When Two plus Two Does Not Equal Four: Event-Related Potential Responses to Semantically Incongruous Arithmetic Word Problems

Kristie J. Fisher (kjfisher@uw.edu)
University of Washington, Department of Psychology, Box 351525
Seattle, WA 98105 USA

Miriam Bassok (mbassok@uw.edu)
University of Washington, Department of Psychology, Box 351525
Seattle, WA 98105 USA

Lee Osterhout (losterho@uw.edu)
University of Washington, Department of Psychology, Box 351525
Seattle, WA 98105 USA

Abstract

Extensive research measuring event-related brain potentials (ERPs) shows that semantic incongruity is indexed by the N400 effect and syntactic/structural incongruity is indexed by the P600 effect. We used these indices to examine how people coordinate their semantic and arithmetic knowledge when they read simple addition and division word problem sentences (e.g., "Twelve roses plus three daisies equals fifteen"). Prior work in problem solving has shown that word-problem solutions are modulated by analogical alignment of semantic and arithmetic relations, such that people avoid or commit errors on misaligned problems (e.g., Aligned: "Twelve roses plus three daisies equals fifteen"; Misaligned: "Twelve cookies plus three jars equals fifteen"). Here, we found that such analogical alignments modulate the comprehension of word-problem sentences. Specifically, we found that analogically Misaligned semantic relations elicited a P600 effect. Furthermore, an N400 effect was elicited by the last number word of Misaligned problem sentences, even when it was a mathematically correct answer. These results show that analogical alignment between semantic and arithmetic relations can be indexed with the P600 effect and provide a foundation for future ERP work on analogical reasoning.

Keywords: ERP; analogy; mathematical cognition; N400 effect; P600 effect

Introduction

A common task facing the cognitive system is *conceptual integration* of individual items into a meaningful whole. For example, language comprehension requires conceptual integration of consecutive words into meaningful sentences. Similarly, comprehension of arithmetic problems requires conceptual integration of numbers and arithmetic operators into correct mathematical expressions. In this paper, we explore the conceptual integration of simple arithmetic word problems, which are unique in that they require conceptual integration of language and of mathematics.

Conceptual Integration & ERPs

The process of conceptual integration, and the conditions under which it can be disrupted, have been investigated in a variety of domains using event-related potential (ERP) methodology, which measures the electrical brain activity elicited by a particular stimulus. Work in this area has shown that two key aspects of conceptual integration, meaning and structure, are indexed by two distinct and highly reliable ERP components—the N400 and P600 components, respectively.

The N400 component is negative-going and peaks around 400ms after presentation of the stimulus. This component is highly sensitive to contextual semantic meaning. The magnitude of this component is larger for semantically incongruous compared to congruous items—a difference known as the *N400 effect*. The N400 effect was first documented in sentence processing. For example, the italicized word in the sentence, "The cat will *bake* the food" will elicit an N400 effect relative to, "The cat will *eat* the food" (Kutas & Hillyard, 1980). Subsequent work has shown that the N400 effect is elicited in response to conceptual incongruities in other domains. For example, incorrect answers to simple symbolic (e.g., "4 x 4 = *21*") and verbal (e.g., "Twelve plus three equals *sixteen.*") arithmetic problems elicit an N400 effect (e.g., Niedeggen & Rosler, 1999; Fisher, Bassok, & Osterhout, 2009). Thus, the N400 effect is generally accepted to be a domain-general index of semantic congruence.

The P600 component is positive and peaks at around 600ms after stimulus presentation. A P600 effect is elicited by violations of syntax within a sentence (e.g., "The cat will *eating* the food I leave on the porch."; Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995) and by violations of structure, such as a wrong note played in a harmonic scale (Patel et al., 1998). Such violations of syntax or structure lead to larger P600 amplitudes, relative to control conditions (i.e., the P600 effect).

