Conceptual Integration in Arithmetic is the Same for Digits and for Words: It's the Meaning, Stupid!

Kristie J. Fisher (kjfisher@u.washington.edu)
University of Washington, Department of Psychology
Box 351525, Seattle, WA 98195-1525 USA

Miriam Bassok (mbassok@u.washington.edu)
University of Washington, Department of Psychology
Box 351525, Seattle, WA 98195-1525 USA

Lee Osterhout (losterho@u.washington.edu)
University of Washington, Department of Psychology
Box 351525, Seattle, WA 98195-1525 USA

Abstract

Research on people's arithmetic knowledge (e.g., *2 + 5 = 7*) suggests that it is organized and accessed in a way analogous to other forms of conceptual knowledge (e.g., *tulips and daisies are flowers*). Evidence for this claim comes in part from research that recorded event-related brain potentials (ERPs) and found that incorrect arithmetic facts evoke a response that is analogous to the "N400" response evoked by semantically incongruous words in sentences. Some researchers debate this conclusion by pointing out various differences between responses to arithmetic and language stimuli as well as differences among the studies on arithmetic. These differences could be due to variations in methodology, properties of the stimuli (digits vs. words), or properties of the semantic networks in question (language vs. arithmetic). To examine these possibilities, we elicited the N400 effect in arithmetic by closely following the ERP methodology used in language research and varying the presentation format of the arithmetic stimuli (*12 + 3 = 15* or *Twelve plus three equals fifteen*). A comparable N400 effect was elicited by incorrect answers in both presentation format conditions. We conclude that the N400 incongruency effect in arithmetic is analogous to that in language. However, the peak of the arithmetic incongruency effect occurred about 100ms earlier than is typically observed in language. We suggest that this onset difference is due to differences in the size and the constraints of the arithmetic and the language networks.

Keywords: arithmetic, conceptual integration, N400, ERP

Introduction

People excel at comprehending meaning. They posses an extensive store of highly interrelated conceptual knowledge in a variety of domains, and they access this knowledge in a way that fits the context of the situations they encounter. Research on language comprehension has documented the rapid and fluent way in which people coordinate the meanings of words with the meaning of preceding words, sentences, and paragraphs (e.g., Kutas & Hillyard 1980, 1984). Indeed, it is well established that the first words of a sentence begin immediately to constrain its meaning and, consequently, determine what following words could and could not be appropriate for maintaining conceptual consistency. For example, a sentence that begins with the words "She likes to go to ___" is conceptually constrained to places where an individual could conceivably like to go, such as "the park," "the theater," etc. If the sentence continues with a word or string of words that is conceptually incompatible, such as "fork," "blue," etc., the reader may have trouble integrating such words with the initial meaning of the sentence. This process of attempting to coordinate the meanings of a particular item, in this case a word, with the meanings of preceding items, in this case a string of other words, is known as *conceptual integration*.

Conceptual Integration & ERPs in Language

The process of conceptual integration in language has been studied extensively using electrophysiology. More specifically, researchers have used the technique of event related brain potentials (ERPs). Kutas and Hillyard (1980, 1984) were the first to use this technique to study conceptual integration in language. They found that, under most conditions, words elicit a negative-going waveform that peaks between 300-500ms after the word is presented. The amplitude of this component is negatively correlated with the semantic congruity between a word and its preceding context—the less congruent it is, the larger the negative-going amplitude of the N400. The difference in the negative-going amplitudes of incongruent and congruent words at this time window is known as the *N400 effect*[1]. To illustrate, in the sentences "The cat will *bake* the food" and "The cat will *eat* the food," "bake" will produce a much larger negative going waveform than will "eat." Furthermore, the N400 effect is modulated by the degree of contextual incongruence of the word in question. The incongruent sentence, "He takes milk and sugar in his *dog*" would elicit an N400 effect of greater magnitude than "He

[1] Note that when a word comes at the end of a sentence, often no baseline N400 wave is observed (e.g., Jost, Henninghausen, & Rösler, 2004). However, an "N400 effect" can still occur in that an incongruent word still elicits a large negative wave relative to the correct condition.

takes milk and sugar in his *juice*" when compared with the congruent sentence "He takes milk and sugar in his *coffee*."

