Published in final edited form as:

Brain Res. 2017 May 01; 1662: 46–56. doi:10.1016/j.brainres.2017.02.019.

The language of arithmetic across the hemispheres: an event-related potential investigation

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

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

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

^cProgram in Neuroscience, University of Illinois, Urbana-Champaign, USA

^dBeckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, USA

Abstract

Arithmetic expressions, like verbal sentences, incrementally lead readers to anticipate potential appropriate completions. Existing work in the language domain has helped us understand how the two hemispheres differently participate in and contribute to the cognitive process of sentence reading, but comparatively little work has been done with mathematical equation processing. In this study, we address this gap by examining the ERP response to provided answers to simple multiplication problems, which varied both in levels of correctness (given an equation context) and in visual field of presentation (joint attention in central presentation, or biased processing to the left or right hemisphere through contralateral visual field presentation). When answers were presented to any of the visual fields (hemispheres), there was an effect of correctness prior to the traditional N400 timewindow, which we interpret as a P300 in response to a detected target item (the correct answer). In addition to this response, equation answers also elicited a late positive complex (LPC) for incorrect answers. Notably, this LPC effect was most prominent in the left visual field (right hemisphere), and it was also sensitive to the confusability of the wrong answer – incorrect answers that were closely related to the correct answer elicited a smaller LPC. This suggests a special, prolonged role for the right hemisphere during answer evaluation.

Keywords

numerical cognition; cerebral hemispheres; arithmetic; P300; late positivity

Corresponding author: Danielle Dickson¹, 603 E Daniel St, Champaign, IL 61820, danielle.s.dickson@gmail.com.

Second author contact: Kara Federmeier, kfederme@illinois.edu

¹Present address: 5511 Pecan Springs Apt 13103, San Antonio, TX 78249

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

The most visible characteristic of the organizational structure of the human brain is that it is made up of two cerebral hemispheres. Although their cellular and broader neurobiological makeup is very similar, cognitive functions are often lateralized to one hemisphere or the other, and this division of labor between the hemispheres seems to be an important principle of typical neural functionality. The first report of such cognitive lateralization – for language production in the left hemisphere – was integral to the foundation of neuropsychology as a field (Broca, 1865). Interest in lateralization of cognitive specialties has since expanded outside of the language domain, with research spanning broad topics such as hemispheric differences in sensitivity to different spatial frequencies (Sergent and Hellige, 1986; Christman et al., 1991), in attentional biases in global versus local features of an object (Martin, 1979), and in bottom-up versus top-down processing (Federmeier, 2007), among others. Here, our report will focus on another domain of research that has been devoting increasing attention to the processing mechanisms and abilities of each hemisphere: numerical cognition, and arithmetic processing in particular.

An interesting analogy can be formed between sentence reading (language) and arithmetic expression reading (math), which might suggest shared underlying cognitive processes. Sentences are made up of subparts (words) that are systemically combined to convey a potentially coherent message, and, similarly, arithmetic equations are made up of subparts (numbers) with combinatorial symbols (+, %, x, –) that can be sensibly completed (e.g., $4 \times 5 = 20$) or not. At the same time, there are important differences in the neural systems that are engaged in these two processes: at a whole-brain level, whereas the processes involved in reading words and sentences are largely associated with fronto-temporal activity (Lau et al., 2008; Price, 2012), the processes involved in understanding numbers and mathematical expressions are more associated with fronto-parietal activity (Chochon et al., 1999). Although the specific subregions involved seem to differ across math and language, it remains unclear to what extent the two share similar patterns of hemispheric lateralization overall, given the complexities of the subprocesses involved and the somewhat mixed evidence in the existing literature (described below). Therefore, the aim of this study is to use a combination of techniques novel to this field to examine the lateralization of function for arithmetic and to then compare it with prior results using the same techniques to study language lateralization.

Within the domain of mathematical cognition, the dominant perspective is that lateralization of function and its relationship to language depends heavily on the numerical skill being tested (Dehaene and Cohen, 1995). If the skill is retrieval of exact facts from memory (e.g., during multiplication), then this has been linked to verbal abilities and has been more associated with language processing and the left hemisphere. However, if the numerical skill being tapped is instead more about approximation or basic number comparison (e.g., judging the size of relative quantities), then there is less association with language and, in turn, more association with bilateral function. The evidence for each of these claims follows, starting with the relationship between lateralization of language function and memorized arithmetic facts (typically multiplication), and then proceeding through the evidence that more general numerosity concepts can be independent of left-lateralized language abilities.

In the particular domain of mathematical expression processing, calculation impairments in patients with unilateral brain lesions have been relatively more associated with damage to the left hemisphere (Jackson and Warrington, 1986; Rosselli and Ardila, 1989; Ashcraft et al., 1992). Furthermore, in patients with severed corpus callosums (i.e., “split-brain” patients), when the right hemisphere alone is forced to perform arithmetic, severe impairments are reported, particularly in multiplication (Gazzaniga and Smylie, 1984; Funnell et al., 2007). In contrast, the isolated left hemisphere is typically able to perform these arithmetic tasks above chance. When patients with left hemisphere lesions were specifically examined for the extent and type of impairments to their language and numerical skills, it was found that more severe language impairments were correlated with impairments in quantitative abilities – again, particularly for multiplication (Delazer et al., 1999). This evidence forms the basis for the argument that arithmetic fact retrieval is dependent on left-lateralized language processing abilities.

That language processing, and production in particular, tends to be lateralized to the left hemisphere is one of the most well-known features of brain organization (e.g., Geschwind and Levitsky, 1968). However, in some individuals, language production is instead lateralized to the right hemisphere (Binder et al., 1996; Bishop, 2013). An interesting question, then, is what happens to arithmetic skills when language is lateralized to the right hemisphere? In patients for whom the lateralization of both language and arithmetic skills was assessed, language lateralization influenced the arithmetic abilities of each hemisphere. Specifically, if language was left-lateralized, then the isolated right hemisphere was impaired at performing multiplication, whereas if language was right-lateralized or bilateral, then the isolated right hemisphere was above chance at multiplication (Delazer et al., 2005). This suggests that it is not just a coincidence that both language and arithmetic skills tend to be left-lateralized. Instead, they seem to track each other, either due to a shared reliance on a higher-level processing mechanism that itself tends to be lateralized, or because arithmetic skills actually depend on language abilities (as is suggested by the association between level of language impairment and level of arithmetic impairment).

The most extreme versions of this conclusion – that, without language, there can be no arithmetic fact knowledge, and that the contralateral hemisphere has no involvement in arithmetic processing – is unlikely and must still contend with evidence for dissociations between these abilities. For example, there is evidence from a TMS study that both hemispheres are causally involved in answer generation for multiplication problems (Andres et al., 2011). There has also been a long history of reporting case studies of patients for whom language and arithmetic abilities are dissociable. In several cases of patients with semantic dementia, which is a progressive neurodegenerative disorder featuring loss of semantic memory (especially in word comprehension), there are reports of successful retention of some (or all) arithmetic skills (Diesfeldt, 1993; Cappelletti et al., 2001, 2012; Crutch and Warrington, 2002). However, even in these reports, it is often the case that multiplication is the most impaired arithmetic ability when verbal memory is compromised. Thus, although there is some evidence implicating right hemisphere involvement in production of multiplication problem answers, there seems to be a strong relationship between left hemisphere language abilities and the production of answers to multiplication problems.

Outside of the domain of producing exact answers to arithmetic problems from verbal memory, and in the domain of symbolic and non-symbolic number comparison, the dependence on left-lateralized language abilities is much less apparent. Instead, there are reports that both hemispheres can compare symbolic numbers for their relative size (Colvin et al., 2005). Indeed, there is evidence directly linking developmental changes in subregions of the right hemisphere's parietal lobe with successful acquisition of symbolic and non-symbolic magnitude judgment abilities (Holloway and Ansari, 2010), and evidence from TMS demonstrating that disruption of these right hemisphere areas results in lower performance on tests of automatic magnitude judgment abilities (Cohen Kadosh et al., 2007).

