Disentangling Neural Sources of Problem Size and Interference Effects in Multiplication

Kerensa Tiberghien[1], Muhammet I. Sahan[1], Bert De Smedt[2],
Wim Fias[1], and Ian M. Lyons[3]

Abstract

■ Multiplication is thought to be primarily solved via direct retrieval from memory. Two of the main factors known to influence the retrieval of multiplication facts are problem size and interference. Because these factors are often intertwined, we sought to investigate the unique influences of problem size and interference on both performance and neural responses during multiplication fact retrieval in healthy adults. Behavioral results showed that both problem size and interference explained separate unique portions of RT variance, but with significantly stronger contribution from problem size, which contrasts with previous work in children. Whole-brain fMRI results relying on a paradigm that isolated multiplication fact retrieval from response selection showed highly overlapping brain areas parametrically modulated by both problem size and interference in a large network of frontal, parietal, and subcortical brain areas. Subsequent analysis within these regions revealed problem size to be the stronger and more consistent "unique" modulating factor in overlapping regions as well as those that appeared to respond only to problem size or interference at the whole-brain level, thus underscoring the need to look beyond anatomical overlap using arbitrary thresholds. Additional unique contributions of interference (beyond problem size) were identified in right angular gyrus and subcortical regions associated with procedural processing. Together, our results suggest that problem size, relative to interference, tends to be the more dominant factor in driving behavioral and neural responses during multiplication fact retrieval in adults. Nevertheless, unique contributions of both factors demonstrate the importance of considering the overlapping and unique contributions of each in explaining the cognitive and neural bases of mental multiplication. ■

INTRODUCTION

Arithmetic is a quintessential mathematical ability, serving in many ways as a cornerstone of everyday numeracy skills (Ashcraft & Guillaume, 2009; Campbell, 2005). Memory retrieval has long been understood to play a key role in arithmetic processing and, in particular, in retrieving multiplication facts (e.g., Fayol & Thevenot, 2012; Campbell & Xue, 2001; Kirk & Ashcraft, 2001). Two of the main factors thought to influence multiplication memory retrieval are problem size (e.g., Stazyk, Ashcraft, & Hamann, 1982) and interference (De Visscher, Berens, Keidel, Noël, & Bird, 2015; De Visscher & Noël, 2014a, 2014b), both on a behavioral and neural level. However, these two factors are often intertwined, and as such, the unique influences of both problem size and interference on memory representation remain poorly understood, particularly at the neural level.

Among the factors thought to influence multiplication memory retrieval, the problem size effect (PSE) has perhaps received the greatest interest. The PSE is a robust finding, wherein poorer performance (slower and more error prone) is typically observed for larger problems relative to smaller problems. Different explanations of the PSE have been suggested. A leading explanation for the PSE suggests that the frequency with which arithmetical problems are taught in school impacts how a given item is stored or represented in memory (Ashcraft & Christy, 1995; McCloskey & Lindemann, 1992; Ashcraft, 1987). Large problems are encountered less frequently and are therefore stored in memory with lower strength, resulting in poorer performance relative to small problems. Similarly, Siegler and Shrager (1984) suggested that learned problems are associated with correct and incorrect answers. Small problems have a lesser history of error, leading to only weak associations with erroneous answers, in turn leading to a high likelihood of retrieving the correct answer. On the other hand, large problems have a larger history of errors because of the possibility of making errors during the execution of calculation procedures (e.g., 6 × 7 = 7 + 7 + 7 + 7 + 7 + 7) at early stages of learning. Consequently, the strength of the association from a large problem to an incorrect answer is relatively high, compared with the strength of small problems. Another explanation for the PSE is put forward by Verguts and Fias's (2005) model of interacting neighbors. Within this model, the PSE can be understood from the

---

[1]Ghent University, [2]University of Leuven, [3]Georgetown University

Journal of Cognitive Neuroscience 31:3, pp. 453–467
doi:10.1162/jocn_a_01359

principles of cooperation and competition. When presenting a problem in an associative network, neighboring problems can either cooperate (if they lead to the same response, because they share the decade or unit; e.g., $4 \times 6 = 24$ and $4 \times 7 = 28$) or compete (if they lead to a different response, because they do not share the decade or unit; e.g., $4 \times 4 = 16$ and $4 \times 5 = 20$). When neighboring problems lead to the same response, they are consistent and hence facilitate the retrieval of the correct answer, whereas inconsistent neighbors will compete with one another and therefore interfere with the retrieval of the correct answer.

Recently, another memory effect has been described that modulates performance of multiplication facts—namely, the interference effect. The multiplication interference effect was first described by De Visscher and Noël (2014a) and is based on feature overlap theory (Oberauer & Lange, 2008; Nairne, 1990), which proposes interference arises directly as a function of the number of overlapping features between two problems. Thus, when more recently learned problems have to be retrieved, they will be subject to proactive interference as a function of the number of overlapping features with previously learned problems. Because multiplication tables tend to be learned sequentially from the two times table to the three times table, up to the nine times table, and so on, it has been suggested that multiplication tables are subject to proactive interference during learning. This interference results in a gradual decrease of memory encoding weight as operand quantities increase. To test this idea, De Visscher and Noël (2014a) developed a measure of the overlap of the proactive interference between multiplication facts, which they termed an "interference parameter." The interference parameter represents the degree of proactive interference by calculating the digit overlap between a given problem (including the solution) with previously learned problems. In this way, each problem is assigned "proactive interference points." One point is given when there is an overlap of two digits (e.g., $\underline{2} \times \underline{8} = 16$ and $\underline{2} \times 4 = \underline{8}$), whereas 3 points are given when three digits overlap (e.g., $\underline{3} \times 9 = \underline{27}$ and $\underline{3} \times \underline{7} = \underline{2}1$). For example, the interference parameter of $3 \times 6 = 18$ is 8 because the proactive interference comes from six previously learned problems ($2 \times \underline{3} = \underline{6}$, $2 \times \underline{6} = \underline{1}2$, $2 \times \underline{8} = \underline{16}$, $\underline{3} \times 2 = \underline{6}$, $\underline{3} \times 4 = \underline{1}2$, $\underline{3} \times 5 = \underline{1}5$). The study showed that problems with a higher interference parameter yielded longer RTs (see also De Visscher & Noël, 2014b).

It is important to point out that problem size and proactive interference have similar predicted effects on performance: Larger problem size should lead to poorer performance, and larger problems involve higher times tables and so, because of increased proactive interference, should also lead to poorer performance. This means that problem size and interference effects are strongly correlated (indeed, $r = .55$ for operands 2 through 9; De Visscher & Noël, 2014a, 2014b). In fact, the network model of arithmetic proposed by Campbell (1995) explicitly posits interference as a key mechanism to explain the PSE. Specifically, large problems are associated with a wider range of other problems (relative to smaller problems), which induces a higher degree of retrieval interference (and hence poorer performance on large problems). An important question is thus whether interference and PSEs are essentially different descriptors of the same underlying mechanism or whether the two factors might be separated into distinct—that is, unique—contributions with respect to arithmetic processing.

To this end, De Visscher, Noël, and De Smedt (2016) directly compared the predictive capacity of problem size and interference parameters on typed multiplication production RTs in fourth-grade children. At the individual subject level, unique effects of both parameters were found (albeit somewhat stronger for interference: $t = 4.4$ vs. 2.5); at the sample level (averaging RTs across all children for a given trial), the authors found that only the interference parameter captured significant unique variance. Together, the De Visscher et al. results suggest that interference is the stronger determinant of multiplication fact retrieval performance in children.

A few questions remain, however. First, De Visscher et al. (2016) examined children still acquiring basic arithmetic knowledge. One question is thus whether interference or problem size is the stronger predictor of multiplication performance in adults whose fact retrieval is more likely to be heavily practiced and thus more efficient. Second, De Visscher et al. used a typing production task (children typed their answers); because the interference parameter is directly related to the number of overlapping digits, this may have inflated the relation with multiplication RTs as answers requiring the same digits to be typed would likely yield similar RTs. Third, it would be potentially informative to examine the relative contributions of problem size and interference to predicting neural responses during retrieval of multiplication facts. Moreover, by focusing on just the retrieval phase, fMRI has the potential to isolate computation from response selection, thus mitigating the potential influence of extraneous aspects of the task not of immediate theoretical interest.

