the pattern of repetition effects seen on the two components, such that N400 effects do not distinguish between numerals rated familiar and unfamiliar, whereas an interaction occurs selectively on the LPC. This was the case in a study involving complex visual shapes (Voss and Paller, 2009b), in which stimuli were too conceptually impoverished to elicit facilitated N400 responses but still elicited enhanced LPCs during recognition. Therefore, failures to find evidence that Arabic numerals are processed for their distinctive amount of meaningfulness on the N400 would not necessarily mean that there is *no* additional information about aspects of their meaning (e.g., learned associations beyond numeracy) in long-term memory, only that accessing this extra meaning is not fast and obligatory.

In general, then, the pattern of N400 and LPC repetition effects for numerals rated as familiar and unfamiliar will help us better understand how people engage with the meaning and long-term memory representation of these symbols outside of their utilization as a representation of specific numerical value. We will take effects of repetition on the N400 as evidence of eased implicit processing of those items' meaning; such effects may be limited to items rated as familiar, but may also occur for all items regardless of personal familiarity ratings (consistent with obligatory meaningfulness processing). We anticipate finding an LPC for second presentation of items, and are interested in whether its presence is unique to numerals rated as personally familiar, which would be consistent with deliberative memory retrieval for those items.

2. Methods

2.1. Participants

Thirty-two participants were included for analysis in this study (one additional participant was dropped due to failure to comply with instructions). All were right-handed native monolingual English speakers (half female) with normal or corrected-to-normal vision, no early exposure to a second language, history of brain trauma, or current use of psychoactive medications. Participants' average age was 21 (range: 18–29), their mean laterality quotient on the Edinburgh handedness inventory (Oldfield, 1971) was 0.82 (range: 0.43–1.00, with 1.0 indicating strongly right-handed and –1.0 strongly left-handed), and 9 reported having a left-handed biological family member. Informed consent was received from all participants in accordance with the Institutional Review Board at the University of Illinois.

2.2. Stimuli

All double-digit numbers were used as critical stimuli, with fillers being either single- or triple-digit numbers (1–9; 100–116). Doubledigit numbers were used for the practical reason that participants were likely to have pre-experimental familiarity with enough of them to generate a meaningful average ERP. Theoretically, the same effects should obtain in familiarity contrasts with any other numbers. From the perspective of the participant, numbers whose familiarity they were rating ranged from 1 to 116. Critical numbers repeated after 2 or 3 intervening numbers. There was no indication that numbers in the double-digits were of special interest other than that they repeated after some intervening items. To avoid unintentional repetitions of nonidentical numbers, there were no consecutive numbers with matching initial digits (e.g., 90 would not be directly followed by 93), and sequential matches of the second digit were rare. To meet the requirements of this complex lag structure, filler items could not repeat.

Three separate orders of numbers and structures of repetitions were generated, so numbers that had 2 intervening items (either fillers or other critical stimuli) in one list might have 3 intervening items in another list, and numbers that appeared chronologically earlier on one list might appear later on another. Each participant saw one of three list orders.

2.3. Procedure

The critical numbers were presented for 500 ms, followed by a blank data collection period of 800 ms, after which participants were prompted by a question mark to enter their response categorizing the item as Familiar or Unfamiliar to them. Once their response to a numeral was registered, they proceeded to the next trial. Participants could respond as soon as they were able; speed was not emphasized.

After a response was registered, a series of white “~~~~” were presented for 1500 ms during which participants could blink and move their eyes around. A series of red “~~~~” directly followed this blinking period to alert the participant that a number was coming up. To discourage complete predictability of stimulus presentation, this remained on the screen for a random period of time within the range of 400 and 600 ms. A small square was continuously presented in the center of the screen just below the stimuli to remind participants where to keep their eyes during otherwise blank periods.

For a numeral to be categorized as Familiar, participants were instructed to use a liberal and subjective definition of either having specific life experiences with the numeral and/or more generally having a high degree of personal familiarity with it (that did not necessarily need to be linked to a concrete reason). This relatively open description was used to try to generate enough endorsements of meaningfulness to have sufficient trial counts for each participant. All participants were provided the following specific examples of how to think about concrete meaningfulness: apartment numbers, people’s ages, dates and years, and athlete jersey numbers. Responses were monitored during the experiment to ensure compliance with instructions, and some participants had to be encouraged to be more liberal with their judgments during the first block of trials (i.e., they started out rating every numeral as Unfamiliar). Familiarity itself was thus defined internally by each subject, so that definitions of Familiar and Unfamiliar were maximally meaningful within a given subject. This is analogous to how one might sort behavioral trials as fast or slow within-subject to model effects of response speed, even though a given subject’s slow trials might be the same speed as another subject’s fast trials (e.g., Payne and Federmeier, 2017). After the first block of trials, no further instructions were provided. Response hand was counterbalanced with participant sex and stimulus list (with half responding Familiar with their right hand and Unfamiliar with their left hand).