Furthermore, Osterhout and Mobley (1995) found that when there is any kind of violation within a sentence, syntactical or semantic, an N400 effect is also elicited by the ending word of the sentence, even though that word is perfectly correct. This last-item N400 effect is likely the result of the experimental paradigm typically used in language research. Participants are typically asked to make binary judgments about the "acceptability" of the sentences they just saw (usually they are not instructed to look for any particular type of error). Thus, when participants reach the end of a sentence that contained a violation, the entire sentence must now be categorized as "unacceptable." The N400 effect to the final word in the sentence may be a result of this judgment processing.

Despite the relatively broad range of studies of conceptual integration, to our knowledge, no previous studies have used these ERP indices to examine a) how people integrate concepts that are presumably organized in distinct conceptual networks and b) the integration of concepts via analogy. Both of these characterize the integration process involved in the solution of mathematical word problems, whereby people are required to apply arithmetic operations in a way that fits the relations among objects in the "real world." This process involves analogical coordination of real world knowledge (e.g., *roses* and *daisies* are *flowers*) with one's knowledge of arithmetic properties (One can add 3 *roses* to 5 *daisies* to create a bouquet of 8 *flowers*). As we explain in the next section, people are highly systematic in the way they coordinate their semantic and arithmetic knowledge. The purpose of our study was to use ERP to examine such cross-network, analogical conceptual integration and, in particular, to test whether the same ERP components that index violations of meaning and structure in language also index violations of analogical alignment.

Mathematical Problem Solving in the "Real World"

Research by Bassok and her colleagues has shown that, when people reason about mathematical word problems, they tend to align structurally analogous semantic and arithmetic relations (Bassok, Chase, & Martin, 1998). Specifically, people align categorically related objects (e.g., cars and trucks) with the commutative addition operation and align functionally related objects (e.g., jars and cookies,) with the non-commutative division operation. Violating such *semantic alignment* (e.g., having to add jars to cookies or having to divide cars by trucks) severely impairs problem-solving performance (Bassok, Wu, & Olseth, 1995; Martin & Bassok, 2005), and even blocks retrieval of arithmetic facts from memory (Bassok, Pedigo, & Oskarsson, 2008).

In the present study we investigated how people conceptually integrate semantic and arithmetic relations while reading simple addition and division word problems presented in a sentence format (e.g., "Twelve roses plus three daisies equals fifteen."). We recorded ERPs as participants read these word-problem sentences. We analyzed the electrical waveforms elicited by the second object word in the sentence, which completed the semantic relation, and by the numerical mathematical answers (e.g., the two underlined words in, "Twelve roses plus three <u>vases</u> equals <u>fifteen</u>"). The semantic object relations were either analogically aligned (Aligned condition) or misaligned (Misaligned condition) with the arithmetic relation in the word problem, and the mathematical answers were either correct or incorrect (see Table 1 for example stimuli). After reading each word-problem sentence, participants were asked to make judgments as to whether or not the problem was "acceptable." As is standard practice in typical language research paradigms, we did not specify the criteria by which participants were to make their judgments.

Table 1: Example Stimuli

Object Alignment	Math Correct	Math Incorrect
Aligned Addition	Twelve limes plus three lemons equals fifteen.	Sixteen cars plus two trucks equals twenty.
Aligned Division	Fifteen roses divided by three bouquets equals five.	Six robins divided by two nests equals eight.
Misaligned Addition	Six questions plus three quizzes equals nine.	Eight cookies plus four jars equals two.
Misaligned Division	Eighteen skirts divided by two dresses equals nine.	Fifteen geese divided by three ducks equals six.