In an fMRI study, De Visscher et al. (2015) assessed neural responses while participants verified whether multiplication problems had been correctly solved. To isolate problem size and interference, the authors chose a subset of problems to fill out an orthogonal 2 (problem size: small, large) × 2 (interference: high, low) design. Results showed that most regions sensitive to one effect were also sensitive to the other (two main effects); indeed, no interactions were found at the whole-brain level. The authors did find that the left angular gyrus was sensitive to interference but not problem size and the opposite pattern in the right intraparietal sulcus (IPS; see also De Visscher et al., 2018).

These two fMRI studies (De Visscher et al., 2015, 2018) converge to show broadly overlapping effects for problem size and interference at the neural level, with some evidence indicating the two effects may be dissociable as well. However, because of the ANOVA-based design, neither study was able to quantify the relative degree of unique and overlapping contribution of each parameter to modulation of neural responses. Only orthogonal components could be identified, and this was specific to a highly selected subset of problems, thus obfuscating the natural correlation between the two parameters when considering multiplication processing more generally. Instead, a parametric fMRI design similar to that used in assessing the relative contributions of the two parameters to behavioral responses (e.g., De Visscher et al., 2016) might be more optimal for this purpose. Furthermore, with respect to both studies (De Visscher et al., 2015, 2018), two additional issues are worth noting: (1) Both used a verification procedure wherein the presence of the proposed solution can alter the nature of retrieval processing (e.g., by a priori narrowing the retrieval search space), and (2) both studies modeled activity across retrieval and response selection, which introduces confounds from the presence and need to process the verification stimulus and the selection of a specific motoric response. fMRI provides the opportunity to separate computation from response selection, which may be particularly crucial in the current case given that interference and problem size may interact differently with response selection than with retrieval processing.

In this study, we used a parametric approach to disentangle the overlapping and unique influences of problem size and interference on behavioral performance and neural response patterns in healthy adults. Specifically, we used a parametric univariate approach for the neural data, wherein we identified brain regions (via whole-brain analysis) whose responses were systematically modulated by interference, problem size, or both. Within these regions, we then assessed the extent to which each factor—interference or problem size—uniquely explained the whole-brain result. We did so via a univariate multiple regression approach (similar to that used for the behavioral data). In this way, we provide a systematic assessment of the relative unique contributions of interference and problem size to multiplication fact retrieval in adults at both the behavioral and neural levels. Moreover, the fMRI paradigm we used was designed to isolate multiplication fact retrieval from response selection, and the behavioral paradigm relied on verbal responses rather than typed responses; together, these two methodological aspects of our approach allowed us to reduce the likelihood that response-related factors may have driven our results. In summary, our aim was to quantify the respective roles of problem size and interference on multiplication retrieval via behavioral and neural measures in healthy adults with mature mental arithmetic systems.

METHODS

Participants

Thirty adults from Ghent University participated in the experiment (22 women, mean age = 24 years, range = 18–27 years, all right-handed). All participants had normal or corrected-to-normal vision and reported no history of neurological or psychiatric illness. Before taking part in the study, participants gave written consent. All participants were paid €40 for their participation. The study was approved by the medical ethics committee of Ghent University and Ghent University Hospital. Six participants were excluded from further analyses (four because of excessive movement, one because of technical difficulties, one was diagnosed with dyscalculia), leaving a final *n* of 24 participants.

Procedure

All experiments were presented via E-Prime (Psychology Software Tools) and displayed on a 1600 × 900 resolution screen. Participants performed an arithmetic task both before and during scanning. For the behavioral task, the computer was placed on average 50 cm in front of the participant. In the scanner, stimuli were presented via a Brainlogics 200MR digital projector visible via a mirror attached to the head coil, with a viewing distance of 120 cm.

Tasks

We should note that the data reported here are part of a larger data set; all results reported here are unique and address hypotheses that do not overlap with any other current or future publications arising from this data set. In particular, both the prescan and fMRI arithmetic tasks included three operations: multiplication, addition, and subtraction. Operation order was fully randomized across participants, so the presence of subtraction problems should not yield systematic biases when considering just the multiplication and addition problems. Here, we limit our attention to just multiplication problems, as these are directly relevant to the hypotheses of theoretical interest here (i.e., memory retrieval). Because, to date, prior work on interference effects have been examined almost exclusively in multiplication, we were primarily interested in investigating the unique contribution of the PSE and interference effect on multiplication.

Prescan Arithmetic Task

The arithmetic task was a production task (i.e., task where the participant needed to generate the answer) containing all permutations of two operands ranging from 0 to 10 (121 total problems) with three different arithmetic operations: addition, multiplication, and subtraction (resulting in a grand total of 363 problems). All problems were presented once, with order randomized across participants. In keeping with prior work (e.g., De Visscher et al., 2016), we focused on multiplication problems from 2 × 2 to 9 × 9, though notably, here we have 64 trials per subject in this range (as opposed to 36 in De Visscher et al.), which should allow for more accurate assessment of problem size and interference effects at the individual subject level. Note that the multiplication tables 0, 1, and 10 were not included. The reason is that these problems are probably solved by means of rule-based strategies (e.g., Sokol, McCloskey, Cohen, & Aliminosa, 1991), and hence, the problem size and interference effects may not necessarily apply for these multiplication problems.

A trial started with a fixation (three squares) presented for 3000 msec followed by the arithmetic problem. The problem remained on the screen until the participant responded. Once the participant said the response out loud, a voice key recorded the onset of speech. Next, the experimenter recorded the participant's response and noted if the voice key triggered correctly. For instance, a participant might occasionally extend a response over a long duration, or he or she may have unwittingly made an extraneous sound such as "uhm" or cough that accidentally triggered the voice key. On such trials, the data were discarded, and that same problem was presented again at a randomly chosen time later in the prescan task. The intertrial interval was 1000 msec. This procedure was maintained until a valid response was recorded for all 363 problems. Short breaks were given after every 33 trials.

Arithmetic Task—fMRI Version

The arithmetic task inside the scanner was kept as similar to the prescan version as possible: All problems ranging from 0 to 10 were used in addition, multiplication, and subtraction, resulting in a grand total of 363 problems. Trials were divided as evenly as possible across six separate runs, with trial (and hence also run) order randomized across participants. As in the prescan task, a trial started with a fixation presented for 3000 msec, followed by the first arithmetic problem. Here, the problem (e.g., 9 × 8) remained on the screen for 2600 msec. Participants were instructed to mentally compute the answer during this period. For most trials, the problem was then replaced by fixation. For 10% of trials, a response was required, in which case the problem was replaced by two response possibilities for 1500 msec. One number was the correct response, whereas the other number was the correct response ±1. Participants pressed either a left or a right key (left or right index finger) to indicate the correct answer. Intertrial interval (fixation) was jittered (range = 1000–8194 msec, M = 3421 msec) for all trials. Response events were modeled as events of no interest.