2.4. EEG recording

Brainwaves were recorded using 26 silver/silver-chloride electrodes mounted in an elastic cap. All electrode impedances were kept below 5 k Ω , and electrodes were evenly distributed over the scalp (see Fig. 2 for the configuration). The data were referenced on-line to the left mastoid and rereferenced offline to an average of the right and left mastoids. A separate frontal electrode acted as ground. The vertical electrooculogram (EOG) signal was monitored with an electrode just below the eye (on the infraorbital ridge) and horizontal EOG was monitored with electrodes placed on the outer canthus of each eye. The EEG signal was sampled at 250 Hz and was subjected to an analog bandpass of 0.02–100 Hz during online amplification by Sensorium amplifiers.

Raw waveforms were assessed for inclusion on a trial-by-trial basis with artifact thresholds separately calibrated by visual inspection for each participant. Trials were excluded from averaging if they included blinks, movement artifacts, signal drift, blocking, or a horizontal eye gaze movement. A participant's endorsement of an item as Familiar or Unfamiliar had to be consistent across first and second presentations for the item to be included. This criterion led to the exclusion of an average of 4 items (range: 0–14) per participant that were initially rated Familiar but were rated Unfamiliar upon second judgment. Similarly, for items that were initially rated Unfamiliar, an average of 6 (range: 0–13) were excluded due to being rated Familiar when the item repeated in the list. After artifact rejection, for items rated Familiar consistently (endorsed on both first and second instance), there were 26 trials on average (range: 8–41; median: 27), and for items rated Unfamiliar consistently, there were 42 trials on average (range: 14–70; median: 39). There were only three participants with fewer than 15 trials in one or more condition, and all critical reported effects were also found without these individuals in the sample.

Epochs of EEG data for each trial were taken from 100 ms prior to item onset until 920 ms after item onset, and the baseline acquired over the 100 ms preceding the onset of each trial was subtracted prior to averaging. ERP mean amplitudes were measured after application of a digital bandpass filter of 0.1–20 Hz. For all tests with more than 1 degree of freedom in the numerator, the Huynh-Feldt corrected p -value is reported. In all of the results below, interactions and effects of electrode-related factors are only reported for those comparisons of theoretical relevance.

3. Results

3.1. Behavior: familiarity ratings

Although there were biases in participants' ratings of individual numbers as Familiar or Unfamiliar, most numbers showed variability across participants, meaning that the distribution of numbers in the Familiar and Unfamiliar bins had a high level of overlap. Of the 90 critical numbers, the only number with universal agreement was 21, which all participants consistently rated as personally relevant, presumably due to the salience of the drinking age for this population in the United States. All other numbers received mixed ratings within participants (i.e., changing their responses across first and second rating) and/or between participants. There was a tendency for numbers that were more often consistently rated as Familiar to be lower in numerical value. This is to be expected given that the numerals were preexperimentally familiar and, in fact, lower numbers are more frequent in the world (Raimi, 1976; Dehaene and Mehler, 1992). However, numerical size was not a perfect predictor of rating. For example, the large numbers 50 and 99 were both more frequently rated as Familiar than the numbers 11 and 12.

3.2. Event-related potentials

Fig. 1 shows the response, over all electrode sites, to numerals that were rated as familiar and unfamiliar (here, collapsed across repetition). Following sensory potentials typical of visual presentation, there is a prominent negativity, peaking just before 400 ms (N400), followed by a positivity, larger over posterior sites (LPC).

3.2.1. N400 timewindow—Analysis of the latency of the N400, as measured by identifying the peak of the negativity in the timewindow from 250–500 ms, found that it was maximum at 372 ms ($SE = 1$ ms) post-stimulus onset. There were no effects of Familiarity or Repetition on its latency (F 's < 3.84). In the analyses that follow, N400 amplitude was measured from 300–500 ms.