The N400 incongruency effect in language is very robust, and numerous researchers have replicated it over the last two decades. In all cases, different violations of semantic congruity change the amplitude of the N400 effect. Though semantic violations reliably affect the amplitude in various ways, the time window at which this effect occurs is very stable (between 300-500ms). It is also important to note that non-semantic sentence anomalies, such as grammatical errors (e.g., "The cat will *eating* the food") (Osterhout & Holcomb, 1992), or irregular, "surprising" surface features (e.g., "The cat will *EAT* the food") do not elicit the N400 effect (Osterhout, McKinnon, Bersick, & Corey 1996). Based on such findings, it is inferred that the N400 effect uniquely reflects the process of conceptual integration as it pertains to semantic meaning.

Conceptual Integration & ERPs in Other Domains

Other researchers have elicited the N400 effect in non-linguistic domains. Researchers have reported N400 effects elicited by incongruent stimuli in pictures (e.g., McPherson & Holcomb, 1999), and films of real-world events (e.g., Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008). An N400 effect can also be elicited in priming paradigms with a variety of stimuli (e.g., Holcomb 1993). Also observed in these domains are modulations of the N400 effect under different degrees of semantic incongruity that seem comparable to modulations of the effect in language. However, the morphology, topographical distribution, and the onset of the N400 effect in all these non-language studies do have some variations. These variations raise the questions of whether all of these effects are indeed versions of the N400 incongruency effect and whether or not the N400 effect can be considered an amodal index of conceptual integration (Kutas & Federmeier, 2000).

More recently, arithmetic processing has also been investigated with ERP methodology. A strong analogy has been drawn between the memory networks of language and mathematical concepts (see Ashcraft, 1992 for a review). These network models of arithmetic have been verified with many studies that have shown priming and interference effects similar to those found in language (e.g., Campbell, 1990; LeFevre, Bisanz, & Mrkonjic, 1988).

When participants are shown incorrect answers to simple arithmetic problems (e.g., *2 + 3 = 7*), an N400-like effect is elicited in comparison to the correct answer (Jost, Henninghausen, & Rösler, 2004; Szucs & Csépe, 2005; Wang, et al., 2000; Zhou et al., 2006). Furthermore, in multiplication, this effect is modulated by how closely-related the incorrect answer is to the correct answer (e.g., *4 x 5 = 25* elicits a smaller-magnitude N400 effect than *4 x 5 = 27*) (Niedeggen & Rösler, 1999). However, numerical distance between the correct answer and an unrelated incorrect answer (e.g. 2 + 3 = *7* vs. = *25*) does not, in itself, affect the N400 amplitude (Szucs & Csépe, 2005). These effects seem analogous to the way that different degrees of semantic incongruency influence the magnitude of the N400 effect in language. In most of these arithmetic ERP studies, though, the N400 effect occurs approximately 100 ms earlier than it does in language. It is possible that this earlier onset in the N400 effect reflects differences between the arithmetic and language memory networks. However, there is also a great deal of variation among the studies in the onset and morphology of the effect, so comparison is difficult and the reason for the differences is unclear.

Arithmetic and semantic conceptual networks vary greatly in size, constraints, and organization. In particular, arithmetic knowledge is rule-based, and for any given simple arithmetic problem, there is only one correct answer. By contrast, sentences vary in their degree of constraint, in that for many sentences there is more than one word that can be used in a particular place while still maintaining a sensible semantic meaning. Furthermore, knowledge of mathematics and language are acquired in very different ways and at different time points during early learning and development (e.g., Siegler, 1981). These differences in initial acquisition may affect the way that the two domains of knowledge are constructed and utilized.

Although it is possible that the N400 effect indexes an analogous conceptual integration process in both the arithmetic and language domains of knowledge, the N400 effect differences between the two domains, particularly in the onset of the effect, leave open the possibility that the conceptual integration process for arithmetic and language is qualitatively different. Our study was designed to help distinguish between these two possibilities.