Along with magnitude judgments, the right hemisphere also seems to be able to engage in generating answers to arithmetic problems through approximation or through step-by-step deliberative procedures. For example, the same left hemisphere lesion patients who can fail to report multiplication answers often retain explicit knowledge of arithmetic operations and the ability to deliberately apply this knowledge in addition and subtraction (e.g., Cohen et al., 2000). In one case study, errors in addition and subtraction were only small distances from the correct answer, suggesting that they were produced through tracking and manipulating conceptual quantities and magnitudes (Funnell et al., 2007).

In sum, it would appear that the right hemisphere's numerical ability is limited to answer approximation and magnitude judgments, leaving the performance of exact recall of mathematical arithmetic facts as the domain of the left hemisphere (Dehaene et al., 1999). However, there are still several aspects of this research field which remain relatively unexamined: (1) the specializations of the hemispheres seem to rely on skills that unfold under different timescales, but no time-sensitive neural measures are reported, and (2) the reports at the level of the hemisphere (i.e., in commissurotomy patients) rely on some kind of end-state explicit report from the participant, which renders it difficult to interpret failures to succeed because it is unclear at what level of processing the failure occurred. In general, these gaps leave it ambiguous whether, for example, the right hemisphere generates exact correct answers and then undergoes another process that renders those answers difficult to access and report, or whether the right hemisphere simply is not a reliable independent source from which to generate the exact answers to multiplication problems.

At present, since there are no studies using time-sensitive measures to understand how the processing of equations dynamically unfolds across the hemispheres, these questions remain open. Addressing this gap is critical given how quickly the answer evaluation process occurs – people are able to verify the correctness of simple arithmetic problems easily after less than a second of exposure to the problem, but whether there are differences in exactly when and how this occurs in each hemisphere is currently undetermined. Fortunately, a successful and productive technique for addressing this type of question has already been developed in order to answer similar questions in the language domain.

Indeed, over a decade of research has combined visual half-field presentation with event-related potentials (ERPs, a time-sensitive neural measure) to examine how the hemispheres differently apprehend words for meaning under different types of priming conditions (e.g.,

Kandhadai and Federmeier, 2010a,b; Dickson and Federmeier, 2014) and different types of sentential contexts (e.g., Federmeier and Kutas, 1999b; Wlotko and Federmeier, 2007). In the visual half-field technique, items of interest are presented to subjects lateralized to either visual field, which provides processing advantages to the contralateral hemisphere (i.e., items presented in the periphery of the left visual field will be preferentially apprehended by the right hemisphere, and vice versa for items in the right visual field.). Although the hemispheres are still connected via the corpus callosum, such that interhemispheric communication will eventually be possible, the advantage afforded by direct stimulation is substantial and what information does become shared is both delayed and degraded (see Banich, 2002, for review).

Under central presentation conditions, a handful of studies have used ERPs to examine arithmetic processing as it unfolds, typically with an equation presented item by item, followed by an answer that is presented for verification (e.g., Niedeggen and Rösler, 1999, Niedeggen et al., 1999; Jasinski and Coch, 2012). Researchers have then used this paradigm to examine the time course with which the provided answer's fit to the prior equation is assessed, enabling them to answer questions about when and how the system appreciates that the answer was correct or incorrect.

At present, there are no such studies using this approach to examine how the time course of mathematical answer evaluation proceeds within each of the two hemispheres. The current paper's primary goal, then, is to begin to fill this gap in the literature by recording ERP responses to answers to multiplication problems when the processing is biased to either hemisphere. Multiplication was chosen as the operation since we wanted to draw comparisons with hemispheric differences in language comprehension, and this operation is most closely associated with verbal language processing (as previously reviewed). Due to the relative novelty of the particular manipulations and stimulus set, we will also report the ERP responses to the same answers in standard, central visual presentation as well.

Since there is no ERP literature looking at processing asymmetries for arithmetic problems, we will describe what has been found in language studies to provide a basis of comparison. ERP studies of language processing have consistently shown that both hemispheres can distinguish between expected and unexpected sentence-final words (Federmeier and Kutas, 1999a, Federmeier et al., 2005). If the verbal retrieval that occurs during multiplication is similar, then this suggests that both hemispheres might similarly be able to categorize correct from incorrect answers to arithmetic problems, counter to the suggestion that only the left hemisphere has access to these exact answers.

However, the effect of expectancy/correctness might not manifest on the same component in arithmetic and sentence processing. In language studies, expectancy effects are seen on a component linked to semantic memory access, the N400. It is unclear whether similar N400 effects obtain in arithmetic processing: although some early work identified effects of correctness (in the form of more positive responses to correct than incorrect answers) as N400 effects, more recent studies have made a strong case that the effect is actually driven by a form of the P300/P3b, which reflects categorization and updating of working memory and is larger (more positive) for target items (Donchin and Coles, 1988; see Jasinski and

Coch, 2012, for an elegant re-examination of the component identification in this paradigm; and, for an example of a typical-looking target P300 response to correct answers when participants have sufficient time to generate the answer, see Fig. 2 in Domahs et al., 2007, and compare it to Fig. 1, when the participants have less time to prepare for the answer). Notably, the P300 is often thought to occur as part of a *complex*, with the P3b accompanied by a preceding/overlapping anterior N2 and then followed by a slow wave (e.g., Sutton et al, 1965). Thus, it is possible that a negativity (the N2, linked to motor inhibition) also contributes to the overall magnitude of the correctness effect, but, interestingly, no arithmetic paradigm (including the present study) has independently manipulated it. As such, we will refer to the effect of correctness as a P300, but readers should note that there could be multiple subparts of this complex that contribute to its summated size.

Part of the reason to think that the kind of effects seen in arithmetic processing reflect the P300 rather than the N400 (and are thus different from language effects) is that such effects can arise early in some circumstances (e.g., Salillas and Wicha, 2012) and, across the literature, are variable in timing. In contrast, perhaps the most defining feature of the N400 is its temporal stability (Kutas and Federmeier, 2011), and, hence, N400 effects in sentences do not show similar early onsets or temporal variability. P300 effects, instead, are well known to shift in time, since P300 latency is linked to stimulus evaluation time and, as such, tends to correlate with reaction times (Kutas et al., 1977). Thus, we expect to see a P300 effect related to correctness and will be focusing on its amplitude and timing. A relevant first question to address, then, is whether both hemispheres categorize correct answers as correct (and distinct from incorrect answers that are more easily confusable with correct answers), and, if they do, whether this categorization completed at the same speed in each hemisphere, as indexed by P300 latency.

Whereas basic appreciation of congruity seems to be similar across the hemispheres in language (and remains to be examined in the context of arithmetic), other manipulations on sentence-final completions have revealed striking hemispheric differences. In particular, subtypes of unexpected sentence final words are processed differently in the two hemispheres. When unexpected endings that came from the same semantic category as the expected ending (e.g., “apples” when the expected ending was “oranges”) were presented as well as unexpected endings that came from a different semantic category as the expected ending (e.g., “carrots”, which are not as closely related to “oranges”), only under right visual field/left hemisphere (RVF/LH) presentation conditions was there facilitation on the N400 for the semantically related unexpected ending. In the left visual field/right hemisphere (LVF/RH), the unexpected, within-category item (“apples”) was processed similarly to the across-category item (“carrots”), consistent with an interpretation in which only the left hemisphere pre-activated the semantics including (and surrounding) the most expected final word. In sum, to the LVF/RH, both subtypes of unexpected items were distinguished from the expected items (and were not different from each other), whereas the RVF/LH distinguished between all three types of completions.

In the interest of comparing these kinds of verbal memory processes across domains, we included two types of unexpected (incorrect) answers, one of which might be more related in long-term memory to the expected ending than the other. Drawing on the existing

multiplication fact retrieval and production literature, which suggests that knowledge of multiplication answers is organized in a table-like fashion (wherein multiples of a given operand are relatively more confusable with each other than with multiples of another operand), we included both table-related and table-unrelated incorrect answers, to be compared with correct answers. With this manipulation, incorrect related answers are correct answers for a different problem that shares an operand with the presented problem (e.g., given the problem setup " 6×7 ", an incorrect related answer would be "36", the correct answer to " 6×6 "), whereas incorrect unrelated answers do not have this property (e.g., given the problem setup " 6×7 ", an incorrect unrelated answer would be "40", which is not evenly divisible by 6 or 7 but is still numerically close to the correct answer of "42"). In production studies, the most common type of error is this type of incorrect related answer (which also takes longer to reject than the incorrect unrelated answers in verification paradigms), presumably because of its heightened confusability in long-term memory with the correct answer (Campbell and Graham, 1985; Campbell, 1987).