It is important to note that, in the scanner, the focus was less on RTs and more on participants' mental calculation of the solution. That is, our goal was to separate the mental calculation aspects of arithmetic processing from response preparation and execution. Therefore, participants were instructed to mentally compute the answer to each problem while it remained on the screen, without an overt response. Responses were therefore probed on only 10% of all problems. These response probes were randomly assigned to different problems (thus, this randomization was different for each participant). Hence, participants could not predict in advance which problems would require a response, and the association between a response event and a given problem was thus not systematic across participants. Accuracy for problems of the scanner task was high (M = 93.9%, SE = 1.5%) and highly similar to problems outside the scanner (prescan task; M = 93.9%, SE = 0.8%). Average RTs were faster than those seen for the prescan behavioral task (prescan: M = 1150 msec, SE = 61 msec; in-scan: M = 681 msec, SE = 19 msec), which is what one would expect if participants were computing the answer during presentation of the problem (2600 msec) before appearance of the (occasional) verification probe. Further evidence that participants were engaging with the task as instructed is the presence of problem size and interference effects in the neural data (see Results). That is, fMRI analyses focused exclusively on the calculation period (before response). It is difficult to explain how a problem size or interference effect could have been observed during this period if participants were simply ignoring problem presentation and waiting until the sporadic presentation of response options to engage with the task.

fMRI Data Acquisition and Preprocessing

Images were collected with a 3-T Siemens Magnetom Trio MRI system (Siemens Medical Systems) using a 32-channel radio-frequency head coil. Participants were positioned headfirst and supine in the magnet bore. Participants were instructed to move their heads as little as possible throughout the entire scanning session. A whole-brain high-resolution anatomical scan was acquired using a standard 3-D MPRAGE sequence (voxel size = 1 mm^3). Functional images were collected using an EPI sequence: repetition time = 2600 msec, echo time = 28 msec, flip angle = 80°; in-plane matrix of 3.3 mm^2 voxels = 64 × 64 (field of view = 211 mm), with slice thickness = 3.3 mm (44 slices, interleaved, no skip), yielded 3.3 mm^3 isometric voxels.

Structural and functional images were analyzed using Brain Voyager QX 20.4 (Brain Innovation). Functional data were interpolated to 3 mm^3 and corrected for slice scan timing using cubic spline interpolation, corrected for head motion (trilinear/sinc interpolation), and finally high-pass filtered using a general linear model (GLM) procedure with a Fourier basis set. Excessive motion was deemed net drift >3 mm in a given run or >1.5 mm sudden movement; participants with runs exceeding these criteria were removed from analysis ($n = 4$). Participants' functional images were then coregistered to their respective anatomical scans using 12-parameter gradient-based affine alignment, and anatomical images were coregistered into Talairach space (Talairach & Tournoux, 1988). Functional data were spatially smoothed at 3 mm FWHM.

Analysis Approach

Problem Size

The problem size can be measured with several indices (i.e., minimum operand, maximum operand, sum, sum squared, product). However, the product of the two operands is more frequently used (Campbell, 1997). For that reason and because the product was a better predictor of performance in the study by De Visscher and Noël (2014a, 2014b), we have used the product as measure of the problem size (see Appendix A). An important difference is that our study has included the commutative pairs (e.g., 2 × 4, 4 × 2), whereas the study of De Visscher and Noël (2014a, 2014b) and De Visscher et al. (2015) did not include them, which increased the available number of trials per participant from 36 to 64, thereby increasing the precision with which parametric effects of problem size and interference could be estimated for each participant.

Interference

As noted in the introduction, the interference parameter is based on the feature overlap theory (Oberauer & Lange, 2008; Nairne, 1990). The idea is that problems that are similar to previously encoded problems will be recalled more poorly due to proactive interference. Hence, similarity can be quantified by the feature overlap between problems. Multiplication tables also contain proactive interference during learning, because these are learned from the two times table up to the nine times table. De Visscher and Noël (2014a, 2014b) calculated the digit overlap (i.e., the proactive interference) between a problem (including its solution) and the previously learned problems, resulting in an interference parameter. The interference parameter is based on "proactive interference points." One point is given when there is an overlap of two digits (e.g., $\underline{2} \times \underline{8} = 16$ and $\underline{2} \times 4 = \underline{8}$), and 3 points are given when three digits overlap (e.g., $\underline{3} \times 9 = \underline{27}$ and $\underline{3} \times \underline{7} = \underline{2}1$). For example, the problem "3 × 6 = 18" has an interference parameter of 8, because of proactive interference from six previously learned problems (1 point for $2 \times \underline{3} = \underline{6}$, $2 \times \underline{6} = \underline{1}2$, $\underline{3} \times 2 = \underline{6}$, $\underline{3} \times 4 = \underline{1}2$, $\underline{3} \times 5 = \underline{1}5$ and 3 points for $2 \times \underline{8} = \underline{16}$). This was done for every multiplication from 2 × 2 to 9 × 9, resulting in an interference parameter that ranged from 0 to 25 (see Appendix A).

Behavioral Analyses

Accuracy levels were high for both the prescan and in-scanner tasks (94% in both cases). Such near-ceiling performance levels provide minimal variability for model fitting; hence, analyses here focus instead on RTs. The two parameters (problem size and interference) were used to predict RTs, separately for each participant. Each problem ranging from 2 × 2 to 9 × 9 was assigned a problem size (based on the product of the operands) and an interference parameter (see above and Appendix A). Parameter fits were estimated in two ways: each predictor treated separately (zero-order relations) and simultaneously (unique relations). Zero-order and unique r values for each predictor were extracted for each participant, which quantified the degree of unique fit between a given parameter and RT; tests were then conducted on Fisher z transformed ($z = \text{atanh}(r)$) r values across participants.

fMRI Analyses

The same parametric approach was used for fMRI analysis. Predictors comprised trials in the same manner as the behavioral results (other trials and response events were modeled as events of no interest) to parallel the behavioral analyses as closely as possible. Note that parametric predictors were scaled from 0 to 1 to make resulting beta-weights more interpretable. A standard voxel-wise GLM was run with parametric predictors for problem size and interference, in addition to the main effect of multiplication (i.e., a single predictor with all multiplication trials weighted equally). Because of the high degree of collinearity between the parametric predictors, if run together in the same GLM, the beta weights associated with problem size and interference would each carry very large error estimates, making their interpretation problematic. Thus, at the whole-brain level, we ran two separate GLMs: each containing just one of the parametric predictors. Note that this also allowed us, initially at least, to identify regions sensitive to each effect separately.[1] For each GLM, ROIs were identified via a conjunction with the overall effect of multiplication: problem size: (Multiplication ∩ Problem Size), interference: (Multiplication ∩ Interference). A given region was thus responsive to multiplication processing in general, and it was modulated systematically as a function of problem size, for instance. Resulting statistical maps were thresholded using an uncorrected voxel-wise threshold of $p < .001$[2] and subsequently cluster-corrected for multiple comparisons using a Monte Carlo simulation procedure (Forman et al., 1995) at $\alpha < .01$.

Within each ROI thus identified, we then partialed out unique modulation of brain activity associated with each parameter, problem size and interference. This was done by creating residualized versions of each parametric predictor, in that each was residualized with respect to its

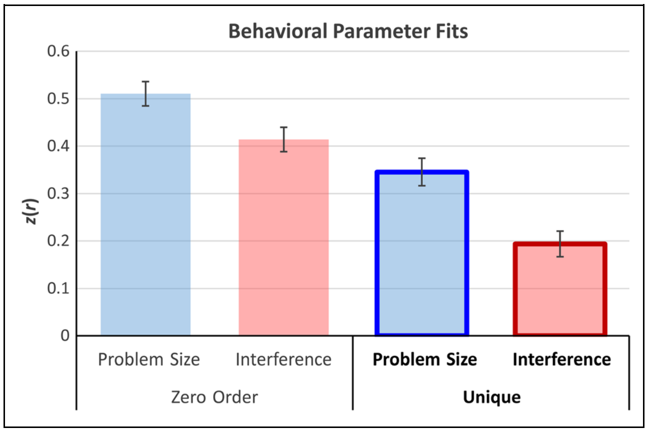

Figure 1. Behavioral results. The figure shows zero-order (left) and unique (right) predictive capacities of problem size and interference (for RTs). Values are Fisher z-transformed r values, which were computed for each participant separately, then averaged across participants. Error bars are *SEM*s.

counterpart. In this way, the residualized problem size predictor no longer contained variance associated with interference, and vice versa. We then computed betas associated with each of these residualized predictors (controlling also for the main effect of multiplication, as with the whole-brain analyses) separately for each participant in each ROI. Finally, these betas were then contrasted against 0 using a standard t test. In this way, within ROIs that responded to problem size, interference, or both, we then computed the extent to which this effect was associated uniquely with each factor.