We had two main predictions. First, we expected that conceptual integration in Aligned word problems should be similar to conceptual integration in arithmetic problems, presented in sentence-form, which do not contain objects. Specifically, we expected that, in the Aligned condition, we would replicate the N400 effect elicited by mathematically incorrect answers to arithmetic problems (Fisher, Bassok, & Osterhout, 2009). Second, and most important, if conceptual integration via analogy is similar to conceptual integration in rule-governed sequence processing, then analogical misalignment of the semantic and arithmetic relations in the problem should elicit a P600 effect. In particular, we expected a P600 effect to occur at the second object word because that word completes a semantic relation that cannot be mapped onto the arithmetic relation in the problem, and thus constitutes a structural violation (e.g., Gentner, 1983). Furthermore, we expected that the mathematically correct answer (the final item of the word problem sentence) in Misaligned problems would elicit an N400 effect relative to correct answers in Aligned problems, replicating previous work by Osterhout and Mobley (1995).

Methods

Participants

The participants were 38 volunteer undergraduate students, graduate students, and staff from the University of

Washington (21 male, 17 female; M_{age} = 22.23 years, SD_{age} = 4.98 years) who were right-handed native English speakers. Participants were either given course extra credit or paid $30 for their participation.

Stimuli

The stimuli were simple word problem sentences that were composed of digit pairs and object word pairs that were either categorically related or functionally related. The digit and object pairs were selected based on pilot testing, as described below.

Arithmetic Problems The arithmetic problems were composed of two operands and satisfied a number of constraints established by cognitive arithmetic literature and required for our experimental manipulations. First, the two operands could be both added and divided to yield a whole-number answer (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, 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[1] (see Zbrodoff & Logan, 2004, for a review). Finally, we controlled for answer parity (LeMaire & Reder, 1999).

Within these constraints, we created a set of 24 problems, 12 addition and 12 division, that were equivalent in difficulty. These problems were selected based on results of a pilot study (error rate and response time), in which 154 undergraduate students solved 48 addition and 48 division problems meeting the above criteria. To create an answer verification task, we constructed two different incorrect answers for each problem. The "Close" incorrect answer for both operations 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 answers for addition were the correct answers to division problems with the same operands (e.g., *12 + 3 = 4*), and the "Other" incorrect answers for division were the correct answers to addition problems with the same operands (e.g., *12 / 3 = 15*)

Object Pairs We initially constructed a set of 163 word pairs that we considered to belong to one of the two semantic relations categories—categorical or functional. The set contained 83 possible categorical pairs and 80 possible functional pairs consisting of concrete, plural nouns (e.g., "cats, dogs"). From this set, we constructed rating surveys that were completed by 202 undergraduate students at the University of Washington as part of a class activity. Instructions asked students to rate, on a seven-point scale, either the extent to which the word pairs were *categorically* related or the extent to which they were *functionally* related. The average categorical and functional ratings in these two conditions were compared for each word pair using an independent t-test with an alpha level of .05. In order to be included in the final set, word pairs had to have significantly different categorical and functional ratings and an average rating of greater than 5 in one dimension and 4 or less in the other. Based on these ratings, we selected 48 categorical and 48 functional pairs. The word pairs in both relation conditions were equivalent in their average number of syllables and letters in each word.

Design

Operation (Addition vs. Division) was manipulated between participants ($N_{Addition}$ = 19; $N_{Division}$ = 19; participants were randomly assigned). Analogical alignment of the mathematical operation and the object sets (Aligned vs. Misaligned), and mathematical correctness of the problems (Correct vs. Close Incorrect vs. Other Incorrect) were manipulated within participants.

Verbal versions of the arithmetic problems (e.g., Twelve plus three" in place of "12 + 3") were created and were then combined with object pairs to create simple word problem sentences (e.g., "Twelve limes plus three lemons equals fifteen."). For the Addition problems, all of the Aligned stimuli were categorically related objects, and all of the Misaligned stimuli were functionally related objects; the reverse was true for the Division problems (see again Table 1). Thus, the same object sets were used for both operations, but for one operation the object sets were Aligned and for another they were Misaligned.