Neural evidence for how sensitive each hemisphere is to this type of manipulation is limited, and, in particular, there is no literature (to our knowledge) on how patients who have undergone commissurotomies respond in their separate hemispheres to these answer types. If one takes the existing evidence to indicate that *all* multiplication fact knowledge is relatively more left-lateralized, including table-relatedness, one might expect only RVF/LH differentiation between incorrect answers that are unrelated to the context generating the correct answer and incorrect answers that are highly related to the context generating the expected correct answer.

Additionally, at a mechanistic level, the left hemisphere relatedness effect in language is often argued to reflect the pre-activation of the expected ending. If the same predictive mechanism is at play during arithmetic processing, then this reasoning would also predict a relatedness effect for arithmetic answers selective to the left hemisphere. On the other hand, some of what the right hemisphere has been argued to do in arithmetic tasks is more methodical and deliberative – breaking down the problems into steps, since direct retrieval mechanisms seem not as effective for it (e.g., Funnell et al., 2007). This kind of processing could potentially also lead to a relatedness effect, but it would likely unfold more slowly across time. Thus, identifying when and how the hemispheres might differently distinguish between subtypes of incorrect answers forms not only an important inroad to relating lateralization during arithmetic processing with lateralization during language processing, but also will provide novel information about the nature of relatedness effects and contextual congruency in arithmetic processing in each hemisphere in general.

Once again, the precise component that might show this differential effect across hemispheres is unlikely to be the same in arithmetic and language. When incorrect unrelated answers have been presented in ERP studies in central vision, the majority of the effect manifests after the initial P300, in the form of a late positivity for incorrect unrelated items relative to incorrect related items. Again, initial reports identified the effect as arising on the N400 and drew inferences about the structure of semantic memory for table-relatedness (Niedeggen and Rösler, 1999; Niedeggen et al., 1999); however, the co-occurrence of the P300 in the same timewindow renders an independent assessment of N400 amplitude

difficult. The bulk of the effect of relatedness is typically much later (see Domahs et al., 2007, for an example of a late relatedness effect). As such, the critical measure for sensitivity to the factor of relatedness is likely to be in comparisons of the size (or presence) of this late positivity in each hemisphere.

Our findings will thus add important new knowledge about how each hemisphere represents mathematical knowledge and uses that information during online processing. If we find that both hemispheres are equally sensitive to answer correctness, then that would suggest that the poorly performing right hemisphere (in patients with left hemisphere damage) is able to recognize the correct answer at some level but then experiences a bottleneck in communicating that knowledge externally (in contrast with the common interpretation that the right hemisphere on its own doesn't extract this information). In terms of sensitivity to answer relatedness, a finding that the left hemisphere is most sensitive to this variable would be consistent with the idea that multiplication facts are stored in a left-lateralized verbal memory store. On the other hand, given that the right hemisphere has been linked with more algorithmic processing of arithmetic problems, it also seems possible that the right hemisphere could show relatedness effects. In combination with a lack of an effect for relatedness in the left hemisphere, such a finding would indicate that the right hemisphere plays a particularly important (and, as yet, underspecified) role in thoroughly analyzing answers.

2. Results

2.1 Behavioral data

On the offline multiplication fluency test, participants on average produced 38 correct answers for the 60 problems, which is a speed of about 1.5 seconds per written answer (range: 18–60 correct answers). Most participants did not finish in the allotted time, leaving many problems at the end blank (problems were randomly ordered, with easier and harder single-digit problems mixed). The accuracy of reported answers was 98% (most participants, 20/30, did not write any wrong answers). Thus, results of the test found that participants' multiplication fluency was high enough not to be a concern for the validity of interpreting the ERP manipulations, and that, as one would expect, there was some (minor) natural variability in multiplication fact retrieval ability across individuals.

2.2 Event-related potentials

In all of the results below, interactions and effects of electrode-related factors are only reported when of theoretical relevance. Moreover, interactions with topographical factors are reported to convey general distributional information, and readers should consult Figures 4 and 5 (topographical effect maps) to aid with interpretation of the obtained interactions involving these factors.

For all F tests reported with more than 1 degree of freedom in the numerator, reported p-values are Huynh-Feldt corrected for potential violations of sphericity. Although we do not report Greenhouse-Geisser corrected values, because they tend to be overly conservative for

our purposes, we note that using them would not change any of our critical results or conclusions.

2.2.1 General waveform morphology—In Figure 1, one can observe the success of the visual field manipulation. By plotting the lateral occipital channels, the effect of visual stimulation in the contralateral visual field is apparent on the N100 component. In the right lateral occipital channel, the visually-evoked N100 is larger for left visual field presentation, reflecting RH stimulation from the contralateral left visual field, and the same is true of N100 in the left lateral occipital channel.

In Figure 2, one can observe over posterior channels that a visually evoked P200 component is generated in response to the onset of the critical item, with an average latency of 235ms (measured over the range from 150 to 300ms post-onset). This effect is almost immediately followed by another positivity, the P300, which is fairly punctate and is over by 400ms post-stimulus. Following the P300, there is a more prolonged late effect that peaks around 500ms post-stimulus onset.

2.2.2 P300 for Correct Answers—Latency analysis of the P300 effect over all channels found that its peak was 311ms (SE = 0.6) after presentation of the critical answer stimulus (measured over the range from 250 to 400ms post-onset), with no effect of Visual Field of presentation ($F(2,58) = 0.73$, $p = 0.49$, $\epsilon = 0.92$) due to remarkably similar timecourses (RH/LVF peak at 309 ms, SE = 1.1; LH/RVF peak at 310 ms, SE = 1.1, Central presentation peak at 314 ms, SE = 1.0).

In order to characterize P300 distribution, we ran an omnibus ANOVA on mean values centered around the effect, from 250 to 400ms, with 3 levels of Visual Field (Right VF/Left Hemisphere, Left VF/Right Hemisphere, and Central Presentation), 3 levels of Condition (Correct, Incorrect Related, Incorrect Unrelated), 3 levels of Laterality (right lateral, central midline, left lateral), and 4 levels of Anteriority (from prefrontal to occipital channels).

This identified a main effect of Condition ($F(2,58) = 9.11$, $p = 0.0004$, $\epsilon = 0.98$), and interactions between Condition and Laterality ($F(4,116) = 3.38$, $p = 0.04$, $\epsilon = 0.60$), Condition and Anteriority ($F(6,174) = 3.90$, $p = 0.025$, $\epsilon = 0.41$), and Condition, Anteriority, and Laterality ($F(12,348) = 5.20$, $p > 0.0001$, $\epsilon = 0.55$). The interactions between Condition and distributional factors reflected that the P300 effect, while widespread, was largest in mediocentral channels, and was quite small in frontal channels (see Figures 4 and 5 for topographical maps of the distribution).

Planned comparisons were then performed to examine Correctness (Correct compared to Incorrect Unrelated answers) and Relatedness (Incorrect Related compared with Incorrect Unrelated answers) separately, in two ANOVAs including 3 levels of VF, over 21 channels that captured the effect (i.e., five frontal channels were dropped). These tests confirmed that there was a significant effect of Correctness ($F(1,29) = 14.39$, $p = 0.0007$) and no effect of Relatedness ($F(1,29) = 0.04$, $p = 0.852$). There was no effect of Visual Field in either the test of Correctness ($F(2,58) = 0.27$, $p = 0.766$, $\epsilon = 0.791$) or the test of Relatedness ($F(2,58) = 0.32$, $p = 0.577$, $\epsilon = 0.745$). In order to report the size of the P300 effect in microvolts, an

effect contrast (i.e., difference wave) was computed by subtracting the response to Incorrect Unrelated answers from Correct answers (Correctness effect) within each subject. This revealed that the main Correctness effect (collapsed across VFs) was due to larger P300s to Correct answers relative to Incorrect Unrelated answers ($\mu V = 0.86$, $SE = 0.05$, $d_z = 17.2$).

