PERGAMON

Neuropsychologia 38 (2000) 325–335

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

The calculating brain: an fMRI study

T.C. Rickard[a], S.G. Romero[b], G. Basso[c], C. Wharton[c], S. Flitman[d], J. Grafman[c,]*

[a]Department of Psychology, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0901, USA
[b]University of Colorado, Boulder, CA, USA
[c]Cognitive Neuroscience Section, National Institutes of Health, 10 Center Drive, MSC 1440, Bethesda, MD 20892-1440, USA
[d]Barrow Neurological Institute, Phoenix, AZ, USA

Received 13 March 1998; received in revised form 15 February 1999; accepted 12 April 1999

Abstract

To explore brain areas involved in basic numerical computation, functional magnetic imaging (fMRI) scanning was performed on college students during performance of three tasks; simple arithmetic, numerical magnitude judgment, and a perceptual-motor control task. For the arithmetic relative to the other tasks, results for all eight subjects revealed bilateral activation in Brodmann's area 44, in dorsolateral prefrontal cortex (areas 9 and 10), in inferior and superior parietal areas, and in lingual and fusiform gyri. Activation was stronger on the left for all subjects, but only at Brodmann's area 44 and the parietal cortices. No activation was observed in the arithmetic task in several other areas previously implicated for arithmetic, including the angular and supramarginal gyri and the basal ganglia. In fact, angular and supramarginal gyri were significantly deactivated by the verification task relative to both the magnitude judgment and control tasks for every subject. Areas activated by the magnitude task relative to the control were more variable, but in five subjects included bilateral inferior parietal cortex. These results confirm some existing hypotheses regarding the neural basis of numerical processes, invite revision of others, and suggest productive lines for future investigation. Published by Elsevier Science Ltd.

Keywords: fMRI; Math; Cognition; Arithmetic; Number processing

4. Introduction

Two central components of numerical cognition are simple arithmetic (e.g., $4 \times 7 = ?$) and magnitude judgment (e.g., 24 or 25, which is larger?). We used fMRI to explore patterns of neural activation for these tasks, with two related goals. First, we sought to confirm and extend findings of the neuropsychological literature. A wealth of patient data suggests that arithmetic is mediated by left or possibly bilateral inferior parietal areas [1,2,14–16,23,29,38,40]. Some evidence suggests that the region around the left angular and supramarginal gyri may be particularly important for these tasks [15,17,18,20,21,39,40]. Other studies have tentatively implicated the left frontal lobe [24,35], the basal ganglia [5,19,41], and the thalamus [27]. However, the complete set and precise localization of critical structures has yet to be conclusively established. In contrast to the apparent left hemisphere bias for arithmetic, there is some evidence suggesting that simple magnitude processing involves right or perhaps bilateral parietal areas [7–10,22]. A number of functional neuroimaging experiments exploring arithmetic in normals have also been conducted [3,10,30,31,33], most of these investigated relatively complex calculations such as counting backward by sevens from a three digit number [30]. Activation has typically been observed in inferior parietal and (or) prefrontal areas in these studies, providing a rough convergence with at least some patterns in the patient data. However, complex arithmetic tasks likely engage a number of cognitive processes not directly associated with arithmetic and thus can not address the issue of which areas of brain acti-

---

* Corresponding author.

E-mail addresses: trickard@ucsd.edu (T.C. Rickard), jgr@box-j.nih.gov (J. Grafman).

0028-3932/00/$ - see front matter Published by Elsevier Science Ltd.
PII: S0028-3932(99)00068-8

vation correspond to specific component processes of interest, such as arithmetic fact retrieval or magnitude comparison.

One exception is a positron emission topography (PET) study by Deheane and colleagues, which was designed explicitly to study arithmetic and magnitude processing [10]. Their subjects were shown two single-digit numbers and they either multiplied them together or identified the larger number. The multiplication task compared to a pure rest condition (i.e., involving no active stimulus processing) revealed right cuneus, bilateral inferior parietal, and left fusiform and lingual gyrus activation. There was also lateral occipital activation, which was interpreted as reflecting visual processing, and precentral and supplementary motor area activation, interpreted as reflecting motor responses associated with subvocalization of the answer. Deheane et al. observed no reliable activation in other areas previously implicated in arithmetic, including prefrontal cortex, thalamus and basal ganglia. Of particular note, there was no reliable angular gyrus activation, and the left supramarginal gyrus was significantly deactivated during multiplication compared to rest. The magnitude task compared to rest did not reliably activate any areas beyond the perceptual and motor areas, but there was a trend toward bilateral inferior parietal activation. Deheane et al. suggested that the left inferio–mesial–occipito–temporal area (including fusiform and lingual gyri) is involved in identifying digits and transmitting their identity to other areas, that basal ganglia may be involved in retrieving rotely memorized arithmetic facts (they were drawing here on the neuropsychological literature; they found no activation in this area), and that the inferior parietal areas represent abstract magnitude information which may be brought to bear for magnitude comparison and also for arithmetic when rote retrieval fails [9].

If the areas suggested in the above summary of patient and neuroimaging data are accurate and exhaustive, then we should find activation in the following areas for simple arithmetic; left or bilateral inferior parietal cortex (possibly centered around the angular gyrus), bilateral thalamus, left basal ganglia, left fusiform and lingual gyri, and perhaps left prefrontal cortex. Predictions for magnitude judgment relative to control are less well established, but the available evidence suggests bilateral inferior parietal cortex with a possible bias toward the right.