RESULTS

Behavioral Results

Behavioral results are summarized in Figure 1. The left side of Figure 1 shows zero-order fits; the right side shows unique fits (the fit between problem size and RTs controlling for the interference, and vice versa). Zero-order fits are provided mainly for context. Our primary theoretical interest here is with the unique predictive capacity of each parameter. Both problem size and interference uniquely predicted longer RTs in that the average (z transformed) unique correlation value across participants was well above zero (problem size: M = .346, SE = .029, t(23) = 12.00, p = 2.2E−11; interference: M = .194, SE = .027, t(23) = 7.15, p = 2.7E−07; see also Figure 1, right). Crucially, problem size captured significantly more unique variance relative to interference (p = 3.7E−04).

To summarize, results show that both problem size and interference predicted performance and that each contributed unique variance, although problem size contributed significantly greater unique variance. In the next section, we examine the overlapping and unique contributions of these factors insofar as each predicts neural activity during mental multiplication fact retrieval, independent of response demands.

fMRI Results

Whole-brain Zero-order Effects

Whole-brain univariate results are shown in Figure 2, with anatomical details (and region abbreviations) given in Table 1. Consistent with the notion that problem size and interference are highly intertwined, the majority of areas sensitive to one parameter were also sensitive to the other (overlap is shown in yellow in Figure 2; see also the "Overlap" section in Table 1). Overlapping voxels included bilateral IPS and right angular gyrus (RAG), and they comprised roughly two thirds (65.2%) of all active voxels across both contrasts, indicating substantial comodulation of neural responses by the two parameters. Several regions showed significant modulation for

Figure 2. Regions showing modulation by problem size and interference. The figure depicts regions significantly modulated by problem size and/or interference at the whole-brain level. Note that the effects shown here are zero-order relations, in that effects of problem size do not take into account interference, and vice versa. For unique effects, see Figure 3. Regions in blue are regions active for problem size, regions in red are active for interference, and regions in yellow comprise overlapping voxels for both problem size and interference. See Table 1 for complete region details and abbreviations.

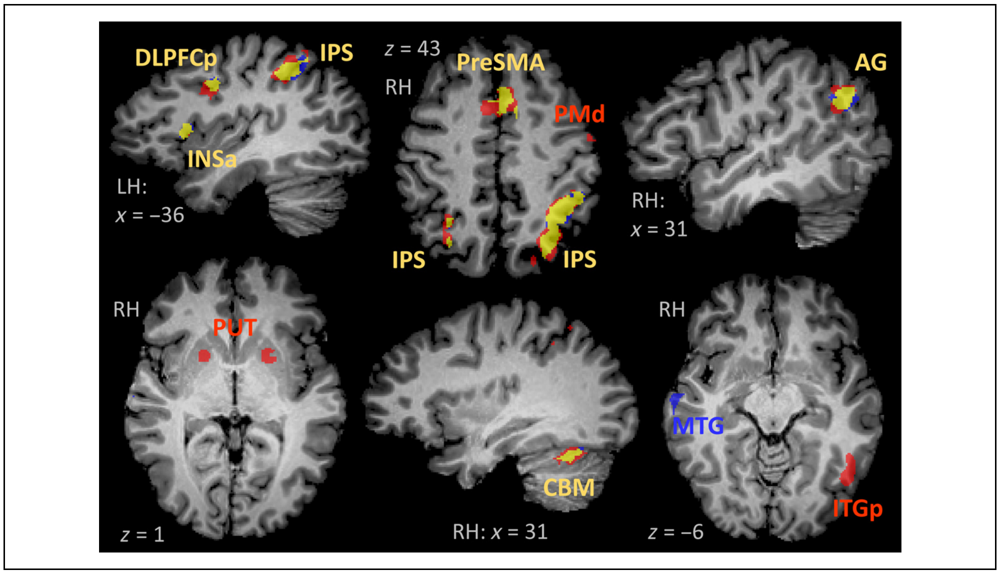

Table 1. Region Details

	Overlap				*Problem Size*				*Interference*			
Region	*x*	*y*	*z*	*Volume*	*x*	*y*	*z*	*Volume*	*x*	*y*	*z*	*Volume*
LDLPFCp	−39	2	32	391	−40	2	32	453	−40	5	31	1085
Pre-SMA	−1	11	47	1292	−1	11	47	1324	−1	10	46	3142
LINSa	−32	18	8	829	−32	18	8	849	−31	18	9	1078
LIPS	−30	−52	42	3608	−31	−52	42	3967	−30	−53	42	6343
RIPS	27	−56	44	240	27	−57	43	299	27	−55	44	899
RAG	47	−56	27	669	47	−56	27	871	47	−56	27	966
RCBM	30	−55	−21	316	30	−56	−21	367	28	−56	−22	865
LACCd[a]					−7	20	34	220				
RMTG[a]					57	−13	−6	344				
LMFG[b]									−27	−1	54	477
LPMd[b]									−47	−5	47	231
LITGp[b]									−45	−57	−5	685
LPUT[b]									−17	7	5	759
RPUT[b]									18	8	4	420

"Overlap" indicates overlapping voxels for a given ROI, where applicable. Volume is measured in mm^3. LDPFCp = left posterior dorsolateral pFC; LMFG = left middle frontal gyrus; LPMd = left dorsal premotor; LINSa = left anterior insula; LACCd = left dorsal ACC; LITGp = left posterior inferior temporal gyrus; RMTG = right middle temporal gyrus; LIPS = left IPS; RIPS = right IPS; LPUT = left putamen; RPUT = right putamen; RCBM = right cerebellum.

[a] ROI showing a significant result only for problem size at the whole-brain level.

[b] ROI showing a significant result only for interference at the whole-brain level.

only one parameter, though it is possible that the other parameter also modulated neural responses, simply just below threshold. Note also that these results effectively constitute zero-order relations between each parameter (problem size, interference) and neural activity. In the next section, we quantify the unique predictive capacity of each parameter within each of the regions identified here.

Unique Effects

Zero-order effects in the previous section indicated substantial overlap in terms of brain areas modulated by both problem size and interference. It is thus important to assess the unique contributions of each parameter within these regions to address the question of whether these contributions are unique or effectively the same. Moreover, although several regions show significant modulation by just one parameter at the whole-brain level, it is important to address the concern that the other parameter modulated activity just below threshold, and hence whether, once one accounts for this potential subthreshold modulation, the original effect still obtains. We address both issues in this section. Note that, although some regions were overlapping (Table 1), there were nevertheless subtle differences in the voxels active for each parameter around the edges of these ROIs. To respect the possibility that these subtle differences may nevertheless prove key, we analyzed regions from the problem size and interference whole-brain contrasts (Figure 2; Table 1) separately. Results for problem size regions can be seen in Figure 3A; results for interference regions can be seen in Figure 3B.