2.2.3 Overall effect of Late Positivity—In order to characterize the distribution of the late positive effect, we began by running an omnibus ANOVA with 3 levels of Visual Field (Right VF, Left VF, and Central Presentation), 3 levels of Condition (Correct, Incorrect Related, Incorrect Unrelated), 3 levels of Laterality (right lateral, central midline, left lateral), and 4 levels of Anteriority (from frontal to posterior channels) over the time period 400–600ms following presentation of the critical answer stimulus.

This revealed main effects of Visual Field ($F(2,58) = 4.44$, $p = 0.017$, $\epsilon = 0.91$), Condition ($F(2,58) = 9.32$, $p = 0.0004$, $\epsilon = 0.84$), Laterality ($F(2,58) = 49.89$, $p < 0.0001$, $\epsilon = 0.65$), and Anteriority ($F(3,87) = 34.75$, $p < 0.0001$, $\epsilon = 0.59$). Additionally, there was a significant four-way interaction between all factors (Visual Field, Condition, Anteriority, and Laterality), ($F(24,696) = 1.8$, $p = 0.036$, $\epsilon = 0.59$), reflecting that effects of Condition between visual fields were distributed unevenly across the scalp. Specifically, when Condition-related effects were present in a VF, they were largest over central and posterior sites, and reduced or absent in other locations.

In order to test prior planned comparisons for the effect of Correctness (Correct versus Incorrect Unrelated) and Relatedness (Incorrect Related versus Incorrect Unrelated) in each hemisphere/VF on the LPC, a cluster of 13 electrodes from the peak of the effect (as indicated by the distributional analysis, see Figure 5) is therefore used in the following analyses.²

2.2.4 Central Presentation—There was a significant effect of Correctness ($F(1,29) = 4.67$, $p = 0.039$). This effect was due to more positivity in response to Incorrect Unrelated than to Correct answers, as found in an effect contrast (i.e., difference wave) subtracting Correct answers from Incorrect Unrelated answers within subject ($1.04 \mu V$, $SE = 0.15$, $d_z = 6.93$). There was no significant effect of Relatedness ($F(1,29) = 0.57$, $p = 0.46$).

2.2.5 Right Hemisphere—There was both an effect of Correctness ($F(1,29) = 9.68$, $p = 0.004$) and Relatedness ($F(1,29) = 7.94$, $p = 0.009$). These effects were driven by significantly more positive responses to Incorrect Unrelated answers compared to both Correct answers and Incorrect Related answers, as found in effect contrasts subtracting Correct answers from Incorrect Unrelated answers ($1.46 \mu V$, $SE = 0.14$, $d_z = 10.43$) and Incorrect Related answers from Incorrect Unrelated answers ($1.06 \mu V$, $SE = 0.12$, $d_z = 8.83$). Numerically and visually (see Figure 3), the effect of Correctness was larger than Relatedness, although a post-hoc follow-up comparison of Incorrect Related and Correct

²The 13 channels used included right lateral, left lateral, and middle occipital channels (RLOc, LLOc, MiOc), left and right medial occipital channels (LMOc, RMOc), left and right dorsal parietal channels (LDPa, RDPa), midcentral and midparietal channels (MiCe, MiPa), left and right midcentral channels (LMCe, RMCe), and left and right midfrontal channels (LMFr, RMFr).

answers found no difference between these conditions in this timewindow ($F(1,29) = 1.64$, $p = 0.21$).

2.2.6 Left Hemisphere—There were no effects of Correctness ($F(1,29) = 2.46$, $p = 0.128$) or Relatedness ($F(1,29) = 0.54$, $p = 0.470$) in this timewindow following RVF/LH presentation. The absence of effects as compared to LVF/RH can be seen in Figures 3 and 5.

3. Discussion

Our data from central presentation generally replicates what has been previously shown in prior papers using similar manipulations, despite moderate changes to the precise stimulus design (e.g., our incorrect answers of both types were generated by different means than has been used before). Specifically, we replicate the presence of an early distinction between correct and incorrect answers in the form of a larger P300 to correct answers (although this same effect has a history of being labeled an N400 by others)³. Importantly, the actual waveform morphology of a positivity to correct answers is the same across these studies (see Jasinski and Coch, 2012). In the broader literature, the P300 is enhanced to target items, even when probability is held constant, and its latency is generally taken as a measure of stimulus evaluation time. The larger P300 to correct items, peaking around 300ms in this experiment, implies that these items were identified and categorized as the target item quite quickly, likely because there was a full second between onset of the second operand and onset of the final answer. In cases with reduced time to predict and expect the final target answer, the P300 has been delayed in time – hence occurring more in the traditional N400 timewindow and/or reduced in amplitude (indeed, if at all present, e.g., Domahs et al., 2007, SOA 150ms; Niedeggen et al., 1999, SOA 200ms). Our result in central presentation implies that the system had sufficient time to process the problem prior to presentation of the answer.

That we did not find support for a relatedness effect in central vision suggests that the generation of this effect is relatively more sensitive to nuances of stimulus design. The timing of presentation is one potential reason for a lack of a significant late positivity: effects of relatedness might be smaller when the system has more time to fully process the problem and retrieve the correct answer. Alternatively, the difference could have arisen because in constructing our incorrect answers, we carefully matched the distribution of even and odd numbers across the related and unrelated conditions. This does not seem to be a feature of the larger literature in the area, leaving open the possibility that sensitivity to relatedness might be partially a sensitivity to those answers being more like correct answers on the dimension of evenness, compared to incorrect unrelated answers (which are more likely to be odd given typical methods of stimulus construction). Given that the P300 is sensitive to subjective probability, it is possible that the relative evenness and oddness of answers across conditions could be influencing its amplitude across studies, although the result of direct manipulation of these variables has not yet been examined.

³In unpublished follow-up work, we have replicated and extended the findings presented here in a manner that supports our identification of the P300 in this paradigm.

Our primary aim was to characterize how the effects of correctness and relatedness might differ as a function of hemisphere, given the prevalence of differences across patient groups with unilateral lesions, (Rosselli and Ardila, 1989; Cohen et al., 1999; for review, see Cappelletti and Cipolotti, 2010) and given the effects of lateralized presentation in patients who have undergone commissurotomies (Gazzaniga and Smylie, 1984; Colvin et al., 2005; Funnell et al., 2007). Through examining these effects, we are able to address two fundamental questions: 1) How well is each hemisphere able to recognize and accurately categorize correct answers from incorrect answers? and 2) How sensitive is each hemisphere to the putative table-based organization of knowledge about mathematical facts, as primed here by different problem contexts? Interpretation and discussion of our results follows for each issue in turn, with consideration given to the corresponding findings in sentence processing.

We found that both hemispheres were able to recognize and categorize the correct answer as correct (indicated by a larger P300 to this condition than to incorrect answers), and were able to do so at the same speed (indicated by no latency difference in the P300 across visual field presentations). In order to reconcile this with prior reports that the right hemisphere is less critical for production of answers than is the left (i.e., patients with left hemisphere lesions are, consistently across studies, more likely to have difficulties with arithmetic), it is important to remember that our recordings do not reflect the hemisphere's abilities to output their knowledge in the form of a response (verbal or otherwise). Instead, our ERPs reflect each hemisphere's appreciation for the answer types in the absence of any form of overt response (which was only requested after a substantial delay). It is well documented that the P300 specifically indexes stimulus evaluation time, and not response selection and planning/execution, which are otherwise difficult to disentangle in reaction time measures (Kutas et al., 1977). Our findings are consistent with the interpretation that the hemispheres are able to recognize correct from incorrect answers at the same timescale, but subsequently undergo different processes when selecting and executing outward expressions of this internal categorization process.