Purpose of Present Investigation

We investigated two possible sources of variation between the ERP studies of arithmetic and of language. First, unlike most previous studies that document an arithmetic N400 effect, we used a task and a presentation methodology that is standard in language research. Second, we examined whether the presentation format, digits in arithmetic and words in language, could be contributing to the differences in the onset of the ERP waves. To this end, we presented individuals with identical arithmetic problems in two different presentation formats. One format condition presented arithmetic concepts as digits that were related by arithmetic symbols (e.g., *12 + 3 = 15*). In the other format condition, the same problems were presented in words, in the form of an English sentence (e.g., *Twelve plus three equals fifteen).* In both cases, participants were asked to verify the correctness of the answers (*15* or *fifteen*, respectively). In controlling for the conceptual meaning while varying the presentation format (digits vs. words), we were able to examine the possible influence of format on the arithmetic N400 effect.

Method

Participants

Participants were 33 volunteer graduate and undergraduate students (18 female, 15 male). They were either compensated $15/hr for their participation or given extra credit for the introductory psychology course in which they were enrolled, as applicable. The average age of the participants was 23.55 years (*SD* = 4.58). All participants were either native English speakers or had been speaking English since the age of six. Participants were randomly assigned to either the Digit format condition (N = 16) or the Word format condition (N = 17).

Stimuli & Design Stimuli were two-operand arithmetic problems with either correct or incorrect answers. The problems were constructed in a way that satisfied a number of constraints established by cognitive arithmetic literature, and constraints that were required by a separate experiment not described here[2]. First, the two operands could be both added and divided (e.g., *12 + 3*; *12 / 3*). Second, we excluded tie problems (e.g., *2 + 2*) and problems containing a one, zero, or 10 as an operand, as evidence from prior work suggests that these types of problems are processed differently (e.g., using rule-based procedures), and often more easily, than other simple arithmetic problems (Ashcraft, 1992; McCloskey, 1992). Third, we only selected problems that fell into the "small" category of division problems, defined as having a divisor lesser than 25, in order to avoid some of the issues of the problem-size effect[3] (see Zbrodoff & Logan, 2004 for a review). Finally, we controlled for answer parity (LeMaire & Reder, 1999). Within these constraints, we created a set of 48 problems (24 addition, 24 division).

Pilot testing was done in order to ensure that the addition and the division problems were of equivalent difficulty. To create an answer verification task for our stimuli, we constructed two different incorrect answers for each problem. The "close" incorrect answer was derived by adding or subtracting the value one or two to or from the correct answer (e.g., *12 + 3 = 14*). The "other" incorrect answer was derived by performing a different arithmetic operation on the operands. For example, the "other" incorrect answer for *12 + 3* was *12 + 3 = 4*, with *4* being a correct answer to the division operation. Based on the results of this pilot study, we selected a set of 12 addition and 12 division problems, composed of identical operand pairs, that were the most closely equivalent in difficulty across the two operations.

We then created our two experimental format conditions. In the Digit-format condition, the problems were presented as a traditional arithmetic combination of digits and symbols (e.g., *12 + 3 = 15*). In the Word-format condition, the digits and symbols were converted to words. The first word was capitalized in order to resemble a sentence format similar to one that might be presented in a traditional language ERP study (e.g., *12 + 3 = 15* was changed to *Twelve plus three equals fifteen)*.

Using these stimuli, we constructed a 2 (Format: Digits, Words) x 2 (Operation: Addition, Division) x 2 (Answer: Correct, Incorrect), mixed factorial experiment. Format was manipulated between participants, and Operation and Answer type were manipulated within participants. Also, the Operation factor was not of theoretical interest to the question we address here, so for the sake of simplicity each operation was analyzed separately (see results section). The experiment consisted of three blocks of trials. During each block, each of the twenty-four problems was presented twice with a correct answer, once with a "close" incorrect answer, and once with an "other" incorrect answer. Though we distinguished between the "close" and "other" answer types when collecting data, responses to these two incorrect answer categories were collapsed in the analyses. There were 96 total trials in each block, with 50% correct equations in each block; trial order was randomized individually within each of the three blocks. There were a total of 288 trials during the experiment.