The experiment consisted of three blocks of trials. There were 96 trials in each block, for a total of 288 trials. Within each block, 50% of the trials were Aligned word problems, and 50% were Misaligned. Within each alignment type, 50% were mathematically Correct, 25% were Close Incorrect, and 25% were Other Incorrect. Trial order was pseudo-randomized within each of the three blocks. Each of the word pairs appeared once per block, and they were combined with different arithmetic problems each time.

Procedure

Participants were seated comfortably in front of a CRT monitor in an isolated room and fitted with electro-encephalography (EEG) recording equipment. Each trial consisted of a fixation point (500ms), and each item of the word-problem sentence was presented alone on a screen (450ms/350ms ISI). The final inter-stimulus interval before the appearance of the YES/NO response screen was 1,000 ms (total trial duration was 7.1 seconds). Participants were given a hand-held controller and were asked to respond *YES* (response hand counter-balanced) using one button if they thought the problem was completely "acceptable" and *NO*, using another button if the problem was "unacceptable" in

[1] 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 problems selected were of equivalent difficulty.

any way. They were told that the instructions were intentionally vague because the criteria by which they would judge the problems were at their discretion. Furthermore, the task did not include object labels, which are usually required in word problem solving. Participants 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. Response time was not recorded and responses triggered onset of the next trial. A break was given after each block. The entire experiment time, including set-up, was less than two hours.

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. Electrical signals were amplified, digitized at a rate of 250Hz, and bandpass filtered at 0.01-40Hz. Impedances at scalp and mastoid electrodes were held below 5 kΩ. Trials associated with blinking, excessive eye movement or amplifier blocking were removed prior to averaging (approximately 11% of all trials). Stimuli were displayed to participants on an 18" CRT monitor approximately three feet from the participants at eye-level with white font on a black background.

Behavioral Responses

Because participants were asked to make open-ended "acceptability" judgments, it is not surprising that there was variation in how they judged the Misaligned problems, particularly in the case where the problem was Misaligned but mathematically correct. These behavioral differences in acceptability judgments corresponded to differences in the magnitude of the overall ERP effects we report here. In this paper we do not discuss these individual differences, as they are not essential to our primary research question.