This pattern aligns with typical results in studies of language comprehension, in which participants passively read sentences and that end in expected or unexpected ways. Both hemispheres have been shown to be able to distinguish "correct" final words from "incorrect" final words (and consistent / inconsistent line drawing images as well), albeit on a different ERP component (the N400, Coulson et al., 2005; Federmeier et al., 2005; Wlotko and Federmeier, 2007). The similarity in response patterns across hemispheres in those studies has been taken as evidence that both hemispheres are able to make sense of the provided sentential context and integrate new words into the message. Our findings here are similarly consistent with the idea that both hemispheres are able to take the context provided by mathematical expressions (multiplication problems) and recognize whether provided answers are good fits or poor fits, although ultimately they might not be equally able to overtly communicate that knowledge (as is also true in language).

That this main effect of congruency occurs on different components across language and arithmetic studies is intriguing. It is perhaps understandable why sentence processing studies typically report no P300 effects. The goal in these experiments tends to be to "read for

understanding”, not to read and make an explicit judgment about the correctness of the critical (typically sentence-final) word. When people do make explicit, categorical judgments about words in a sentence (or at least about the sentence as a whole) then the critical words often elicit positivities as well (see, e.g., Geyer et al., 2006, and Kim and Osterhout, 2005, for positivities elicited in the context of plausibility judgments). Further, it is possible that some part of the reported P300 effect here includes contributions from a negativity (N2), largest for incorrect answers, which is often associated with inhibitory control mechanisms (i.e., preventing the participant from pressing “yes” to incorrect answers). This is also not a typical cognitive process demanded by typical sentence reading contexts (although it was recently reported in a self-paced reading paradigm, Payne et al., 2016).

What is harder to reconcile is why there is not also an N400 effect of congruency in the context of math equation processing when digits are used as stimuli. Arabic numerals presumably convey meaning (in the form of numerosity and other long term memory representations), but perhaps this information is not fully appreciated by the N400 timewindow. We have previously observed that some features of wordforms that have been thought to be part of the meaning of the word only are appreciated later in the form of a late positive complex (LPC), and the same might be true of the nuanced meaning of symbolic numbers (Fischer-Baum et al., 2014).

On the other hand, it is possible that the P300 effect is overlapping with a co-occurring (smaller) N400 effect, and that the effect of semantic facilitation is present, but hard to detect. Given that all potentially meaningful stimuli elicit N400 components, and that digits fall within that category, we believe this to be possible (Kutas and Federmeier, 2011). However, the effects in the present study are clearly unlike typical N400 effects of semantic facilitation in both morphology (with a distinct positivity for the “facilitated” correct condition) and in timing (occurring almost 100ms prior to the traditional N400 timewindow.) To more clearly examine semantic processing of digits, the experimental design would need to contain tasks that do not also elicit substantial P300 effects.

Regarding sensitivity to subtypes of incorrect answers, we only find evidence for an effect of relatedness in the LVF/RH (in the form of a larger late positivity for incorrect unrelated answers compared to incorrect related answers), and no significant difference between these answer types in either the left hemisphere or in central presentation (although there was a numerical trend for the effect during central viewing). This result provides the first evidence, to our knowledge, that, whereas the right hemisphere appreciates a given incorrect answer’s similarity to the correct answer, the left hemisphere does not.

This pattern is a remarkable reversal of what has been previously reported for a similar manipulation in sentence processing, in which unexpected sentence-final words that were semantically related to the expected completion were only distinguished from unexpected unrelated words following presentation to the RVF (i.e., the left hemisphere had a relatedness effect) (Federmeier and Kutas, 1999b). In that work, it was instead the right hemisphere that had a more categorical approach in comprehension of completions. However, at a mechanistic level, direct comparisons of these results to those reported here is

difficult because, whereas the effect of relatedness in sentence processing was observed on the N400, the effect of relatedness we report is on a late positivity.

Effects in language processing on late positivities are common across studies, although the particular eliciting conditions and nomenclature for the subsequently reported effects varies (see Van Petten and Luka, 2012, for a review of the ongoing differences of opinion and complicated discussions unfolding in this domain). A few studies have looked at late positivities in each of the hemispheres. In one, there was greater late positive complex (LPC) priming for backward ordered word pairs in the left hemisphere (e.g., “meat” followed by “butcher”, where “meat” does not strongly prime “butcher”, but the words are clearly related) than in the right, which suggests that there is a left hemisphere advantage for strategically reshaping meaning activation for weakly related and/or non-canonically ordered word pairs (Kandhadai and Federmeier, 2010a). In a subsequent publication, it was observed that the right hemisphere could be forced to appreciate the relatedness in these types of word pairs (Kandhadai and Federmeier, 2010b), but, on the whole, in language contexts there seems to be a clear bias favoring the *left* hemisphere for sensitivity to semantic relatedness on late positive effects. Our result of a right hemisphere relatedness effect thus serves as strong evidence that the memory system engaged by multiplication arithmetic problems and the memory system engaged in sentence and word reading can be dissociable, and that their interrelationship is thus more complicated than it might appear on the surface.

Because there are currently no published reports of late positivities *selective* to the right hemisphere (to our knowledge), and because there are also no studies looking specifically at behavioral responses to these particular conditions with visual half-field presentation, there are many alternative interpretations for this pattern. That is, at a mechanistic level, the source of the right hemisphere appreciation of related answers could be the result of a number of differences in how learned information is structured and/or in how incoming information is processed by each hemisphere. We will therefore review three possible sources of the effect below.

Under a first interpretation, this pattern could arise because arithmetic fact knowledge is organized in a table-like fashion in the right hemisphere and not organized in this way in the left hemisphere. Then, when related answers are provided, only the right hemisphere is prepared to recognize this property. Although possible, there is no evidence suggesting this in the larger literature, and our finding for similar effects of correctness across the hemispheres would seem to weigh against the idea that arithmetic knowledge is structured fundamentally differently in the hemispheres. However, a second possibility is that the left hemisphere also contains the table-relatedness information but does not access it in this context. In fact, this might be advantageous: it could be harmful for the successful execution of the task at hand to be entertaining multiple answers after seeing the full arithmetic expression. Under this view, the left hemisphere could be strategically focusing on the search for only the correct answer. Interestingly, this idea is consistent with the proposal that the right hemisphere generates a broader range of concepts than does the left (Jung-Beeman, 2005), which, in this case, would imply that the right hemisphere activated a broader range of potential answers. If the right hemisphere does this, it might explain some of its apparent difficulty in response selection, as seen in commissurotomy patients performing poorly in

verification tasks for problems presented to the right hemisphere (e.g., Funnel et al., 2007). However, in the language domain, multiple direct tests of the coarse coding hypothesis have failed to find support for it (Kandhadai and Federmeier, 2007, 2008). Also problematic for this view is that one might have expected to see earlier effects of relatedness on the P300 in the right hemisphere, reflecting a lack of focus on the correct answer as a target item, which is not what we observed.

A final explanation places the locus of this effect more in the domain of explicit evaluation of the provided answer. This is in contrast to the prior two explanations, which see the source of the difference as arising at the level of memory representations or in the spread of activation across them. Under this third view, once the left hemisphere sees an answer, it categorizes it as correct or not (i.e., generating a P300), and then does no more follow-up evaluation of the presented number. The right hemisphere, instead, could be engaging in a deliberative appreciation for additional features of an answer once it has been presented for consideration. This might be functionally important for fully analyzing arithmetic problems beyond an instinctual recall of correct answers, and having more information about the qualities of the provided answers could in theory be helpful for deliberative mathematical reasoning tasks. At the same time, the left hemisphere's theoretically more categorical approach would be less likely to be vulnerable to error production.

In general, this interpretation of the relatedness effect in the right hemisphere is also consistent with the idea that the calculating right hemisphere is involved in more nuanced evaluations of approximate values (perhaps drawing on enhanced spatial abilities), whereas the left hemisphere's approach to answer evaluation has been attributed to relatively more simple memorized retrieval of mathematical facts (i.e., recognizing the correct answer as correct) (Dehaene and Cohen, 1995; Rosselli and Ardila, 1989). However, it is important to note that we have evidence that the right hemisphere initially does engage in the categorization process that characterizes the left hemisphere; it is only later in time that the right and left hemispheres' processing strategies diverge, and the right begins what appears to be a process of appreciating more detailed properties of the proposed answers. On the whole, we therefore find this final interpretation to be most consistent with the existing literature on both arithmetic and language, and it is the only interpretation of the three we introduced that does not also make a prediction that the P300 in the right hemisphere should have a relatedness effect.