To assess effects of Repetition and Familiarity on the N400 and to characterize the scalp distribution of any such effects, we performed an omnibus ANOVA with 2 levels of Familiarity (Familiar, Unfamiliar), 2 levels of Repetition (First Presentation, Second Presentation), 2 levels of Hemisphere (Left Hemisphere, Right Hemisphere), 2 levels of Laterality (Lateral, Medial), and 4 levels of Anteriority (from front to back) on the 16 electrode sites that could be classified along all three spatial dimensions (from front to back: left and right lateral prefrontal; left and right medial prefrontal; left and right lateral frontal, left and right medial frontal; left and right medial central, left and right lateral temporal; left and right medial occipital; left and right lateral occipital). This analysis revealed a significant main effect of Familiarity ($F(1,31) = 14.92, p < 0.001; \eta_p^2 = 0.48$), which interacted with the distributional variables of Laterality ($F(1,31) = 12.32, p = 0.001; \eta_p^2 = 0.40$) and Anteriority ($F(3,93) = 10.93, p = 0.002, \epsilon = 0.487; \eta_p^2 = 0.34$), as well as with Laterality and Anteriority together ($F(3,93) = 4.00, p = 0.02, \epsilon = 0.595; \eta_p^2 = 0.13$). There was no main effect of Repetition ($F(1,31) = 1.67, p = 0.21; \eta_p^2 = 0.05$), but there were interactions between Repetition and the distributional variables of Laterality ($F(1,31) = 5.57, p = 0.02; \eta_p^2 = 0.18$) and Anteriority ($F(3,93) = 4.38, p = 0.04, \epsilon = 0.43; \eta_p^2 = 0.14$). Finally, there was an interaction between Familiarity, Repetition, and the three distributional factors ($F(3,93) = 3.06, p = 0.03, \epsilon = 1; \eta_p^2 = 0.10$).

The reported interactions between critical factors and the distributional factors reflected that effects were generally larger at medial sites (generating effects of Laterality), as is relatively typical when a mastoid reference is used. There were also differences in how these effects patterned across frontal and posterior channels (generating effects of Anteriority). To better characterize the effects of Familiarity and Repetition across levels of Anteriority, separate analyses of these factors were performed with the 11 frontal channels and with the 15 posterior channels.

At frontal channels, there was only a main effect of Familiarity ($F(1,31) = 18.23, p < 0.001; \eta_p^2 = 0.59$), and no effect of Repetition ($F(1,31) = 0.01, p = 0.94; \eta_p^2 < 0.001$) or interaction between Repetition and Familiarity ($F(1,31) = 0.46, p = 0.50; \eta_p^2 = 0.01$). The main effect of Familiarity was due to a more positive response to items rated as Familiar ($3.64 \mu V, SE = 0.15$) compared to items rated as Unfamiliar ($2.30 \mu V, SE = 0.12$).

At posterior channels, however, there were both main effects of Familiarity ($F(1,31) = 9.33, p = 0.005; \eta_p^2 = 0.30$) and Repetition ($F(1,31) = 7.67, p = 0.01; \eta_p^2 = 0.25$). There was no

interaction between Familiarity and Repetition ($F(1,31) = 0.82, p = 0.37; \eta_p^2 = 0.03$). The main effect of Familiarity was due to a more positive response to items rated as Familiar ($5.22 \mu\text{V}, SE = 0.11$) compared to items rated as Unfamiliar ($4.50 \mu\text{V}, SE = 0.10$), and the main effect of Repetition was due to Second Presentation items ($5.18 \mu\text{V}, SE = 0.11$) being more positive than First Presentation items ($4.53 \mu\text{V}, SE = 0.10$).¹

In sum, both frontal and posterior channels had an effect of Familiarity that was unrelated to Repetition, with more positivity (or less negativity) for items rated as Familiar than items rated as Unfamiliar (see Fig. 1). In contrast, posterior channels revealed an additional effect of Repetition that was unrelated to Familiarity (with facilitation for repeated items). Waveforms showing the N400 repetition effect at two representative posterior channels are plotted in Fig. 2. The leftmost plots in Fig. 3 display topographical maps of the distribution of the N400 repetition effect, separately for items rated Familiar and Unfamiliar.

3.2.2. LPC timewindow—Analysis of the latency of the late positivity, as measured by identifying the maximum of the positivity in the timewindow from 500–900 ms, found that it was maximum at 611 ms ($SE = 2$ ms) poststimulus onset. There were no effects of Familiarity or Repetition on its latency (F s < 0.74). In the analyses that follow, LPC amplitude was measured from 500–800 ms.