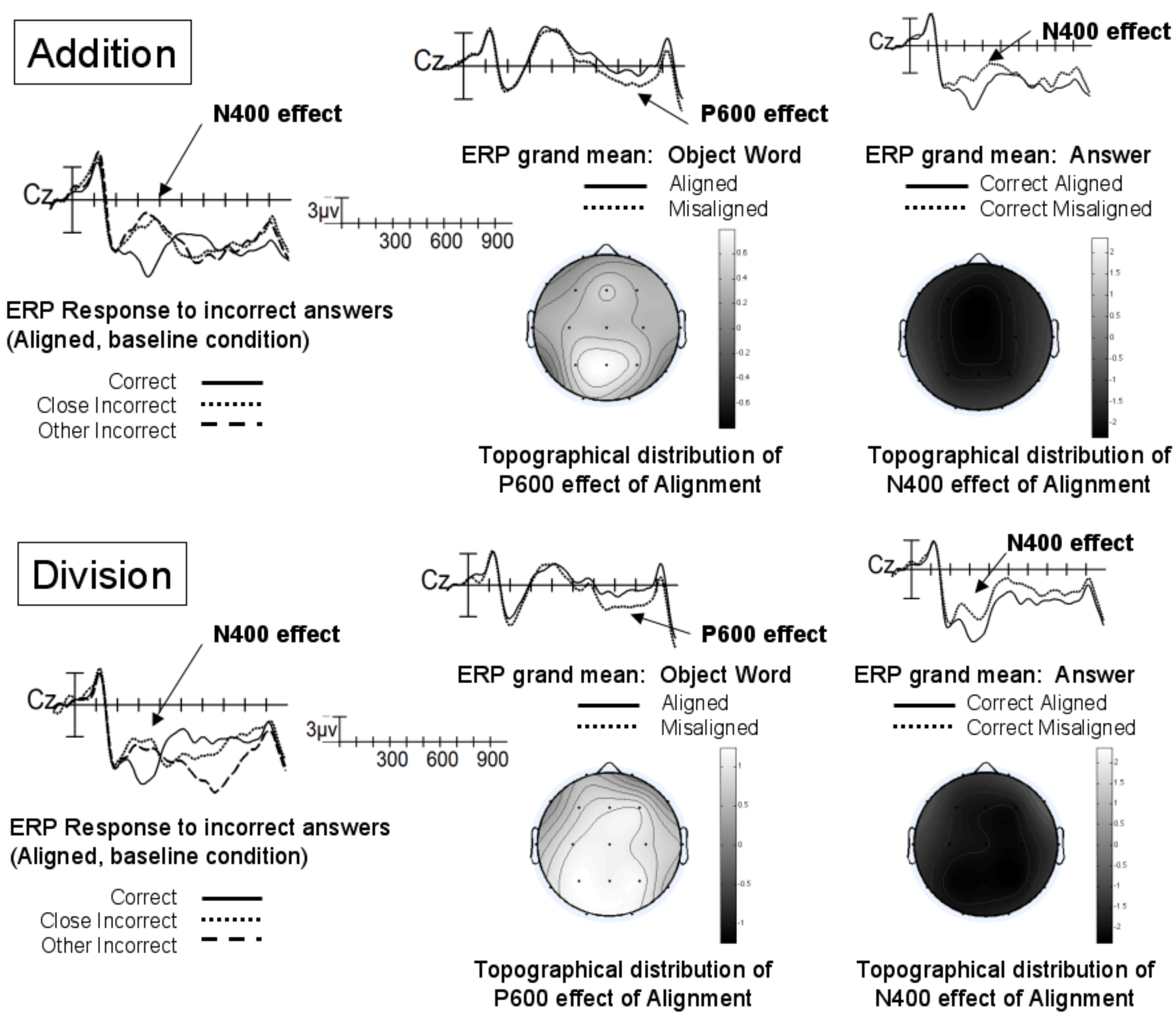

Figure 1: ERP responses to mathematically incorrect answers and semantic alignment.

ERP Responses

EEG amplitudes elicited by the second object word and by the last word (the mathematical answer to the problem) were averaged according to their respective Alignment and Answer conditions. Mean amplitudes were compared separately for the Addition and Division operations in the 250-450ms time window (N400 effect) and in the 500-700ms time window (P600 effect) following stimulus onset. For all analyses, separate ANOVAs[2] were conducted for midline (Fz, Cz, Pz), medial (Fp1, F3, C3, P3, O1, Fp2, F4, C4, P4, O2), and lateral (F7, T7, P7, F8, T8, P8) electrode sites, with electrode location and hemisphere included as factors in each ANOVA.

Incorrect Answers We first tested whether we replicated previous work with sentence-form arithmetic problems (Fisher, Bassok, & Osterhout, 2009), which found an N400 effect to incorrect numerical answers. We tested for this effect within the Aligned condition, which served as our baseline. Separately for Addition and Division, a 3-way (Answer Type - Correct, Close Incorrect, Other Incorrect) repeated measures ANOVA was conducted at each electrode grouping. Indeed, we found a main effect of answer types, such that mathematically incorrect answers elicited a significant N400 effect relative to correct answers for both Addition and Division word problems (see Figure 1) [Addition: $F_{Midline}(2,36) = 15.36$, $MSE = 11.99$, $p < .001$; $F_{Medial}(2,36) = 25.54$, $MSE = 17.71$, $p < .001$; $F_{Lateral}(2,36) = 22.29$, $MSE = 5.86$, $p < .001$; Division: = $F_{Midline}(2,36) = 6.08$, $MSE = 8.10$, $p < .01$; $F_{Medial}(2,36) = 10.37$, $MSE = 13.69$, $p < .001$; $F_{Lateral}(2,36) = 7.24$, $MSE = 6.23$, $p < .01$] Planned contrasts revealed no significant differences between the Close and Other incorrect answer types except at the lateral electrode sites for Addition[3].