One important factor that has not been discussed thus far is that of the amount of time our participants had to process the full multiplication expression prior to the answer appearing on the screen. This one-second interval between seeing the second operand (i.e., the complete problem) and seeing the answer is not very pressured, especially compared with some paradigms wherein this interval was reduced to 150ms (in particular, in Domahs et al., 2007; this presentation latency visually appears to have substantially delayed and reduced the amplitude of the P300 compared to the when the P300 was elicited with a 550ms SOA). Moving forward, then, it could be of interest to see how the hemispheres respond when they are under more pressure (i.e., in speeded response conditions, and/or with less time between viewing the operands and viewing the answers). It is possible that the left hemisphere, which, in the language domain, has been associated with rapid top-down processing of input,

might be differently affected by such a manipulation than would the right hemisphere, which has been associated with slower, more bottom-up processing (Federmeier, 2007).

This work is the first to characterize the temporal dynamics of answer evaluation processing in the two hemispheres, and, as such, our findings make a novel and important contribution to the growing literature in the neurobiology of arithmetic processing. We found that, similar to language, both hemispheres are sensitive to whether or not an expression has been completed in an expected / correct or unexpected / incorrect manner, although the mechanisms supporting this sensitivity seem to be functionally different. Most intriguingly, we uncovered a unique sensitivity to properties of the answer, beyond their status as correct or incorrect, that occurs later in time only in the right hemisphere, which appears to reflect its more deliberative processing strategy. Given that a similar effect has not yet been observed in language processing, these results are consistent with a view that there are important differences in how we read and assess sentences in language and expressions in arithmetic problems.

4. Experimental Procedures

4.1 Participants

Data were analyzed from 30 young adults who were recruited from the University of Illinois community and compensated with either course credit or cash (\$10/hour). Participants consented in accordance with the Institutional Review Board of the University of Illinois at Urbana-Champaign. All participants were right-handed native monolingual English speakers (15 female, 15 male) with normal or corrected-to-normal vision, and no early exposure to a second language, history of brain trauma, or current use of psychoactive medications. Their mean age was 20 years old (Range: 18–23), and their mean laterality quotient on the Edinburgh handedness inventory (Oldfield, 1971) was 0.85 (Range: 0.45 – 1.00, with 1.0 indicating strongly right-handed and –1.0 strongly left-handed). Nine participants reported having a left-handed biological family member.

In order to reach our target of 30 participants, we collected data from a total of 33 participants because three had to be dropped prior to analysis due to excessive noise either in the global recording or in an individual channel.

4.2 Materials

Participants saw familiar multiplication problems presented sequentially on the screen (e.g., $3 \times \dots \times \dots 4 \times \dots = \dots 12$). All 60 critical multiplication problem stems had single digit operands. The only single digit operations that were not included were ones for which creation of all the necessary answer types was not feasible due to how small the operand numbers were: 2×3 (and its counterpart 3×2), and 2×4 (and its counterpart 4×2). Square operations (e.g., 6×6) were included.

Answer types were Correct, Incorrect Unrelated, or Incorrect Related. Since Correct answers are naturally distributed to be even numbers about 75% of the time (odd answers are only the result of an odd number times another odd number), both types of Incorrect answers were forced to match this naturally-existing distribution of even numbers. Thus, critical

Correct answers were 73% even, Incorrect Unrelated answers were 72% even, and Incorrect Related answers were 75% even. The average answer size participants encountered was about 36, including all Correct/Incorrect answer types.

Incorrect Unrelated answers were generated by taking the Correct answer and either adding and subtracting a number between 1 and 5 to it (with rare cases of 6 when 5 or less generated related answers), thus allowing the creation of Incorrect Unrelated answers that were even in the desired distribution and also that deviated from the Correct answer in an unpredictable manner. Incorrect Unrelated answers were also deliberately selected so as not to accidentally be a correct answer to the sum of the operands (e.g., $5 \times 2 = 7$ was not permitted as an Incorrect Unrelated answer).

Incorrect Related answers were always a multiple of the first operand, generated by multiplying it with ± 1 of the second operand. For stems with 9 as the second operand, 8 was used to generate the Incorrect Related answers (since multiples of 10 might be perceived as implausible), and similarly, when 2 was the second operand, Incorrect Related answers were generated by multiplying by 3, not 1.

Half of each type of Incorrect answer was numerically larger than the problem's corresponding Correct answer and the other half was numerically smaller than the Correct answer. On average, Incorrect Related answers were numerically more distant from the Correct answer than were Incorrect Unrelated answers. This is because any time the first operand was above 5, the Incorrect Related answers would necessarily deviate from the Correct answer by a multiple of 6–9, whereas Incorrect Unrelated answers were almost always limited to deviations of 5 from the Correct answer.

There were additionally 30 filler multiplication problems that used 1 as an operand and that used common double-digit operands in very simple problems (e.g., $20 \times 3 = 60$; $2 \times 13 = 26$). The answer to these problems was always correct, such that, from the perspective of the participant, the distribution of Correct/Incorrect answers throughout the experiment was 50/50⁴.

For the 3 Visual Field presentation conditions (LVF/RH; RVF/LH; Central), there were 3 separate sets of 30 critical items – one set for each of the 3 answer types. The 60 critical problem stems were divided such that there were no exact matches in any group of 30. There could be rare cases of operand order swaps (e.g., a set of 30 could have both 3×5 and 5×3).

Because the responses were delayed to prevent artifacts during collection of critical data, and because we wanted to prevent early button-pressing, only a subset of trials were selected for participants to respond to. All 90 filler trials were probed, and a subset of the critical trials was probed (33% of the incorrect answers, evenly distributed between unrelated and

⁴The P300 is sensitive to both subjective and global probability, with larger P300s for less probable target items. A 50/50 probability design here therefore reduces the likelihood of observing particularly large target P300s, but even in 50/50 manipulations, targets elicit larger P300s than non-targets (see Johnson, 1986, for factors influencing P300 amplitude).

related conditions across visual fields). From the participants' perspective, just under half of trials were queried (42% total), and 60% were correct.

Six lists were generated so that each combination of answer type (Correct, Incorrect Unrelated, Incorrect Related) and level of VF of presentation (RVF/LH, LVF/RH, Center) was rotated. For example, on List 1, the 30 problem stems in set 3 might have assigned to have Incorrect Unrelated answers in LVF/RH, whereas on List 2, set 3 might have been assigned to have Correct answers in Central vision. Across each set, the distribution of even/odd answers and the distribution of Incorrect answers that were larger/smaller than the Correct answer were matched as closely as possible (although having full rotation of the sets minimized concerns at this level). Across lists, the order of the 360 problem stems (270 critical: $3 \times 3 \times 30$) was the same.

4.3 Procedure

Equations were presented one element at a time on the screen: Operand 1 ... $\times$... Operand 2 ... = ... Answer. All elements of the equation context were presented centrally, whereas the answers could appear either lateralized or in central vision. Each element was presented for 200ms, then was off for 300ms (i.e., SOA of 500ms). After the answers were presented, there was a 1000ms blank period for ERP collection, during which participants were instructed to refrain from blinking or moving their eyes. After the blank, a subset of trials was queried for correctness. Queries occurred after the answer was presented in any of the visual fields in an unpredictable manner. Since queries were postponed, speed was not emphasized to participants.

After the response, if probed, a series of white “~~~~” were presented for 1200ms, during which subjects could blink and move their eyes. If there was no probe, this blinking period directly followed the blank period. A series of red “~~~~” directly followed the blinking period, serving as a warning that the next critical stimulus was about to appear. This warning signal was presented for an unpredictable amount of time ranging from 400ms to 600ms (to reduce the contribution of sustained expectancy-related potentials to the time-locked average for the critical items). A small square was continuously presented in the center of the screen just below the stimuli to remind participants where the center was; many participants reported focusing on the dot to keep their eyes steady when the lateralized items appeared. There were periodic self-paced breaks in which yellow “~~~~” were presented after a response; this indicated that the subjects could take a break until they were ready to do more trials.