Using the same approach as for the N400 timewindow, to assess effects of Repetition and Familiarity on the LPC and to characterize its scalp distribution, we performed an omnibus ANOVA with 2 levels of Familiarity (Familiar, Unfamiliar), 2 levels of Repetition (First Presentation, Second Presentation), 2 levels of Hemisphere (Left Hemisphere, Right Hemisphere), 2 levels of Laterality (Lateral, Medial), and 4 levels of Anteriority (from front to back). This analysis revealed a main effect of Familiarity ($F(1,31) = 29.73, p < 0.0001; \eta_p^2 = 0.96$), which interacted with the distributional variable of Laterality ($F(1,31) = 16.56, p < 0.001; \eta_p^2 = 0.53$), and with both Laterality and Anteriority ($F(3,93) = 3.81, p = 0.03, \epsilon = 0.68; \eta_p^2 = 0.12$). There was no main effect of Repetition ($F(1,31) = 0.29, p = 0.59; \eta_p^2 = 0.01$), but this factor interacted with the distributional variable of Anteriority ($F(3,93) = 16.27, p < 0.001, \epsilon = 0.415; \eta_p^2 = 0.52$), as well as with Laterality and Anteriority ($F(3,93) = 3.23, p = 0.045, \epsilon = 0.827; \eta_p^2 = 0.10$) and Hemisphere and Laterality ($F(1,31) = 2.07, p = 0.04; \eta_p^2 = 0.15$). Additionally, there was an interaction between Repetition, Familiarity, and the distributional variable of Anteriority ($F(3,93) = 6.60, p = 0.01, \epsilon = 0.396; \eta_p^2 = 0.21$), as well as an interaction between Repetition, Familiarity, and all three distributional factors ($F(3,93) = 2.83, p = 0.04, \epsilon = 0.903; \eta_p^2 = 0.09$).

Following the strategy used for the N400 timewindow, we examined the effects of Familiarity and Repetition and their interactions with Anteriority by performing separate

¹Note that although the Familiar condition had lower trial numbers on average than the Unfamiliar, the standard errors for both frontal and posterior analyses are similar and low, mitigating concerns about signal to noise differences skewing the results.

analyses over the frontal 11 channels and the posterior 15 channels. As in the earlier timewindow, over frontal channels, there was only a main effect of Familiarity ($F(1,31) = 21.87, p < 0.001; \eta_p^2 = 0.71$), wherein items rated as Familiar were more positive ($5.33 \mu\text{V}$, $SE = 0.16$) than were items rated as Unfamiliar ($3.81 \mu\text{V}$, $SE = 0.13$). This effect can be seen in Fig. 1.

However, over posterior channels, there were main effects of Familiarity ($F(1,31) = 21.32, p < 0.001; \eta_p^2 = 0.69$) and Repetition ($F(1,31) = 5.97, p = 0.02; \eta_p^2 = 0.19$), and an interaction between Familiarity and Repetition ($F(1,31) = 5.53, p = 0.03; \eta_p^2 = 0.18$). Thus, in the omnibus ANOVA, the interactions between Familiarity and Repetition with Anteriority largely reflected that, at posterior channels, there was an interaction between Familiarity and Repetition, whereas at frontal channels, there was a main effect of Familiarity that was not impacted by Repetition.

To examine the interaction over posterior channels (which is also where LPC effects tend to be largest), the effect of Repetition was assessed separately for items rated as Familiar and Unfamiliar. For items that were rated Unfamiliar, there was no effect of Repetition ($F(1,31) = 0.53, p = 0.47; \eta_p^2 = 0.02$). However, for items that were rated as Familiar, there was an effect of Repetition ($F(1,31) = 7.90, p = 0.01; \eta_p^2 = 0.25$), with Familiar Second Presentation (repeated) items being more positive ($6.07 \mu\text{V}$, $SE = 0.17$) than Familiar First Presentation (unrepeated) items ($4.82 \mu\text{V}$, $SE = 0.18$). The rightmost plots in Fig. 3 display topographical maps of the distribution of the LPC effect, separately for items rated Familiar and Unfamiliar, and the waveforms can be seen in Fig. 2.

4. Discussion

Our primary goal was to examine when and how people assess the meaning of numerals, beyond the basic quantity information that those numerals represent. As such, we asked participants to provide subjective ratings of familiarity for double-digit Arabic numerals and examined the impact on participants' (1) conceptual processing of them (indexed by N400 repetition effects), and (2) explicit retrieval of detailed information about them (indexed by LPC repetition effects). Importantly, these effects were assessed within-subject (based on their subjective ratings of personal experience with Arabic numerals), so that, across participants, any particular numeral varied in whether it was categorized as familiar or unfamiliar.

We observed a significant N400 repetition effect for numerals. This effect had the canonical (centroparietal) distribution and timecourse (peaking around 370 ms) for N400 effects as reported in studies using words and other types of text, suggesting that the kind of semantics people engage with when reading numerals can be similar to what is engaged during word reading. Further, critically, there was not a difference in this effect across levels of familiarity: numerals rated familiar and numerals rated unfamiliar elicited indistinguishable (small) N400 repetition effects (see Fig. 2 for these waveforms). This replicates findings from a similar task with words and letter strings (Laszlo et al., 2012). There, even

unpronounceable strings of letters rated as unfamiliar (“TXQ”) elicited N400 repetition effects when participants were judging whether or not they were meaningful (i.e., whether or not they were an acronym). Interestingly, this is not what has been observed when participants rated novel shapes (squiggles)—only shapes rated as familiar elicited N400 repetition effects (Voss et al., 2010b). That is, unlike truly pre-experimentally novel geometric shapes, double-digit numbers were similarly processed for their meaningfulness regardless of how participants eventually subjectively rated them. This might indicate that having long-term, pre-experimental experience extracting meaning from a stimulus class renders even those items rated as unfamiliar more meaningful than items that are similarly rated as unfamiliar but for which participants have no pre-experimental experience.