Procedure After giving written informed consent, the participant was fitted with an electrode cap. The participant did not complete any practice trials, but were given a trial demonstration and a thorough explanation of the task from the experimenter. Participants initiated the start of each block by pressing a key and were given a break, of the duration of their choosing, after each 96-trial block. Within each block, equations appeared in a randomized order. Each trial consisted of a fixation point (an asterisk, '*') of duration 500ms, followed by the presentation of a complete arithmetic problem in a delayed-verification paradigm. Each item of the problem was presented alone on a screen for 350ms followed by a 250ms inter-stimulus interval, for a total trial duration of 3.75sec. A response screen followed the problem presentation. Participants responded using a hand-held controller and were asked to strive for 100% accuracy. The *YES* response hand was counterbalanced across participants, and their response triggered the onset of the next trial. They were asked not to blink between the onset of the fixation point and the appearance of the response screen. They were permitted to blink and take a short break while the response screen was displayed, as response time was not recorded. ERPs of interest were those evoked during the presentation of the answer (e.g., *15* in *12 + 3 = 15*). Most participants finished the experimental task within 45 minutes, making the entire experiment time, including set-up, less than two hours.

[2] This particular experiment required that: 1) Addition and division were the only operations used, and 2) One of the "incorrect" answer conditions was derived from the correct answer of the alternative operation. For example, for the equation *12 + 3 = 15*, one of the incorrect answers would be *14*, and one would be *4*.

[3] Note, however, that "small" division problems translate into "large" addition problems. As described in this section, the stimuli selection pilot study was conducted primarily to ensure that the addition and division problems selected were of equivalent difficulty.

Data Acquisition & Results

EEG recording

Continuous EEG was recorded from 19 tin electrodes attached to an elastic cap (Eletro-cap Inernational) in accordance with the extended 10-20 system. Vertical eye movements and blinks were monitored by two electrodes, one placed beneath the left eye and one placed to the right of the right eye. The 19 electrodes were referenced to an electrode placed over the left mastoid and were amplified with a bandpass of 0.01-100Hz (3db cutoff) by an SAI bioamplifier system; an additional low-pass filter at 15Hz was applied prior to final analysis. Impedances at scalp and mastoid electrodes were held below 5 μΩ. Trials associated with blinking, excessive eye movement or amplifier blocking were removed prior to averaging (approximately 9% of all trials). Stimuli were displayed to participants on an 18" CRT monitor approximately three feet from the participants at eye-level (participants were seated during the experiment) with white font on a black background.

N400 Component All analyses were performed on mean voltage amplitudes in the following time windows: 200-400ms, 250-450ms, and 300-500ms following the presentation of the answer to each arithmetic problem, which was the critical item of each trial. These three time intervals were chosen based on the typical onset observed for the N400 effect in other work. Effects were not significant in the 300-500ms window, and were strongest in the 200-400ms time window. Thus, these are the only results reported here. It is important to note that this time window represents a peak of the N400 effect in a time window that is approximately 100ms earlier than is typically observed in language studies, a finding that is consistent with prior studies on arithmetic.

Analyses were performed separately for Addition and Division. We conducted three factorial ANOVAs[4] in the 200-400ms time window for different locations. For the answer type variable (correct vs. incorrect), planned comparisons revealed no significant differences between the two types of incorrect answers ("close", and "other"). Thus, these two answer types were collapsed into an overall "Incorrect" condition. The midline analysis was a 2 (Format: Numbers, Words) x 2 (Answer Type: Correct, Incorrect) x 3 (Electrode Location: Fz, Cz, Pz) mixed ANOVA; the lateral analysis was a 2 (Format: Numbers, Words) x 2 (Answer Type: Correct, Incorrect) x 2 (Hemisphere: Left, Right) x 5 (Electrode Location: Fp, F, C, P, O, respective to each hemisphere) mixed ANOVA; the peripheral analysis was a 2 (Format: Numbers, Words) x 3 (Answer Type: Correct, Incorrect) x 2 (Hemisphere: Left, Right) x 5 (Electrode Location: F, T, P, respective to each hemisphere) mixed ANOVA.