4.4 EEG Recording Parameters

The EEG was recorded using 26 silver/silver-chloride electrodes mounted in an elastic cap. All electrode impedances were kept below 5k , and electrodes were evenly distributed over the scalp (see the icon in Figure 2 for the arrangement). The data were referenced on-line to the left mastoid and re-referenced 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 to 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 subject. Trials were excluded from averaging if they included blinks, movement artifacts, signal drift, blocking, or a horizontal eye gaze movement. Because participants were only queried following some (but not all) trials, all trials were included in the analysis regardless of subsequent accuracy.⁵ For five subjects who had at least 20 blink-contaminated trials, the trials that were excluded only due to blink artifacts were corrected and reintroduced to the average (in a procedure described by Dale, 1994). After artifact rejection, critical bins included 26 trials on average, and no individual subject had fewer than 16 critical trials in any bin.

Epochs of EEG data for each trial were taken from 100ms prior to item onset until 920ms after item onset, and the baseline acquired over the 100ms 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 to 20 Hz.

Component latency values, when reported, were determined by measuring the average time point of the maximum (for a positivity) or minimum (for a negativity) value in the provided time range.

4.5 Arithmetic fluency measure

Following the ERP experiment, participants took a 60-second paper and pencil multiplication test in which they provided answers to as many of the critical single-digit problems in the study as they could. The 60 problems were in the same scrambled order for all participants. In order to minimize test anxiety, this was always described as an “exercise” and never as a test, quiz, or examination.

Acknowledgments

The authors would like to thank Dan Hyde, Monica Fabiani, and Chuck Hillman for their feedback on previous versions of this manuscript. Funding: This work was supported by the James S. McDonnell Foundation and the National Institutes of Health under grant AG026308.

Abbreviations

LVF/RH Left visual field/right hemisphere

RVF/LH right visual field/left hemisphere

References

Andres M, Pelgrims B, Michaux N, Olivier E, Pesenti M. Role of distinct parietal areas in arithmetic: An fMRI-guided TMS study. *Neuroimage*. 2011; 54(4):3048–3056. [PubMed: 21073958]

⁵Trials are not removed on the basis of accuracy in language comprehension / sentence reading tasks either. It is typical in paradigms analyzing expression completions for participants' confidence and interpretation to vary across individual trials.

- Andres M, Seron X, Olivier E. Hemispheric lateralization of number comparison. *Cognitive Brain Research*. 2005; 25(1):283–290. [PubMed: 16005617]
- Ashcraft MH, Yamashita TS, Aram DM. Mathematics performance in left and right brain-lesioned children and adolescents. *Brain and Cognition*. 1992; 19(2):208–252. [PubMed: 1642861]
- Babkoff H, Faust M. Lexical decision and visual hemifield: An examination of the RT—Accuracy relationship. *Neuropsychologia*. 1988; 26(5):711–725. [PubMed: 3211289]
- Banich, MT. The divided visual field technique in laterality and interhemispheric integration. In: Hugdahl, K., editor. *Experimental methods in neuropsychology*. Boston, MA: Kluwer; 2002. p. 47–64.
- Binder JR, Swanson SJ, Hammeke TA, Morris GL, Mueller WM, Fischer M, et al. Determination of language dominance using functional MRI: A comparison with the Wada test. *Neurology*. 1996; 46(4):978–984. [PubMed: 8780076]
- Bishop DV. Cerebral asymmetry and language development: Cause, correlate, or consequence? *Science* (New York, NY). 2013; 340(6138):1230531.
- Broca P. Sur le siège de la faculté du langage articulé. *Bulletins De La Société d'Anthropologie De Paris*. 1865; 6(1):377–393.
- Campbell JI. Production, verification, and priming of multiplication facts. *Memory & Cognition*. 1987; 15(4):349–364. [PubMed: 3670055]
- Campbell JI, Graham DJ. Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*. 1985; 39(2):338–366.
- Cappelletti M, Butterworth B, Kopelman M. Sparing numerical abilities in a case of semantic dementia. *Neuropsychologia*. 2001; 39(11):1224–1239. [PubMed: 11527560]
- Cappelletti M, Butterworth B, Kopelman M. Numeracy skills in patients with degenerative disorders and focal brain lesions: A neuropsychological investigation. *Neuropsychology*. 2012; 26(1):1–19. [PubMed: 22122516]
- Cappelletti, M., Cipolotti, L. The neuropsychology of acquired calculation disorders. In: Gurd, JM, Kischka, U., Marshall, JC., editors. *Handbook of Clinical Neuropsychology*. New York, NY: Oxford University Press; 2010. p. 401–417.
- Chochon F, Cohen L, Van De Moortele P, Dehaene S. Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience*. 1999; 11(6): 617–630. [PubMed: 10601743]
- Christman S, Kitterle FL, Hellige J. Hemispheric asymmetry in the processing of absolute versus relative spatial frequency. *Brain and Cognition*. 1991; 16(1):62–73. [PubMed: 1854470]
- Cohen L, Dehaene S, Chochon F, Lehericy S, Naccache L. Language and calculation within the parietal lobe: A combined cognitive, anatomical and fMRI study. *Neuropsychologia*. 2000; 38(10): 1426–1440. [PubMed: 10869586]
- Colvin MK, Funnell MG, Gazzaniga MS. Numerical processing in the two hemispheres: Studies of a split-brain patient. *Brain and Cognition*. 2005; 57(1):43–52. [PubMed: 15629214]
- Coulson S, Federmeier KD, Van Petten C, Kutas M. Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2005; 31(1):129–147.
- Crutch SJ, Warrington EK. Preserved calculation skills in a case of semantic dementia. *Cortex*. 2002; 38(3):389–399. [PubMed: 12146663]
- Dale, AM. Unpublished Dissertation. University of California; San Diego, La Jolla, CA: 1994. Source localization and spatial discriminant analysis of event-related potentials: linear approaches.
- Dehaene S, Cohen L. Towards an anatomical and functional model of number processing. *Mathematical Cognition*. 1995; 1(1):83–120.
- Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. Sources of mathematical thinking: Behavioral and brain-imaging evidence. *Science* (New York, NY). 1999; 284(5416):970–974.
- Delazer M, Girelli L, Semenza C, Denes G. Numerical skills and aphasia. *Journal of the International Neuropsychological Society*. 1999; 5(3):213–221. [PubMed: 10217921]