Interestingly, although the N400 repetition effect itself was not affected by ratings of subjective familiarity, there nevertheless was a substantial effect of familiarity that began at the same time as the N400. This frontally-distributed, long-lasting effect was present during both first and second presentation of numerals and took the form of a sustained negativity for numerals that participants self-rated as unfamiliar compared to numerals that they self-rated as familiar (see Fig. 1). A first-pass interpretation of this effect might attribute it to the “FN400” commonly reported in studies of memory. In these reports, words recognized as old at test (familiar words) elicit less negative-going ERPs over frontal channels compared to words categorized as new at test (unfamiliar words). This is essentially a frontally distributed N400 repetition effect, which appears likely to reflect conceptual priming (Voss and Paller, 2009b; Voss and Federmeier, 2011). Additionally, the effect can be extended beyond the N400 timewindow (as in the present study), and it is less clear what that continued facilitation signifies (see, e.g., Wilding and Rugg, 1996; Woodruff et al., 2006).

Two major issues with interpreting the present main effect of familiarity as a frontally-distributed N400 of the type reported in the memory literature is that the effect was not impacted by repetition (which would be expected of an N400 effect), and that we obtained it on first presentation (whereas in memory paradigms it is always reported on second presentation, during recall tests). That is, if the obtained effect truly reflected the kind of conceptual priming thought to be giving rise to the (F)N400 effect in memory paradigms, then it should have been similarly affected by the priming induced by repetition. Because it was not, we suspect the difference between numerals rated familiar and unfamiliar must instead be caused by a different mechanism.

One possibility is that the frontal negativity for unfamiliar items is related to the family of sustained negativities that are often associated with differences in working memory load (with larger sustained negativities for larger loads; e.g., Ruchkin et al., 1995; Chao and Knight, 1997). For example, sustained negativities are a feature of difficult-to-parse sentences compared to less demanding sentences (King and Kutas, 1995). Sustained, frontally-distributed negativities have also been observed to words that are ambiguous (e.g., “duck”, which has different meanings across its noun and verb senses) compared to unambiguous words (e.g., the noun “beer”) (Federmeier et al., 2000; Lee and Federmeier, 2006). Under this interpretation, then, numerals that elicit more disparate associations might be more likely to be rated as unfamiliar. Such an interpretation would also suggest that reading numerals with the aim of sorting them into meaningfulness categories involves a

similar disambiguation process as is induced by word reading. Clearly, more work is necessary to fully assess the possible relationship among word ambiguity effects, working memory processes, and the presently reported effects of familiarity judgment. Whether the present effect is ultimately found to be related to working memory or not, perhaps the most interesting feature of the mechanisms involved in this effect is that they are independent of item repetition. That is, to the extent that this measure indicates some kind of processing effort, it is just as demanding to categorize unfamiliar items the second time as it was the first time they were encountered.

In contrast, there was a different effect that was both influenced by repetition and that distinguished between items rated familiar and items rated unfamiliar: the late positive complex (LPC). This effect was selectively enhanced (i.e., more positive) for second presentation of numerals rated as familiar (see Fig. 2 for waveforms and Fig. 3 for the topography of this effect). Items that were rated as unfamiliar only had an effect of repetition on the N400. One way to view the distinction between this pattern of N400 repetition effects (which were present regardless of how the item was rated) and LPC repetition effects (which were present only for numerals rated familiar) is that, whereas the N400 repetition effect measure primarily reflected *an attempt* to access numeral meaning, the LPC instead reflected actual retrieval success.

This perspective is consistent with the interpretation of LPC effects in memory paradigms, wherein ERPs are recorded during a test in which old (previously viewed) and new (previously unseen) items are classified by participants. There, larger LPCs are typically associated with more successful recollection of old (repeated) items, especially when compared with successfully rejected new items or falsely rejected old items. Additionally, retrieval efforts that yield more explicit details about previously viewed items (e.g., which list they were seen on) and retrieval effects which result in higher-confidence judgments of an item being old are both associated with larger LPCs (see Voss and Paller, 2008, for review).