Semantic Alignment When comparing ERP responses *between* semantic alignment conditions, we first compared the ERP waveforms elicited by the second object word in Aligned versus Misaligned problems (e.g., Twelve cars plus three *trucks* equals fifteen. vs. Twelve roses plus three *vases* equals fifteen.). Consistent with our predictions, we found that the second object word in the Misaligned condition elicited a P600 effect relative to the Aligned condition in both operations [Addition: $F_{Midline}(1, 18) = 5.28$, $MSE = 4.08$, $p = .03$; $F_{Medial}(1, 18) = 6.39$, $MSE = 6.66$, $p = .02$; $F_{Lateral}(1, 18) = 5.78$, $MSE = 1.68$, $p = .03$; Division: $F_{Midline}(1, 18) = 4.94$, $MSE = 4.25$, $p = .04$; $F_{Medial}(1,18) = 5.62$, $MSE = 7.16$, $p = .03$; $F_{Lateral}(1,18) = 3.20$, $MSE = 1.95$, $p = .09$]. As noted earlier, this effect occurred regardless of participants' behavioral response as to whether or not the problem was "acceptable."

Next, we compared ERP amplitudes elicited by the mathematically correct answer (the final item of the word problem sentence) between the Aligned and Misaligned conditions. Correct answers of Misaligned word problems elicited an N400 effect relative to the correct answers of Aligned word problems [Addition: $F_{Midline}(1,18) = 9.00$, $MSE = 13.69$, $p < .01$; $F_{Medial}(1,18) = 11.16$, $MSE = 23.81$, $p < .01$; $F_{Lateral}(1,18) = 8.62$, $MSE = 6.67$, $p < .01$; Division: $F_{Midline}(1,18) = 8.01$, $MSE = 15.67$, $p =.01$; $F_{Medial}(1,18) = 9.34$, $MSE = 32.98$, $p < .01$; $F_{Lateral}(1,18) = 7.26$, $MSE = 8.01$, $p = .02$].

This pattern of ERP results mirrors those found in studies of language processing (e.g., Osterhout & Mobley, 1995). That is, structural/syntactic violations within a sentence typically elicit a P600 effect, and the final word of sentences containing such violations elicits an N400 effect, even when those words contained no violations. In the case of our particular stimuli, the sentences were simple arithmetic word problems, and the structural violations were violations of analogical alignment between the semantic and arithmetic relations in the problem. The final words in the sentences were the mathematical answers to the word problems, and an N400 effect occurred for mathematically correct answers in the Misaligned, relative to the Aligned, condition.

Discussion

The goal of this study was to examine the conceptual integration process with respect to arithmetic word problems and how it compares to conceptual integration for sentences and other meaningful sequences. Arithmetic word problems are unique in that they combine elements of language and math and provide the opportunity for analogical alignment or misalignment between the semantic relations and the arithmetic relations in the problem (e.g., Bassok, Pedigo, & Oskarsson, 2008; Bassok, Wu, & Olseth, 1995).

Overall, our results provide evidence for the fluid integration of arithmetic and semantic knowledge during word problem processing. More broadly, our results suggest that the conceptual integration process does not change significantly when people must integrate concepts *across* two distinct knowledge networks. That is, the same ERP effects were elicited by violations of structure and meaning in word problems as are usually found in sentences and in arithmetic problems not containing objects. These results suggest that the N400 and P600 effects could be used as dependent measures in investigations of other situations wherein an individual has to integrate distinct types of knowledge, such as in reasoning problems that involve the applications of formal logic rules to object sets, or in song writing wherein one has to coordinate lyrics with a melody.