Note that Figure 3 gives both zero-order effects and partial (unique) effects. The former are shown in lighter shade and labeled "zero order" in the figure legends. These effects are similar to zero-order correlations between the parametric predictor and neural activity in that they do not control for the influence of the other parameter. As these are the effects identified by the whole-brain analysis, interpreting them in and of themselves would essentially constitute double dipping. We provide them here (1) as visual confirmation of the whole-brain results, but more importantly (2) to provide a frame of reference for the main focus of Figure 3: the unique effects of each parameter. The unique effects are shown in slightly darker shade with bold outline (labeled as "unique" in the figure legends). These effects constitute the fit between a given parameter and neural activity, controlling for (i.e., over and above) the influence of the other parameter. As expected, the unique effects tend to be smaller than the zero-order effects; the key,

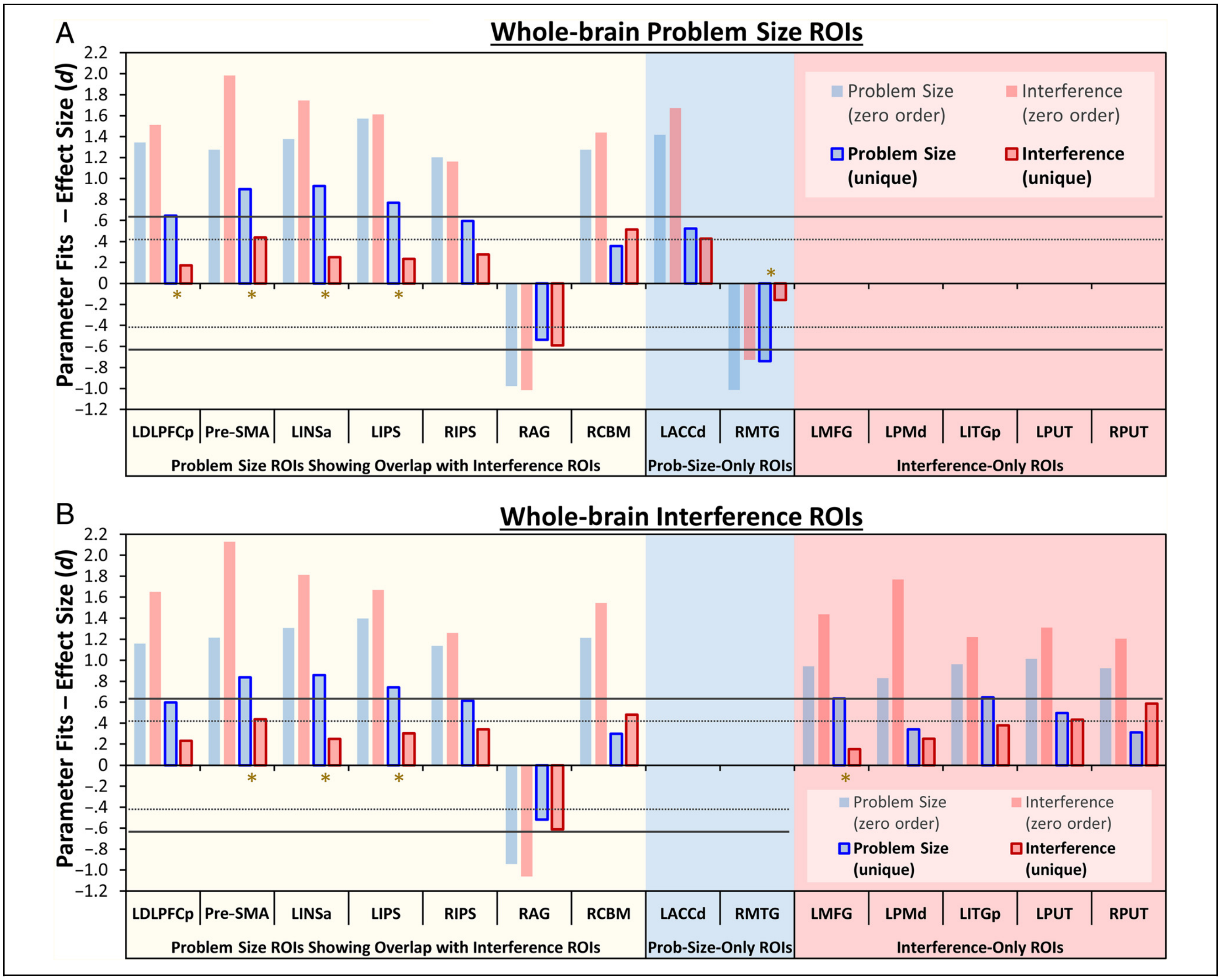

Figure 3. Unique neural effects of problem size and interference. This figure shows predictive effects of neural activity for problem size and interference in (A) problem size and (B) interference regions from the whole-brain analyses (Figure 2; Table 1). Effects are reported here as effect sizes; means and *SE*s are given in Table 2. Lighter shaded bars correspond to zero-order effects (the same as those used to identify the regions in the whole-brain analyses). They are provided here to give visual context to the unique effects, which are shown in darker bars in bold outline, which are the primary focus of this analysis. Regions in the yellow "Overlap" subpane showed overlapping whole-brain effects. Regions from the separate contrasts were treated separately, though they are aligned here across subfigures (A) and (B) to aid in visual comparison. The dotted gray line corresponds to the significance threshold $p < .05$; the solid gray line corresponds to the more conservative threshold of $p < .005$ (see text for details). See Table 1 for region abbreviations.

though, is in assessing which parameter—problem size, interference, both, or neither—retains significant unique predictive capacity of neural responses. Finally, note that effect sizes are given in Figure 3 to aid generalizability across studies (raw means and standard errors can be found in Table 2). Two statistical thresholds are given: the traditional $p < .05$ (dotted gray line) and the more conservative $p < .005$ (solid gray line). The latter roughly corresponds to Dunn–Šidák correction for multiple comparisons: $p = .0057$ in Table 2A (nine regions) and $p = .0043$ in Table 2B (12 regions).

We first address overlap regions (yellow portions of Figure 3). We found significant unique modulation of neural activity by problem size in six of the seven overlap regions: LDLPFCp, pre-SMA, left anterior insula, left IPS, right IPS, and RAG. Indeed, most problem size results were highly significant (obtaining at the more stringent significance threshold), and problem size contribution was significantly greater than that of interference in four of the seven overlap regions (denoted by a gold asterisk in Figure 3). Results were overall similar whether one looked at overlap regions defined by the problem size (Figure 3A) or interference (Figure 3B) whole-brain contrasts. In contrast, we found significant unique modulation of neural activity by interference in only three of the seven overlap regions: pre-SMA, RAG, and right cerebellum, each only at the more liberal threshold of $p < .05$. This suggests that the whole-brain overlap effects were largely driven by problem size. That said, as with the behavioral results, additional unique variance

Table 2. Means and Standard Errors of Neural Effects

(A) Whole-brain Problem Size ROIs

	ROI	*Problem Size (Zero-order)*	*Interference (Zero-order)*	*Problem Size (Unique)*	*Interference (Unique)*
Overlap	LDLPFCp	.87 (.13)	.73 (.10)	.90 (.28)	.17 (.21)
	Pre-SMA	.98 (.16)	.88 (.09)	.89 (.20)	.30 (.14)
	LINSa	.89 (.13)	.75 (.09)	.86 (.19)	.17 (.14)
	LIPS	.87 (.11)	.75 (.10)	.89 (.24)	.20 (.18)
	RIPS	.66 (.11)	.61 (.11)	.55 (.19)	.24 (.18)
	RAG	−.81 (.17)	−.55 (.11)	−.56 (.21)	−.39 (.14)
	RCBM	.70 (.11)	.66 (.09)	.57 (.33)	.47 (.19)
Problem size	LACCd	.73 (.10)	.81 (.10)	.62 (.24)	.38 (.18)
	RMTG	−.69 (.14)	−.37 (.10)	−.69 (.19)	−.09 (.11)
Interference	LMFG				
	LPMd				
	LITGp				
	LPUT				
	RPUT				

(B) Whole-brain Interference ROIs

		Problem Size (Zero-order)	*Interference (Zero-order)*	*Problem Size (Unique)*	*Interference (Unique)*
Overlap	LDLPFCp	.82 (.14)	.71 (.09)	.80 (.27)	.21 (.18)
	Pre-SMA	.84 (.14)	.78 (.07)	.77 (.19)	.26 (.12)
	LINSa	.86 (.13)	.75 (.08)	.86 (.20)	.16 (.13)
	LIPS	.80 (.12)	.74 (.09)	.79 (.22)	.24 (.16)
	RIPS	.61 (.11)	.59 (.10)	.46 (.15)	.28 (.17)
	RAG	−.80 (.17)	−.55 (.11)	−.55 (.21)	−.41 (.14)
	RCBM	.61 (.10)	.63 (.08)	.48 (.33)	.46 (.20)
Problem size	LACCd				
	RMTG				
Interference	LMFG	.62 (.13)	.65 (.09)	.70 (.22)	.12 (.16)
	LPMd	.47 (.12)	.47 (.05)	.53 (.32)	.19 (.15)
	LITGp	.60 (.13)	.61 (.10)	.56 (.18)	.27 (.15)
	LPUT	.47 (.10)	.55 (.09)	.39 (.16)	.28 (.13)
	RPUT	.49 (.11)	.52 (.09)	.29 (.19)	.42 (.15)

The table gives means (standard errors) for problem size and interference effects in each ROI, averaged across participants. Effect sizes comparing these values against 0 are given in Figure 3. LDPFCp = left posterior dorsolateral pFC; LMFG = left middle frontal gyrus; LPMd = left dorsal premotor; LINSa = left anterior insula; LACCd = left dorsal ACC; LITGp = left posterior inferior temporal gyrus; RMTG = right middle temporal gyrus; LIPS = left IPS; RIPS = right IPS; LPUT = left putamen; RPUT = right putamen; RCBM = right cerebellum.

in neural activity attributable to interference was seen in several regions; hence, it would be premature to discount interference altogether. A final point is that modulation in all regions save one was positive, indicating greater positive deflection of neural activity as problem size and/or interference increased; the RAG showed the opposite pattern.