Our current finding is also strikingly similar to what has previously been observed when participants performed familiarity ratings for novel shapes: only novel shapes rated as familiar elicited LPC effects during a subsequent memory test (Voss et al., 2010b). Of note is that there is a substantial difference between our paradigm and these memory paradigms. Whereas traditional memory paradigms clearly require more recollection and retrieval during the test (which is always second presentation) than during the encoding phase (first presentation), in our paradigm, first and second presentations both involved a memory test of sorts (to the extent that generating a rating of familiar/ unfamiliar involved explicit long-term memory retrieval).

It is interesting to see that this enhanced LPC selective to the second presentation of familiar items still persisted in the current task design. This suggests that the recollection processes that occurred during first presentation for numerals that people rated as familiar somehow selectively facilitated the same retrieval processes at second presentation—either by augmenting the amount of information recollected or by rendering the judgment more confident (both of which are mechanisms that are associated with increased LPC size).

Given that trials were only included if participants' judgments were consistent (e.g., if a participant only rated an item as familiar on second encounter, it was dropped from analysis), both interpretations (confidence/increased retrieval) are plausible—it seems unlikely that one would retrieve less information or be less confident when rating a numeral as familiar for the second time (especially in a relatively short amount of time). Most importantly, to our knowledge, this is the first piece of evidence linking a traditional measure of recollection (LPC) not only to numerals, but to numerals that are specifically perceived by individuals as familiar due to personal (and not explicitly numerical) experience.

In sum, we found evidence that people are able to use numerals to access a different set of information than is typically studied (i.e., not specifically related to the numerosity or number value associated with the numeral as a symbol), manifesting across multiple effects when numerals were sorted by participants' subjective ratings of familiarity. There was a sustained frontal negativity larger for numerals rated unfamiliar, and there was also an LPC repetition effect larger for numerals rated familiar. However, some aspects of numeral processing were not sensitive to the act of making these familiarity judgments. Specifically, the size of the N400 repetition effect was unaffected by whether people rated the numerals as familiar or not, suggesting that the amount of meaning that is more effortlessly accessed from numerals is relatively consistent across items. These findings more broadly suggest that potentially non-numerical information related to our personal experience with Arabic numerals (e.g., apartment numbers, legally relevant age markers like 18 and 21, and numbers associated with sports figures)—a relatively unexplored domain of numerical cognition—can importantly shape the processing elicited when those numerals are encountered.

Acknowledgments

Author notes

The authors would like to thank Dan Hyde, Monica Fabiani, and Chuck Hillman for their feedback on previous versions of this manuscript. We would also like to acknowledge Jessica Kosapatti and Svetlana Sadovskaia for contributing to data entry and collection. This work was supported by a James S. McDonnell Foundation Scholar Award to K.D.F. and the National Institutes of Health under grant AG026308.

References

- Cappelletti M, Jansari A, Kopelman M, Butterworth B, 2008 A case of selective impairment of encyclopaedic numerical knowledge or 'when December 25th is no longer Christmas day, but '20+ 5' is still 25'. *Cortex* 44 (3), 325–336. 10.1016/j.cortex.2006.07.005. [PubMed: 18387561]
- Chao LL, Knight RT, 1997 Prefrontal deficits in attention and inhibitory control with aging. *Cereb. Cortex* 7 (1), 63–69. 10.1093/cercor/7.1.63. [PubMed: 9023433]
- Curran T, Cleary AM, 2003 Using ERPs to dissociate recollection from familiarity in picture recognition. *Cogn. Brain Res* 15 (2), 191–205. 10.1016/s0926-6410(02)00192-1.
- Dehaene S, Mehler J, 1992 Cross-linguistic regularities in the frequency of number words. *Cognition* 43 (1), 1–29. 10.1016/0010-0277(92)90030-L. [PubMed: 1591901]
- Duzel E, Yonelinas AP, Mangun GR, Heinze HJ, Tulving E, 1997 Event-related brain potential correlates of two states of conscious awareness in memory. *Proc. Natl. Acad. Sci. USA* 94 (11), 5973–5978. 10.1073/pnas.94.11.5973. [PubMed: 9159185]
- Duzel E, Vargha-Khadem F, Heinze HJ, Mishkin M, 2001 Brain activity evidence for recognition without recollection after early hippocampal damage. *Proc. Natl. Acad. Sci. USA* 98 (14), 8101–8106. 10.1073/pnas.131205798. [PubMed: 11438748]