Moreover, in the word problems used in our study, semantic and arithmetic knowledge had to be coordinated via analogy. Thus, our results also demonstrate that the

[2] A Greenhouse- Geisser correction for sphericity violations was used when necessary

[3] Specific results for different answer conditions and the interactions between the semantic alignment variable, the answer type variable, and behavioral response pattern are not central to the research question addressed here and thus are not elaborated upon for the sake of brevity.

P600 effect can serve as an index of the integrity of analogical structure within arithmetic word problems just as it indexes syntactic integrity in sentences (e.g., Osterhout & Holcomb, 1992; Osterhout & Mobley, 1995) and structural integrity in other meaningful sequences (e.g., Patel et al., 1998). As such, our results provide a foundation for future ERP investigations of the cognitive processes related to analogical reasoning, using the P600 effect as an index of structure-mapping (Gentner, 1983). That is, when the structures of two relations cannot be mapped in an analogy task (e.g., Bird:Nest as Bear:Cave vs. Bear:Desert; Spellman, Holyoak, & Morrison, 2001), the size of the P600 effect could be used to discriminate the degree of relational structure mismatch. Such investigations could examine the analogical conceptual integration within one domain (e.g., animals and their habitats) or across two domains of conceptual knowledge (e.g., Bird:Nest as Car:Garage).

Interestingly, we found these ERP effects for violations of analogical alignment and mathematical correctness across all participants even though we observed distinctly different patterns of "acceptability" judgments and corresponding ERP effect magnitude within our sample. Though we are unable to elaborate on these differences here, initial analyses suggest that these patterns are consistent with prior work in mathematical reasoning suggesting that some people are better than others at coordinating their mathematical and "real world" knowledge when constructing and solving more complex mathematical expressions than the ones presented in this study (e.g., algebraic equations; Fisher & Bassok, 2009). Because this ability is arguably relevant to our simpler task, it is not surprising that there are individual differences among our sample of participants such that some were more sensitive than others to violations of semantic alignment in simple word problems, particularly because our sentences did not include labels as part of the solution.

Of course, further work is required to fully explore these individual difference patterns and elucidate the reason behind them. To expand on our current findings, we also plan to more thoroughly investigate the processes of analogical conceptual integration. Lastly, in the future we hope other researchers will continue to use ERP for investigations of conceptual integration in more complex, knowledge-diverse situations.

Acknowledgments

This research was partially funded by the University of Washington's Royalty Research Fund, with a grant awarded to Miriam Bassok and NIDCD Research Grant R01DC01947 awarded to Lee Osterhout. Thanks to the members of the Cognitive Neuroscience of Language Lab for help with data collection and theoretical insights and to Melody Sherry and Louis Wei for help with pilot data collection and analysis.

References

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

Bassok, M., Chase, V., & Martin, S. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology, 35*, 99-134.

Bassok, M., Pedigo, S. F., & Oskarsson, A. (2008) Priming addition facts with semantic relations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 343-352.

Bassok, M., Wu, L.-L., & Olseth, K. L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition, 23,* 354-367.

Fisher, K.J., & Bassok, M. (2009) Analogical alignments in algebraic modeling. In B. Kokinov, K. Holyoak, & D. Gentner (Eds.), *Proceedings of for the 2nd International Conference on Analogy in Sofia, Bulgaria*, pp. 137-144.

Fisher, K. J., & Bassok, M., & Osterhout, L (2009). Conceptual integration is the same for digits and for words: It's the meaning, stupid! In N. A. Taatgen, H. Van Rijn, L. B. J. Schomaker, & J. Nerbonne (Eds.), *Proceedings of the 31st Annual conference of the Cognitive Science Society* (pp. 2142-2147).

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

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

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.

Martin, S. & Bassok, M. (2005). Effects of semantic cues on mathematical modeling: Evidence from word problem solving and equation construction. *Memory & Cognition, 33,* 471-478.

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

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

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

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739-773.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*, 717-733.

Spellman, B. A., Holyoak, K. J., & Morrison, R. (2001). Analogical priming via semantic relations. *Memory & Cognition, 29*, 383-393.

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.