- Delazer M, Karner E, Unterberger I, Walser G, Waldenberger P, Trinka E, Benke T. Language and arithmetic—a study using the intracarotid amobarbital procedure. *Neuroreport*. 2005; 16(12):1403–1405. [PubMed: 16056148]
- Dickson DS, Federmeier KD. Hemispheric differences in orthographic and semantic processing as revealed by event-related potentials. *Neuropsychologia*. 2014; 64:230–239. [PubMed: 25278134]
- Diesfeldt HF. Progressive decline of semantic memory with preservation of number processing and calculation. *Behavioural Neurology*. 1993; 6(4):239–242. [PubMed: 24487142]
- Domahs F, Domahs U, Schlesewsky M, Ratnckx E, Verguts T, Willmes K, et al. Neighborhood consistency in mental arithmetic: Behavioral and ERP evidence. *Behavioral and Brain Functions*. 2007; 3(1):66. [PubMed: 18163911]
- Donchin E, Coles MG. Is the P300 component a manifestation of context updating. *Behavioral and Brain Sciences*. 1988; 11(3):357–427.
- Federmeier KD, Kutas M. A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*. 1999a; 41(4):469–495.
- Federmeier KD, Kutas M. Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*. 1999b; 8(3):373–392. [PubMed: 10556614]
- Fischer-Baum S, Dickson DS, Federmeier KD. Frequency and regularity effects in reading are task dependent: Evidence from ERPs. *Language, Cognition and Neuroscience*. 2014; 29(10):1342–1355.
- Funnell MG, Colvin MK, Gazzaniga MS. The calculating hemispheres: Studies of a split-brain patient. *Neuropsychologia*. 2007; 45(10):2378–2386. [PubMed: 17420034]
- Gazzaniga MS, Smylie CS. Dissociation of language and cognition. A psychological profile of two disconnected right hemispheres. *Brain*. 1984; 107(1):145–153. [PubMed: 6697151]
- Geschwind N, Levitsky W. Human brain: Left-right asymmetries in temporal speech region. *Science (New York, NY)*. 1968; 161(3837):186–187.
- Geyer, A., Holcomb, P., Kuperberg, G., Perlmutter, N. Plausibility and sentence comprehension. An ERP study. Poster presented at the annual meeting of the Cognitive Neuroscience Society; San Francisco, CA. 2006 Apr.
- Holloway ID, Ansari D. Developmental specialization in the right intraparietal sulcus for the abstract representation of numerical magnitude. *Journal of Cognitive Neuroscience*. 2010; 22(11):2627–2637. [PubMed: 19929327]
- Jackson M, Warrington EK. Arithmetic skills in patients with unilateral cerebral lesions. *Cortex*. 1986; 22(4):611–620. [PubMed: 3816245]
- Jasinski EC, Coch D. ERPs across arithmetic operations in a delayed answer verification task. *Psychophysiology*. 2012; 49(7):943–958. [PubMed: 22563982]
- Johnson R. For distinguished early career contribution to psychophysiology: Award address, 1985. *Psychophysiology*. 1986; 23(4):367–384. [PubMed: 3774922]
- Hellige, Joseph B. Hemispheric asymmetry: What's right and what's left. Cambridge, MA: Harvard University Press; 1993.
- Jung-Beeman M. Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*. 2005; 9(11):512–518. [PubMed: 16214387]
- Kadosh RC, Kadosh KC, Schuhmann T, Kaas A, Goebel R, Henik A, et al. Virtual dyscalculia induced by parietal-lobe TMS impairs automatic magnitude processing. *Current Biology*. 2007; 17(8):689–693. [PubMed: 17379521]
- Kandhadai P, Federmeier KD. Multiple priming of lexically ambiguous and unambiguous targets in the cerebral hemispheres: The coarse coding hypothesis revisited. *Brain Research*. 2007; 1153:144–157. [PubMed: 17459344]
- Kandhadai P, Federmeier KD. Summing it up: Semantic activation processes in the two hemispheres as revealed by event-related potentials. *Brain Research*. 2008; 1233:146–159. [PubMed: 18675257]
- Kandhadai P, Federmeier KD. Automatic and controlled aspects of lexical associative processing in the two cerebral hemispheres. *Psychophysiology*. 2010a; 47(4):774–785. [PubMed: 20136731]

- Kandhadai P, Federmeier KD. Hemispheric differences in the recruitment of semantic processing mechanisms. *Neuropsychologia*. 2010b; 48(13):3772–3781. [PubMed: 20638397]
- Kutas M, Federmeier KD. Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*. 2011; 62:621–647.
- Kutas M, McCarthy G, Donchin E. Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science (New York, NY)*. 1977; 197(4305):792–795.
- Lau EF, Phillips C, Poeppel D. A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*. 2008; 9(12):920–933. [PubMed: 19020511]
- Martin M. Hemispheric specialization for local and global processing. *Neuropsychologia*. 1979; 17(1): 33–40. [PubMed: 431807]
- Niedeggen M, Rösler F. N400 effects reflect activation spread during retrieval of arithmetic facts. *Psychological Science*. 1999; 10(3):271–276.
- Niedeggen M, Rösler F, Jost K. Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. *Psychophysiology*. 1999; 36(03):307–324. [PubMed: 10352554]
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*. 1971; 9(1):97–113. [PubMed: 5146491]
- Payne BR, Stites MC, Federmeier KD. Out of the corner of my eye: Foveal semantic load modulates parafoveal processing in reading. *Journal of Experimental Psychology: Human Perception and Performance*. 2016; 42(11):1839–1857. [PubMed: 27428778]
- Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*. 2012; 62(2):816–847. [PubMed: 22584224]
- Rosselli M, Ardila A. Calculation deficits in patients with right and left hemisphere damage. *Neuropsychologia*. 1989; 27(5):607–617. [PubMed: 2739887]
- Salillas E, Wicha NY. Early learning shapes the memory networks for arithmetic evidence from brain potentials in bilinguals. *Psychological Science*. 2012; 23(7):745–755. [PubMed: 22707225]
- Sergent J, Hellige JB. Role of input factors in visual-field asymmetries. *Brain and Cognition*. 1986; 5(2):174–199. [PubMed: 3964471]
- Sutton S, Braren M, Zubin J, John ER. Evoked-potential correlates of stimulus uncertainty. *Science*. 1965; 150(3700):1187–1188. [PubMed: 5852977]
- Van Petten C, Luka BJ. Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology*. 2012; 83(2):176–190. [PubMed: 22019481]
- Wada J, Rasmussen T. Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance: Experimental and clinical observations. *Journal of Neurosurgery*. 1960; 17(2):266–282.
- Wlotko EW, Federmeier KD. Finding the right word: Hemispheric asymmetries in the use of sentence context information. *Neuropsychologia*. 2007; 45(13):3001–3014. [PubMed: 17659309]

- Both cerebral hemispheres are sensitive to correctness of arithmetic answers
- Right hemisphere has additional late positive response to incorrect answers
- Response to correct answers likely to be P300
- No evidence for arithmetic N400 effects with Arabic numeral stimuli

Figure 1. Lateral Occipital Channels

Displayed are the recordings from left and right lateral occipital channels (LLOC at left and RLOC at right) following presentation to the right visual field (biased to the left hemisphere, plotted with solid black), and presentation to the left visual field (biased to the right hemisphere, plotted with dashed black). The N1 visual evoked potential is larger for the scalp channels corresponding to the stimulated hemisphere. There is also a later, slow-wave effect (starting here around 400 ms post-stimulus) that is characteristic of visual field manipulations.

Figure 2. Response to Answers in Central Presentation

Responses to centrally presented critical answer types from selected channels over the head (starred on the channel configuration plot at left) are plotted. The P300 elicited by correct (red) answers, as compared to incorrect related (solid black) and incorrect unrelated (dashed black), can be seen in these centromedial and posterior channels.

Figure 3. Response to Answers in Each Visual Field

The ERPs to critical answer types after presentation to RVF/LH (left plots) or LVF/RH (right plots) from selected channels over the head (same as in Figure 2) are plotted. The P300 elicited by correct (red) answers, as compared to incorrect related (solid black) and incorrect unrelated (dashed black) can be seen in both LVF/RH and RVF/LH, whereas the later effect of relatedness (difference between two black lines) can only be seen in the LVF/RH plots.

Figure 4. Distribution of Effects Following Central Presentation

Topographical maps of the effects of correctness (left) and relatedness (right), as measured by the subtraction of the response to incorrect unrelated answers from correct answers during the P300 timewindow (measured from 250 – 400 ms post stimulus onset), and subtraction of the response to incorrect related answers from incorrect unrelated answers during the later timewindow (measured 400 – 600 ms post stimulus onset), as viewed over the back of the head. Darker (red) areas correspond to more substantial effects. It can be seen that there is an effect of correctness in the P300 timewindow and no effect of relatedness in the later timewindow.

Figure 5. Distribution of Effects Following Left and Right Visual Field Presentation

Topographical maps of the effect of correctness (plots on left) and relatedness (plots on right), as measured by the subtraction of the response to incorrect unrelated answers from correct answers during the P300 timewindow (measured from 250 – 400 ms post stimulus onset), and the response to incorrect related answers from incorrect unrelated answers during the later timewindow (measured 400 – 600 ms post stimulus onset), viewed over the back of the head. It can be seen that there is an effect of correctness in the P300 timewindow for both RVF/LH (bottom left plot) and LVF/RH (top left plot), but an effect of relatedness only for responses following LVF/RH presentation (top right plot). Darker (red) areas correspond to more substantial effects. (Note that the scale used across the maps differs in order to accommodate different effect ranges, so readers should be careful not to interpret similar colors as an indication of similar magnitude of effects across these plots.)