- Federmeier KD, Segal JB, Lombrozo T, Kutas M, 2000 Brain responses to nouns, verbs and class-ambiguous words in context. *Brain* 123 (12), 2552–2566. 10.1093/brain/123.12.2552. [PubMed: 11099456]
- King JW, Kutas M, 1995 Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *J. Cogn. Neurosci* 7 (3), 376–395. 10.1162/jocn.1995.7.3.376. [PubMed: 23961867]
- Kutas M, Federmeier KD, 2000 Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci* 4 (12), 463–470. 10.1016/S1364-6613(00)01560-6. [PubMed: 11115760]
- Kutas M, Federmeier KD, 2011 Thirty years and counting: finding meaning in the N400 component of the event related brain potential (ERP). *Annu. Rev. Psychol* 62, 621–647. 10.1146/annurev.psych.093008.131123. [PubMed: 20809790]
- Laszlo S, Armstrong BC, 2014 PSPs and ERPs: applying the dynamics of post-synaptic potentials to individual units in simulation of temporally extended event-related potential reading data. *Brain Lang* 132, 22–27. 10.1016/j.bandl.2014.03.002. [PubMed: 24686264]
- Laszlo S, Federmeier KD, 2007 Better the DVL you know: acronyms reveal the contribution of familiarity to single-word reading. *Psychol. Sci* 18 (2), 122–126. 10.1111/j.1467-9280.2007.01859.x. [PubMed: 17425530]
- Laszlo S, Stites M, Federmeier KD, 2012 Won't get fooled again: an event-related potential study of task and repetition effects on the semantic processing of items without semantics. *Lang. Cogn. Process* 27 (2), 257–274. 10.1080/01690965.2011.606667. [PubMed: 22518068]
- Lee C, Federmeier KD, 2006 To mind the mind: an event-related potential study of word class and semantic ambiguity. *Brain Res* 1081 (1), 191–202. 10.1016/j.brainres.2006.01.058. [PubMed: 16516169]
- Nagy ME, Rugg MD, 1989 Modulation of event-related potentials by word repetition: the effects of inter-item lag. *Psychophysiology* 26 (4), 431–436. 10.1111/j.1469-8986.1989.tb01946.x. [PubMed: 2798692]
- Oldfield RC, 1971 The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9 (1), 97–113. 10.1016/0028-3932(71)90067-4. [PubMed: 5146491]
- Olichney JM, Van Petten C, Paller KA, Salmon DP, Iragui VJ, Kutas M, 2000 Word repetition in amnesia. electrophysiological measures of impaired and spared memory. *Brain* 123 (9), 1948–1963. 10.1093/brain/123.9.1948. [PubMed: 10960058]
- Paller KA, Kutas M, 1992 Brain potentials during memory retrieval provide neurophysiological support for the distinction between conscious recollection and priming. *J. Cogn. Neurosci* 4 (4), 375–392. 10.1162/jocn.1992.4.4.375. [PubMed: 23968130]
- Paller KA, Kutas M, McIsaac HK, 1995 Monitoring conscious recollection via the electrical activity of the brain. *Psychol. Sci* 6 (2), 107–111. 10.1111/j.1467-9280.1995.tb00315.x.
- Payne BR, Federmeier KD, 2017 Pace yourself: intraindividual variability in context use revealed by self-paced event-related brain potentials. *J. Cogn. Neurosci* 29 (5), 837–854. 10.1162/jocn_a_01090. [PubMed: 28129064]
- Pylkkänen L, Llinás R, Murphy GL, 2006 The representation of polysemy: MEG evidence. *J. Cogn. Neurosci* 18 (1), 97–109. 10.1162/089892906775250003. [PubMed: 16417686]
- Raimi RA, 1976 The first digit problem. *Am. Math. Mon* 83 (7), 521–538.
- Ruchkin DS, Canoune HL, Johnson R, Ritter W, 1995 Working memory and preparation elicit different patterns of slow wave event-related brain potentials. *Psychophysiology* 32 (4), 399–410. 10.1111/j.1469-8986.1995.tb01223.x. [PubMed: 7652117]
- Rugg MD, 1990 Event-related brain potentials dissociate repetition effects of high-and low-frequency words. *Mem. Cogn* 18 (4), 367–379. 10.3758/bf03197126.
- Rugg MD, Nagy ME, 1987 Lexical contribution to nonword-repetition effects: evidence from event-related potentials. *Mem. Cogn* 15 (6), 473–481. 10.3758/bf03198381.
- Rugg MD, Nagy ME, 1989 Event-related potentials and recognition memory for words. *Electroencephalogr. Clin. Neurophysiol* 72 (5), 395–406. 10.1162/jocn.1992.4.1.69. [PubMed: 2469564]