The two "problem size" regions (blue portions of Figure 3) both showed significant effects of problem size; left dorsal ACC also showed a significant effect of interference. Note also that the right middle temporal gyrus, similar to RAG, showed negative effects, indicating significantly reduced activation as problem size increased.

Perhaps surprisingly, regions that, at the whole-brain level, appeared to show only effects of interference (red portions of Figure 3) in fact tended to show somewhat larger unique effects of problem size. Although the zero-order effects of interference tended to be larger (albeit nonsignificantly so), each of these regions also showed a zero-order effect of problem size (which was thus likely just below the critical whole-brain threshold). Crucially, taking into account these subthreshold effects revealed a somewhat different picture. Although three of five showed significant unique effects of problem size (left middle frontal gyrus, left posterior inferior temporal gyrus, left putamen), only two of the five showed unique effects of interference (bilateral putamen), though it is important to note that the difference between effects reached significance only in left middle frontal gyrus. These results underscore two points: First, it is crucial to take into account potential competing (correlated) predictors, even if these predictors obtain only at subthreshold levels at the whole-brain level. Second, these regions conform to the broader pattern whereby problem size accounts for a greater unique proportion of multiplication-related variance, with interference accounting for additional significant variance in a smaller set of regions.

DISCUSSION

Arithmetic is a quintessential mathematical ability used by many children and adults on a more or less daily basis. Mental multiplication is one of the most common forms of arithmetic, which is thought to be primarily solved via direct retrieval from memory. The key factors that govern memory retrieval in arithmetic—and in particular retrieval of multiplication facts—remain a major source of interest for researchers in the domain of numerical cognition and more broadly in the domain of memory retrieval as well. Two of the main factors thought to influence multiplication memory retrieval both behaviorally and neurally are problem size (e.g., Stazyk et al., 1982) and interference (De Visscher et al., 2015; De Visscher & Noël, 2014a, 2014b). However, because these two factors are often intertwined, here we sought to identify the unique influences of problem size and interference on both performance and neural responses during multiplication fact retrieval in healthy adults. Behavioral results showed that both problem size and interference explained separate unique portions of RT variance, but with a significantly stronger contribution from problem size. Whole-brain fMRI results using a paradigm that isolated multiplication fact retrieval from response selection showed highly overlapping brain areas parametrically modulated by both problem size and interference. Within these regions, problem size was the stronger and more consistent unique modulating factor. This result was obtained in both overlapping regions as well as those that appeared to respond only to problem size or interference at the whole-brain level, thus underscoring the need to look beyond anatomical overlap and arbitrary thresholds by accounting for unique modulatory contributions. Problem size, relative to interference, appears to be the more dominant factor in driving both behavioral and neural responses during multiplication fact retrieval in adults. That said, additional unique contributions of interference (beyond problem size) were identified in RAG and several subcortical regions associated with procedural processing. Hence, the unique contributions of both factors demonstrate the importance of considering the overlapping and unique contributions of each in explaining the cognitive and neural basis of mental multiplication.

One of the key features of the current study is that we allowed problem size and interference to covary across trials, which is what in turn allowed us to assess both overlapping and unique contributions of each in a single study. Previous fMRI work looking at problem size and interference preselected trials so as to create an orthogonal Problem Size × Interference design (De Visscher et al., 2015, 2018). Such an approach is highly useful for isolating specific effects, as was the intent of the authors of the papers in question. However, such an approach (1) a priori assumes that problem size and interference indeed comprise independent factors and (2) it thus does not allow one to explicitly test the validity of this assumption across a broader array of commonly encountered problems.

By adopting a parametric approach, we were able to do precisely this. On the one hand, our results support the validity of this assumption. The behavioral data showed highly significant unique contributions of both problem size and interference with respect to predicting multiplication RTs (Figure 1). That said, problem size captured more unique behavioral variance. That is, the average standardized fit across participants (mean partial *r*) was nearly twice as large for problem size as that seen for interference (.346 vs. .194). Interestingly, this is roughly the opposite of what De Visscher et al. (2016) found when using a similar approach: Those authors found significant unique effects of both parameters on RTs at the individual level, but somewhat stronger for interference rather than problem size (*t*s = 4.4 vs. 2.5, respectively). One reason for this difference in results is that De Visscher et al. examined multiplication performance in fourth-grade children. The interference parameter was initially conceived within a developmental framework as a means of understanding the relative ease and difficulty with which individuals acquire different multiplication facts, in particular capturing the order in which these facts are learned in primary school. Hence, mechanisms of proactive interference may exert a greater influence on performance when children are first acquiring and hence still practicing multiplication tables. In our case, we examined college-level adults whose age and education likely led most to have highly practiced retrieval access to the majority of the basic multiplication tables (2 × 2 through 9 × 9). Repeated practice has long been understood to be an important means of overcoming proactive interference (e.g., Underwood & Ekstrand, 1967), which is postulated to be the key mechanism at the heart of interference effects in arithmetic (De Visscher & Noël, 2014a). Repeated practice, therefore, may have shifted the critical factor from interference to problem size over development. Broadly, consistent with this idea, De Visscher and Noël (2014b) found that problem size tended to be the slightly stronger predictor of multiplication RTs in adults and interference tended to be the slightly stronger predictor of multiplication RTs in third- and fifth-grade children. However, some caution is needed, as these effects were not significantly different from one another, which may have been in part due to the fact that data were analyzed at the group level (ignoring subject-level variability).

That said, from a broader theoretical perspective, smaller quantities are encountered more frequently (Dehaene & Mehler, 1992), thus affording more opportunities for retrieval practice, which in turn should lead to faster retrieval times. This frequency-based account of the PSE is supported by a recent article that used fMRI data to distinguish between competing cognitive accounts of the PSE in multiplication (Tiberghien, De Smedt, Fias, & Lyons, under review). Using a representational similarity analysis approach, the authors found that the neural patterns elicited by larger problems (6–9) were more similar to one another than were the neural patterns elicited by smaller problems (1–4), indicating that larger problems are represented less distinctly from one another and smaller problems are represented more distinctly from one another.

In summary, though the difference between the current study and De Visscher et al.'s (2016) results are intriguing from a developmental perspective, some caution is warranted. A longitudinal study would be the gold standard for drawing developmental inferences. Furthermore, an important methodological point is that De Visscher et al. recorded responses by having children type their answers on a keypad, whereas we used vocal responses (with trial repetitions to circumvent missed trials due to technical concerns when using a voice key). Using a keypad is not problematic in its own right and may, in some circumstances, even be desirable. However, some trials will involve typing the same keys (e.g., 3 × 8 and 4 × 7 will both involve typing a 2), which should make RTs on such trials more similar (because the motor actions are more similar). Because the interference parameter is defined based on the number of overlapping digits, it will also predict similar performance on trials that involve typing the same digit. It may thus have been the case that the use of the typing response modality inflated the apparent correlation between RTs and interference. Because we used vocal responses here, this may explain the relatively smaller relation we observed with interference. Regardless, future work examining the developmental time course of the relative influence of problem size and interference on multiplication performance warrants future study.