There were no main effects of the between-participants variable of format (Digits vs. Words) for either operation at any electrode location group. This finding indicates that the conceptual processing of correct and incorrect arithmetic problems does not depend on the presentation format. A robust N400 effect was found for incorrect answers in both addition and division operations (see Figure 1), and it was comparable for both format conditions. The magnitude and equivalence of the effects is best represented by the difference waveforms in Figure 1. All ANOVA values for the main effect of answer type are listed in Table 1.

To summarize, incorrect answers to arithmetic problems elicited a large-amplitude, negative-going waveform relative to correct answers. This N400 effect did not depend on the presentation format (Digits vs. Words), and was similar to the N400 effect found in language for semantic violations. Nonetheless, as in previous studies on the arithmetic N400, this effect peaked about 100ms earlier than it does in language.

Table 1: ANOVA values for the main effect of Answer

Location	Operation	*F* & *MSE*
Midline	Addition	$F(1, 31) = 58.12$*** $MSE = 2.94$
	Division	$F(1, 31) = 24.92$*** $MSE = 6.13$
Lateral	Addition	$F(1, 31) = 59.99$*** $MSE = 5.74$
	Division	$F(1, 31) = 37.82$*** $MSE = 8.67$
Peripheral	Addition	$F(1, 31) = 54.49$*** $MSE = 1.66$
	Division	$F(1, 31) = 34.35$*** $MSE = 2.70$

* $p < .05$, ** $p < .01$, *** $p < .001$

Topography The effect mostly displayed a centro-parietal topographical distribution, with a slight bias to the right hemisphere. Although there were no main effects of format on the N400 effect, there were significant topographical differences between the two formats in the distribution of the effect, as can be seen in Figure 2.

For addition, the N400 effect was slightly stronger over the midline sites in the Words vs. the Digits format, as evidenced by an interaction effect at those sites between answer type and format, $F(1, 31) = 8.03$, $MSE = 2.85$, $p < .01$ and a three-way interaction between answer type, midline electrode position, and format, $F(2, 62) = 5.32$, $MSE = .56$, $p < .05$. There were also interactions between format and electrode sites in the lateral and peripheral regions, indicating that the electrode locations where the N400 effect was strongest differed between Words and Digits. In particular, there was a significant interaction between answer type, lateral electrode position, and condition, $F(4, 124) = 7.41$, $MSE = 2.12$, $p < .01$; a significant four-way interaction between answer type, hemisphere, lateral electrode position, and format condition, $F(4, 124) = 2.90$, $MSE = 1.90$, $p < .05$; and a three-way

[4] A Greenhouse- Geisser correction for violation of sphericity was used whenever necessary.

interaction between answer type, peripheral electrode position, and format $F(2, 62) = 5.88$, $MSE = .91$, $p < .01$.

For division, similar differences in topography were observed, though there is somewhat less variability between the Digits and Words conditions as compared to addition. For division, the N400 effect was more strongly distributed on the right hemisphere at both lateral $F(1, 31) = 2.98$, $MSE = 2.08$, $p = .09$, and peripheral sites, $F(1, 31) = 5.03$, $MSE = 3.74$, $p < .05$, though the effect at the lateral sites was only marginally significant. It is possible that these differences in topography reflect differences in memory organization and access. We address this possibility in the discussion.