- Rugg MD, Furda J, Lorist M, 1988 The effects of task on the modulation of event-related potentials by word repetition. *Psychophysiology* 25 (1), 55–63. 10.1111/j.1469-8986.1988.tb00958.x. [PubMed: 3353486]
- Schacter DL, 1987 Implicit expressions of memory in organic amnesia: learning of new facts and associations. *Hum. Neurobiol* 6 (2), 107–118. [PubMed: 3624022]
- Van Petten C, Senkfor AJ, 1996 Memory for words and novel visual patterns: repetition, recognition, and encoding effects in the event-related brain potential. *Psychophysiology* 33 (5), 491–506. 10.1111/j.1469-8986.1996.tb02425.x. [PubMed: 8854736]
- Vogel EK, Luck SJ, Shapiro KL, 1998 Electrophysiological evidence for a postperceptual locus of suppression during the attentional blink. *J. Exp. Psychol.: Human. Percept. Perform* 24 (6), 1656–1674. 10.1037/0096-1523.24.6.1656. [PubMed: 9861716]
- Voss JL, Paller KA, 2008 Neural substrates of remembering—electroencephalographic studies. In: *Learning and Memory: A Comprehensive Reference* 3 Elsevier, Oxford, pp. 79–97. 10.1016/b978-012370509-9.00106-6.
- Voss JL, Federmeier KD, 2011 FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology* 48 (4), 532–546. 10.1111/j.1469-8986.2010.01085.x. [PubMed: 20701709]
- Voss JL, Paller KA, 2009a An electrophysiological signature of unconscious recognition memory. *Nat. Neurosci* 12 (3), 349–355. 10.1038/nn.2260. [PubMed: 19198606]
- Voss JL, Paller KA, 2009b Remembering and knowing: electrophysiological distinctions at encoding but not retrieval. *NeuroImage* 46 (1), 280–289. 10.1016/j.neuroimage.2009.01.048. [PubMed: 19457375]
- Voss JL, Lucas HD, Paller KA, 2010a Conceptual priming and familiarity: different expressions of memory during recognition testing with distinct neurophysiological correlates. *J. Cogn. Neurosci* 22 (11), 2638–2651. 10.1162/jocn.2009.21341. [PubMed: 19702474]
- Voss JL, Schendan HE, Paller KA, 2010b Finding meaning in novel geometric shapes influences electrophysiological correlates of repetition and dissociates perceptual and conceptual priming. *NeuroImage* 49 (3), 2879–2889. 10.1016/j.neuroimage.2009.09.012. [PubMed: 19770053]
- Warrington EK, Weiskrantz L, 1970 Amnesic syndrome: consolidation or retrieval? *Nature* 228 (5272), 628. 10.1038/228628a0. [PubMed: 4990853]
- Warrington EK, Weiskrantz L, 1982 Amnesia: a disconnection syndrome? *Neuropsychologia* 20 (3), 233–248. 10.1016/0028-3932(82)90099-9. [PubMed: 7121792]
- Wilding EL, Rugg MD, 1996 An event-related potential study of recognition memory with and without retrieval of source. *Brain* 119 (3), 889–905. 10.1093/brain/119.3.889. [PubMed: 8673500]
- Woodruff CC, Hayama HR, Rugg MD, 2006 Electrophysiological dissociation of the neural correlates of recollection and familiarity. *Brain Res* 1100, 125–135. 10.1016/j.brainres.2006.05.019. [PubMed: 16774746]
- Woroch B, Gonsalves BD, 2010 Event-related potential correlates of item and source memory strength. *Brain Res* 1317, 180–191. 10.1016/j.brainres.2009.12.074. [PubMed: 20051237]

Main effect of number rating

Fig. 1.

Grand average waveforms for the response to the presentation of numerals rated familiar and unfamiliar are plotted. In frontal channels (top of figure), an effect of familiarity can be seen whereby numerals that were rated as unfamiliar across presentations elicited a long-lasting negativity compared to numerals that were rated familiar.

Fig. 2.

Grand average waveforms are plotted showing the effect of repetition in left and right posterior channels (starred on the channel configuration plot at left). The solid box denotes the N400 timewindow, whereas the dashed box denotes the timewindow of the late positive complex (LPC). The top plot corresponds to repetition effects for numerals subjectively rated as familiar, reflected in the difference between second presentation compared to first presentation, and the bottom plot shows the same effects for numerals subjectively rated as unfamiliar. The LPC for second presentation is only for numerals rated as familiar, not numerals rated unfamiliar. The N400 repetition effect is small but present in all plots.

Fig. 3.

Topographical maps of the distribution of the effect of repetition, as measured by the subtraction of the second presentation of numerals from first presentation (viewed over the top of the head), are displayed. Darker areas correspond to more substantial effects of repetition. On the left are repetition effects for the N400 timewindow (measured from 300–500 ms post stimulus onset). On the right are repetition effects during the LPC timewindow (measured from 500–800 ms post stimulus). It can be seen that there is an effect of repetition in N400 timewindow for numerals rated familiar and unfamiliar, but that in the LPC timewindow there is only a repetition effect for numerals rated familiar.