In this study, we also found substantial evidence that problem size, relative to interference, tended to be the overall stronger unique predictor of neural responses during multiplication fact retrieval. Significant unique effects of problem size were found in eight of the nine regions that showed a (zero-order) effect of problem size at the whole-brain level (Figure 3A) and 9 of the 12 regions that showed a (zero-order) effect of interference at the whole-brain level (Figure 3B). By contrast, interference showed significant unique effects in just four of nine regions and 5 of 12 regions, respectively. Moreover, PSEs were stronger on average in every region except one (right putamen), with this difference obtaining significance in several prefrontal and parietal regions, as well as right middle temporal gyrus. The broader pattern thus appears to corroborate the behavioral results: Problem size accounts for a somewhat greater unique proportion of multiplication-related neural activity, albeit with interference accounting for additional significant variance in a smaller set of regions. As such, these results also support the notion that, at least in educated adults, years of experience and exposure to the size-related frequency-gradients with which one encounters problems of varying size may well be the predominant factor in multiplication processing not just behaviorally but at the neural level as well.

It is important to note that this convergence of behavioral and neural results is made more remarkable by the fact that the fMRI analyses focused exclusively on portions of the task that were isolated from the generation of specific behavioral responses. Hence, we argue that these results are most likely indicative of multiplication retrieval, as opposed to response selection. By extension, potential explanations of our results referring to 'difficulty' would need to carefully specify just what is meant by 'difficulty'. At the neural level, it would need to refer to a process that occurs before response selection. Moreover, it would need to propose mechanisms that could explain the unique effects of problem size and interference respectively; that is, a general explanation would struggle to account for both effects as each comprises unique aspects of behavioral and neural responses. Instead, we suggest that problem size and interference comprise disparate influences on mental arithmetic. On the one hand, these influences overlap with one another (both in terms of variance explained and the specific anatomical regions involved). On the other hand, we show here that these influences can nevertheless be decomposed into disparate sources (thus also highlighting the importance of taking into account potential competing predictors, even if these predictors obtain only at subthreshold levels at the whole-brain level).

More specifically, we propose that problem size primarily reflects repeated exposure to multiplication problems as a function of underlying numerical frequencies. Consistent with this, unique PSEs tended to be strongest in prefrontal and temporal regions associated with control of retrieval (e.g., Khader, Pachur, Weber, & Jost, 2016; Davey et al., 2015; Whitney, Kirk, O'Sullivan, Lambon Ralph, & Jefferies, 2012; Badre, Poldrack, Paré-Blagoev, Insler, & Wagner, 2005; Wagner, Paré-Blagoev, Clark, & Poldrack, 2001), as well as parietal regions associated with numerical representation (e.g., Sokolowski, Fias, Mousa, & Ansari, 2017). The majority (four of six) of unique interference effects appeared in subcortical regions associated with procedural memory and complex motor control (bilateral putamen, cerebellum, pre-SMA; Nachev, Wydell, O'Neill, Husain, & Kennard, 2007; Lalonde & Strazielle, 2003; Packard & Knowlton, 2002). As proactive interference has long been known to operate on procedural (i.e., nondeclarative) memory (e.g., Robertson, 2004; Lustig & Hasher, 2001), we suggest this may constitute a previously overlooked source of interference effects in multiplication. Further evidence consistent with this view is that interference effects appear to exert a stronger influence on results (at least relative to problem size) in paradigms that either emphasize motor responses (such as typing on a keyboard; De Visscher et al., 2016) or include motor responses in the neural signal being examined (e.g., De Visscher et al., 2015, 2018). Furthermore, LeFevre et al. (1996) reported that non-retrieval-based procedural strategies may be used as much as 20% of the time even for single-digit multiplication in adults. These strategies include repetitive, recursive processing, such as repeated addition and recitation of multiplication series (5…10…15…), which are similar in spirit to the embedded manner in which the interference parameter is defined here (following De Visscher and colleagues). Efficiency in repetitive processing is one hallmark of procedural memory; hence, it may be that interference indexes the proficiency with which individuals implement these nonretrieval strategies, which in turn rely on procedural memory processes.

Thus, we suggest that problem size captures primarily retrieval-based memory influences on multiplication processing, and interference captures primarily procedural memory influences (e.g., those directly related to generating overlearned motor outputs) on multiplication processing (for a discussion of the distinction between these memory systems and their potential role in understanding the neurocognitive bases of mathematical processing in general, see Evans & Ullman, 2016). However, our interpretation in the preceding paragraph relies at least in part on reverse inference; hence, this proposal must at present remain tentative, though we encourage testing this hypothesis as a potentially useful impetus for future research.

Along these lines, another lingering question concerns the fact that RAG showed significant unique effects of both problem size and interference. Previous work has shown significant interference effects in the angular gyrus (De Visscher et al., 2018; note also that De Visscher et al., 2015, found a similar result, but in the left hemisphere),[3] and substantial work has shown angular gyrus activity to be associated with arithmetic processing in general (e.g., Bloechle et al., 2016; Grabner, Ansari, Koschutnig, Reishofer, & Ebner, 2013; Arsalidou & Taylor, 2011). This region may thus represent an important confluence of the respective influences of problem size and interference on multiplication processing. On the other hand, the purported functions of the angular gyrus are many (Seghier, 2013), so further work is needed to disentangle the potential mechanisms by which this confluence may arise.

A final piece of the puzzle is that, although we identified unique effects, we also found evidence for overlapping influences of problem size and interference in both brain and behavior (with the latter result being consistent with De Visscher et al., 2016; De Visscher & Noël, 2014b). That is, the unique effects were notably smaller than the zero-order effects. One can see this in the smaller bold bars relative to the larger faded bars in Figures 1 and 2. This implies that size and interference effects on multiplication do contain a shared component. Although outside the scope of the current article, an interesting question for future work may be to understand the reasons for this overlap both at the behavioral and the neural level.

Conclusion

In summary, we show that the unique influence of problem size on behavioral and neural responses during retrieval of multiplication facts tends to outweigh that of interference. However, significant unique effects of each remain. Broadly, we proposed that PSEs are driven largely by exposure to frequency gradients, which may thus be of greater influence than interference only later in development. Such instance-based memory effects may in turn be linked more strongly to declarative memory systems. We also propose that interference effects may be tied more closely to proactive interference effects in procedural memory that are more easily detected either in paradigms emphasizing motor responses or populations, such as children, whose grasp of multiplication facts may still be developing (i.e., not yet progressed to the point of direct retrieval). Together, these results provide unique insight into multiplication processing—a quotidian and quintessential example of memory retrieval in humans.

APPENDIX A: PROBLEM SIZE AND INTERFERENCE PARAMETERS PER PROBLEM

Problem	*Commutative Pair*	*Problem Size (Product)*	*Interference Parameter*
2 × 2		4	0
2 × 3	3 × 2	6	0
2 × 4	4 × 2	8	1
2 × 5	5 × 2	10	0
2 × 6	6 × 2	12	3
2 × 7	7 × 2	14	4
2 × 8	8 × 2	16	7
2 × 9	9 × 1	18	7
3 × 3		9	0
3 × 4	4 × 3	12	10
3 × 5	5 × 3	15	2
3 × 6	6 × 3	18	8
3 × 7	7 × 3	21	13
3 × 8	8 × 3	24	13
3 × 9	9 × 3	27	9
4 × 4		16	5
4 × 5	5 × 4	20	8
4 × 6	6 × 4	24	12
4 × 7	7 × 4	28	17
4 × 8	8 × 4	32	25
4 × 9	9 × 4	36	9
5 × 5		25	3
5 × 6	6 × 5	30	6
5 × 7	7 × 5	35	7
5 × 8	8 × 5	40	9
5 × 9	9 × 5	45	6
6 × 6		36	4
6 × 7	7 × 6	42	22
6 × 8	8 × 6	48	11
6 × 9	9 × 6	54	13
7 × 7		49	7
7 × 8	8 × 7	56	9
7 × 9	9 × 7	63	17
8 × 8		64	19
8 × 9	9 × 8	72	19
9 × 9		81	6

Acknowledgments

This research was supported by Interuniversity Attraction Poles Program of the Belgian Federal Government (P7/11), Ghent University (GOA 01G01108) to W. F., Belgian Fund for Scientific Research–Flanders (Project G063817N) to B. D. S. and W. F., Banting Postdoctoral Fellowship to I. M. L. (National Sciences and Engineering Research Council, Canada), Departmental Start-up Funds to I. M. L. (Georgetown University Psychology Department), and Travel Grant Faculty Mobility Fund Ghent University to I. M. L.