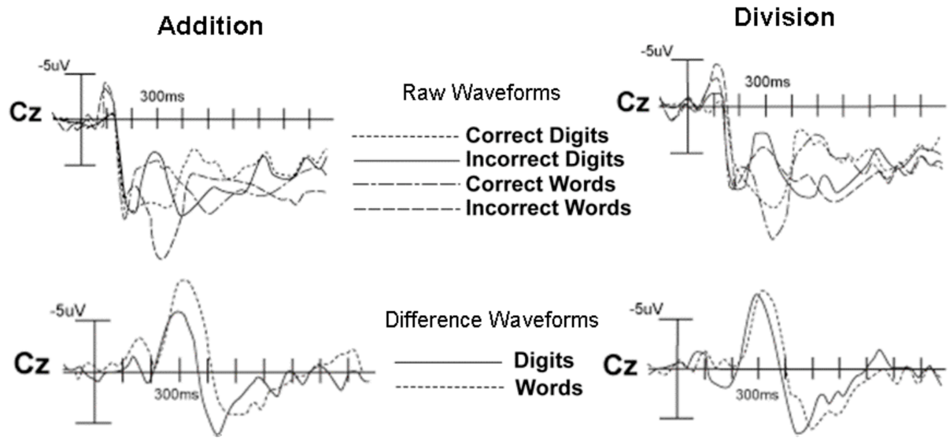

Figure 1. Raw and difference waveforms of the Digit and Word Condition on a representative midline electrode

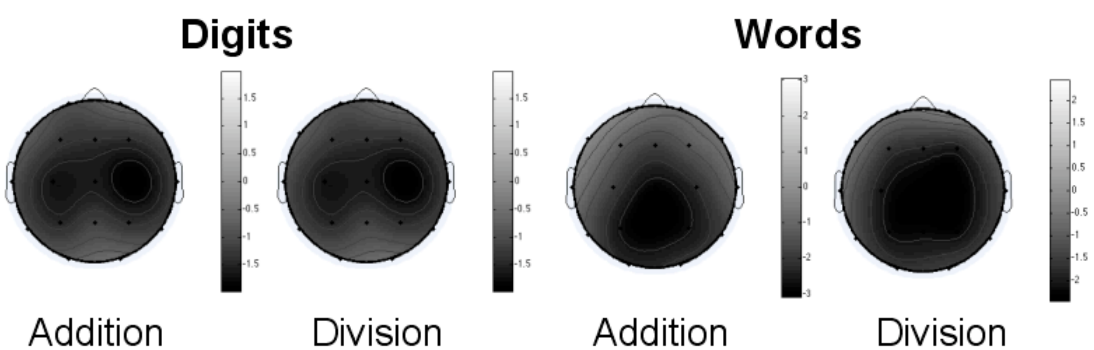

Figure 2. Topographical distribution of the N400 effect (200-400ms time window) in each condition

Discussion

In this study, we wanted to investigate possible reasons for various differences between the N400 effect in arithmetic and in language observed in previous work. To this end, we compared the ERP responses elicited by the conceptual integration of identical mathematical concepts presented either in Digit format (e.g., *12 + 3 = 15*) or in Word format (e.g., *Twelve plus three equals fifteen*). For both, we used a methodology that is standard in language research. In doing so, we examined two possible sources of difference between previous N400 studies in arithmetic and in language.

We found that incorrect answers to addition and division problems, relative to correct problems, elicited a robust, negative-going waveform, similar to that elicited by semantic violations in language. The same effect was elicited regardless of whether the arithmetic problems were presented in Digit or in Word format. These results provide strong support to the conjecture made by other researchers (e.g., Niedeggen & Rösler, 1999; McPherson & Holcomb, 1999) that the N400 effect is not specific to language. Rather, despite small variations, it is probably an index of conceptual integration in a variety of semantic knowledge domains, including arithmetic.

Interestingly, the N400 effect in our study peaked approximately 100ms earlier than is typically observed in traditional language studies. We suspect that this is possibly due to differences between the arithmetic and language memory networks. Among other differences, the arithmetic memory network is much smaller and more tightly constrained, in terms of connections among items, compared to the language network. The fact that we conducted our study with such a small subset of arithmetic problems may have added further constraints. Thus, conceptual integration within this smaller, more constrained "conceptual space" may occur more quickly. Of course, more work needs to be done to determine if this is the case, possibly by trying to find conditions under which the onset of the language N400 can be reliably shifted 100 ms earlier.

Even though there were no significant differences of arithmetic problem format on the onset or overall amplitude of the N400 effect, there were topographical differences in the distribution of the N400 effect between digits and words. These differences were most pronounced in addition. Previous work in cognitive arithmetic has suggested that the format in which numerical concepts are presented (digits vs. words) can influence memory encoding and access (e.g., Noel, Fias, & Brysbaert, 1997). Therefore, it is possible that the variance observed in the topographical distributions reflects such differences. The fact that addition problems, which have a stronger representation in verbal memory (e.g., Ashcraft, 1992), showed more variation between the two formats would be consistent with this idea.

Importantly, our results demonstrate that regardless of the avenue in memory used to access the concepts in the arithmetic problem, the process of integrating those concepts together to be able to verify an answer is equivalent for both digits and words.

Acknowledgments

This research was partially funded by the University of Washington's Royalty Research Fund, with a grant awarded to Miriam Bassok. Thanks to the members of the Cognitive Neuroscience of Language Lab for help with data collection and theoretical insights.

References

Ashcraft, M.H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*, 75-106.

Campbell, J. I. D. (1990). Retrieval inhibition and interference in cognitive arithmetic. *Canadian Journal of Psychology, 44*, 445-464.

Campbell, J. I. D., & Penner-Wilger, M. (2006). Calculation latency: The mu of memory and the tau of transformation. *Memory & Cognition, 34*, 217-226.

Holcomb, P. J. (1993). Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. *Psychophysiology, 30,* 47-61.

Jost, K., Hennighausen, E., & Rösler, F. (2004). Comparing arithmetic and semantic fact retrieval: Effects of problem size and sentence constraint on event-related brain potentials. *Psychophysiology, 41*, 46-59.

Kutas, M. & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4*, 463-470.

Kutas, M., & Hillyard, S. A. (1980). Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science, 207*(4427), 203-205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

LeFevre, J., Bisanz, J., & Mrkonjic, L. (1988). Cognitive arithmetic: Evidence for obligatory activation of arithmetic facts. *Memory & Cognition, 16,* 45-53.

LeMaire, P. & Reder, L. (1999). What affects strategy selection in arithmetic? The example of parity and five effects on product verification. *Memory & Cognition, 27*, 364-382.

McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, *44*, 107-157.

McPherson, W. B., & Holcomb, P. J. (1999). An electrophysiological investigation of semantic priming with pictures of real objects. *Psychophysiology, 36*, 53-65.

Niedeggen, M., & Rösler, F. (1999). N400 effects reflect activation spread during retrieval of arithmetic facts. *Psychological Science, 10*, 271-276.

Noel, M.-P., Fias, W. & Brysbaert, M. (1997). About the influence of the presentation format on arithmetical-fact retrieval processes. *Cognition, 63*, 335-374.

Osterhout, L. McKinnon, R., Bersick, M., & Corey, V. (1996). On the languge specificity of the brain response to syntactic anomalies: Is the syntactic positive shift a member of the P300 family? *Journal of Cognitive Neuroscience, 8*, 507-526.

Osterhout, L. & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.

Siegler, R. (1981). Developmental sequences within and between concepts. *Monographs of the Society of Research in Child Development, 46*, 1-84.

Sitnikova, T., Holcomb, P., Kiyonaga, K., & Kuperberg, G. (2008). Two neurocognitive mechanisms of semantic integration during the comprehension of visual real-world events. *Journal of Cognitive Neuroscience, 20*, 2037-2057.

Szucs, D., & Csépe, V. (2005). The effect of numerical distance and stimulus probability on ERP components elieicted by numerical incongruencies in mental addition. *Cognitive Brain Research, 19,* 10-27.

Wang, Y., Kong, J., Tang, D., Zhuang, D., & Li, S. (2000). Event-related potential N270 is elicited by mental conflict processing in human brain. *Neuroscience Letters, 293*, 17-20.

Zbrodoff, N. J., & Logan, G. D.(2005).What everyone finds: The problem size effect. In J. I. D. Campbell (Ed.), Handbook of mathematical cognition (pp. 331-346). New York: Psychology Press.

Zhou, X., Chen, C., Dong, Q., Zhang, H., Zhou, R., Zhao, H., et al. (2006). Event-related potentials of single-digit addition, subtraction, and multiplication. *Neuropsychologica, 44*, 2500-2507.