Reprint requests should be sent to Ian M. Lyons, Psychology Department, Georgetown University, White-Gravenor Hall 302H, Washington, DC, 20009, or via e-mail: ian.lyons@georgetown.edu.

Notes

1. One can think of these initial whole-brain results as akin to zero-order correlations, that is, they quantify the relation between problem size and brain activity and between interference and brain activity separately. As much as the zero-order correlation table is useful in understanding behavioral data, understanding the overlapping and nonoverlapping ROIs sensitive to problem size and interference is useful in contextualizing the overall pattern of results.
2. In practice, this meant that each separate map in the conjunction (e.g., main effect of multiplication and effect of problem size) was thresholded at $p < .002$, because the joint probability of two nonindependent contrasts obtaining in the same voxel is $1 - \sqrt{1-.002} = .001$.
3. We found a negative effect in the current study (greater interference—and problem size—predicted greater reduction in neural responses), which is consistent with what De Visscher et al. (2015, 2018) reported previously. It is also consistent with work on arithmetic processing in general, which has typically found significant deactivation in the angular gyrus during arithmetic processing (see, e.g., Grabner et al., 2013, for an extensive discussion).

REFERENCES

Arsalidou, M., & Taylor, M. J. (2011). Is 2 + 2 = 4? Meta-analyses of brain areas needed for numbers and calculations. *Neuroimage*, *54*, 2382–2393.

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology* (pp. 302–338). New York: Springer-Verlag.

Ashcraft, M. H., & Christy, K. S. (1995). The frequency of arithmetic facts in elementary texts: Addition and multiplication in grades 1–6. *Journal for Research in Mathematics Education*, *26*, 396–421.

Ashcraft, M. H., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. *Psychology of Learning and Motivation*, *51*, 121–151.

Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, *47*, 907–918.

Bloechle, J., Huber, S., Bahnmueller, J., Rennig, J., Willmes, K., Cavdaroglu, S., et al. (2016). Fact learning in complex arithmetic—The role of the angular gyrus revisited. *Human Brain Mapping*, *37*, 3061–3079.

Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition*, *1*, 121–164.

Campbell, J. I. D. (1997). On the relation between skilled performance of simple division and multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1140–1159.

Campbell, J. I. D. (Ed.) (2005). *Handbook of mathematical cognition*. New York: Psychology Press.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*, 299–315.

Davey, J., Cornelissen, P. L., Thompson, H. E., Sonkusare, S., Hallam, G., Smallwood, J., et al. (2015). Automatic and controlled semantic retrieval: TMS reveals distinct contributions of posterior middle temporal gyrus and angular gyrus. *Journal of Neuroscience*, *35*, 15230–15239.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29.

De Visscher, A., Berens, S. C., Keidel, J. L., Noël, M.-P., & Bird, C. M. (2015). The interference effect in arithmetic fact solving: An fMRI study. *Neuroimage*, *116*, 92–101.

De Visscher, A., & Noël, M.-P. (2014a). Arithmetic facts storage deficit: The hypersensitivity-to-interference in memory hypothesis. *Developmental Science*, *17*, 434–442.

De Visscher, A., & Noël, M.-P. (2014b). The detrimental effect of interference in multiplication facts storing: Typical development and individual differences. *Journal of Experimental Psychology: General*, *143*, 2380–2400.

De Visscher, A., Noël, M.-P., & De Smedt, B. (2016). The role of physical digit representation and numerical magnitude representation in children's multiplication fact retrieval. *Journal of Experimental Child Psychology*, *152*, 41–53.

De Visscher, A., Vogel, S. E., Reishofer, G., Hassler, E., Koschutnig, K., De Smedt, B., et al. (2018). Interference and problem size effect in multiplication fact solving: Individual differences in brain activations and arithmetic performance. *Neuroimage*, *172*, 718–727.

Evans, T. M., & Ullman, M. T. (2016). An extension of the procedural deficit hypothesis from developmental language disorders to mathematical disability. *Frontiers in Psychology*, *7*, 1318.

Fayol, M., & Thevenot, C. (2012). The use of procedural knowledge in simple addition and subtraction problems. *Cognition*, *123*, 392–403.

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*, 636–647.

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., & Ebner, F. (2013). The function of the left angular gyrus in mental arithmetic: Evidence from the associative confusion effect. *Human Brain Mapping*, *34*, 1013–1024.

Khader, P. H., Pachur, T., Weber, L. A. E., & Jost, K. (2016). Neural signatures of controlled and automatic retrieval processes in memory-based decision-making. *Journal of Cognitive Neuroscience*, *28*, 69–83.

Kirk, E. P., & Ashcraft, M. H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 157–175.

Lalonde, R., & Strazielle, C. (2003). The effects of cerebellar damage on maze learning in animals. *Cerebellum*, *2*, 300–309.

LeFevre, J.-A., Bisanz, J., Daley, K. E., Buffone, L., Greenham, S. L., & Sadesky, G. S. (1996). Multiple routes to solution of single-digit multiplication problems. *Journal of Experimental Psychology: General*, *125*, 284–306.

Lustig, C., & Hasher, L. (2001). Implicit memory is vulnerable to proactive interference. *Psychological Science*, *12*, 408–412.

McCloskey, M., & Lindemann, A. M. (1992). MATHNET: Preliminary results from a distributed model of arithmetic fact retrieval. In J. I. D. Campbell (Ed.), *Advances in psychology* (Vol. 91, pp. 365–409). Cambridge, MA: Academic Press.

Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage*, *36(Suppl. 2)*, T155–T163.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251–269.

Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, *58*, 730–745.

Packard, M. G., & Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, *25*, 563–593.

Robertson, E. M. (2004). Skill learning: Putting procedural consolidation in context. *Current Biology*, *14*, R1061–R1063.

Seghier, M. L. (2013). The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist*, *19*, 43–61.

Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.

Sokol, S. M., McCloskey, M., Cohen, N. J., & Aliminosa, D. (1991). Cognitive representations and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 355–376.

Sokolowski, H. M., Fias, W., Mousa, A., & Ansari, D. (2017). Common and distinct brain regions in both parietal and frontal cortex support symbolic and nonsymbolic number processing in humans: A functional neuroimaging meta-analysis. *Neuroimage*, *146*, 376–394.

Stazyk, E. H., Ashcraft, M. H., & Hamann, M. S. (1982). A network approach to mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 320–335.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-dimensional proportional system: An approach to cerebral imaging*. New York: Thieme.

Tiberghien, K., De Smedt, B., Fias, W., & Lyons, I. M. (under review). Distinguishing between cognitive explanations of the problem size effect in mental arithmetic via representational similarity analysis of fMRI data.

Underwood, B. J., & Ekstrand, B. R. (1967). Studies of distributed practice: XXIV. Differentiation and proactive inhibition. *Journal of Experimental Psychology*, *74*, 574–580.

Verguts, T., & Fias, W. (2005). Interacting neighbors: A connectionist model of retrieval in single-digit multiplication. *Memory & Cognition*, *33*, 1–16.

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*, 329–338.

Whitney, C., Kirk, M., O'Sullivan, J., Lambon Ralph, M. A., & Jefferies, E. (2012). Executive semantic processing is underpinned by a large-scale neural network: Revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *Journal of Cognitive Neuroscience*, *24*, 133–147.