A second and related goal of this study was to provide new insight into possible relations between cognitive representations for arithmetic and for numerical magnitude. One possibility, consistent with modular and connectionist models of number processing [25,26,37], is that arithmetic facts are stored and represented as part of an abstract network dedicated to representing numerical magnitude. This perspective suggests that substantially overlapping brain areas should be activated by arithmetic fact retrieval and magnitude judgment. An alternative framework advanced by Dehaene posits separate systems for representing analog magnitude, verbal, and visual number forms [6]. In this model, arithmetic is mediated primarily by the verbal code, with semantic, or magnitude, codes providing a supportive role if the verbal pathway fails. This perspective would be consistent with a finding that arithmetic and magnitude comparison activate largely distinct brain regions. One viable possibility is a left hemisphere bias for arithmetic and a right hemisphere bias for magnitude processes.

The experimental design involved three conditions: (1) *multiplication verification*, in which a problem was presented along with a candidate answer, and subjects responded whether the candidate answer was true or false (e.g., $4 \times 7 = 35$; true or false?), (2) *magnitude judgment*, in which two two-digit numbers were presented (e.g., 24 25) and subjects responded whether the left or right side number was larger, (3) and a *detect ones* control condition, in which four digits were presented and subjects were required to determine whether one of them was a 1 (e.g., 4 2 1 7; is there a 1 present?). Each type of stimulus required a dichotomous response decision (true/false for verification, left/right for magnitude judgment, and yes/no for detect ones). Areas more activated by verification relative to the detect ones control task should primarily reflect processes involved in solving arithmetic problems (i.e., factoring out perceptual, digit, and motor processes), and areas more activated by magnitude judgment relative to detect ones condition should reflect primarily processes uniquely involved in magnitude judgment. The direct comparison of verification and magnitude judgment should reveal areas more activated by each task of interest relative to the other.

These tasks have several properties that make them well suited for initial exploration of basic arithmetic and magnitude processes. First, each task condition consists only of digits and arithmetic symbols and subtends identical vertical and horizontal visual angles. Thus, basic perceptual differences among tasks are negligible and are not likely to result in significant activation differences outside of primary visual areas which are of little interest in this study. Second, each type of task is nevertheless clearly distinct from the others. Subjects will therefore have no confusion about which task is to be performed on the stimuli during scanning. In pilot behavioral work, we observed that, if tasks are perceptually identical and only task goals are changed over conditions, subjects sometimes become confused about which task is to be performed. For example, in a pilot behavioral study we presented items such as '4 7 35' under alternating instructions either to multiply the first two digits and compare the result to the second two (i.e., a verification task) or to determine the larger magnitude of the two two-digit numbers. Subjects occasionally confused these two task goals during performance. Further, even if overt confusion did not occur, exact perceptual equivalence of stimuli might promote automatic processing of the inappropriate task. These factors clearly would cause serious problems with interpretability of fMRI results. By incorporating relatively minor perceptual differences in the items, we completely avoid the problem of task goal confusions, at the sacrifice of introducing only modest perceptual differences.

A third desirable property of the tasks is that they all involve numerical stimuli. Our primary interest in this study was isolation of processes involved in arithmetic and magnitude processing, rather than processing of digits per se. The presence of four digits for each stimuli roughly equates basic digit processing in each condition and thus activation reflecting basic digit processing should not be present in the task comparisons. The importance of using a digit control task can be demonstrated by considering an alternative design in which the control is either a pure rest condition [10] or a non-numerical perceptual-motor control. In such a design, activation in the magnitude task relative to the control task might reflect the magnitude processes of interest, or, alternatively, either basic perceptual processes (in the case of the pure rest control), or non-magnitude aspects of visual digit processing (in the case of a non-numerical perceptual-motor control). Magnitude processes can only be isolated by using a control condition which involves every process except magnitude processing. Note that the detect one's control may also involve some magnitude processing, provided that such processing is automatically triggered by the perception of any numerical stimuli. However, it is reasonable a priori to expect magnitude processing to be of greater magnitude in the number comparison task, which explicitly requires accessing of magnitude information to execute the task goal. Thus, we would still expect to see activation in magnitude processing areas in the magnitude judgment vs detect 1's control task.

A final property of the tasks is that they all involve roughly equivalent, dichotomous motor responses that provide accuracy data. Our decision to use the verification task rather than a simpler production task in which subjects produce the answer (e.g. $4 \times 7 = ?$) was motivated by an overriding goal of collecting accuracy data. Not only could accuracy data not be collected for the somewhat simpler production task using the dichotomous response buttons at our disposal, but subject vocalization of the answer in that task would result in head movement which might seriously compromise image quality.

If verification is to reflect the underlying task of primary interest (production), we need to assure that subjects use a produce-compare strategy, in which they first produce the answer (probably by subvocalization), and then compare it to the candidate answer given, to solve the problems. To maximize use of a produce-compare strategy, we (a) selected subjects who reported that they were easily able to retrieve answers to problems (this is important because ease of retrieval is likely to increase use of a produce-compare strategy in verification) [32], (b) used problem-answer combinations which elicited a high frequency of produce-compare strategy reports in previous research [32], and (c) instructed subjects to use a produce-compare strategy. These steps do not guarantee that subjects used produce-compare during scanning for all verification problems, but they do allow for reasonable confidence that this was the dominant strategy.

2. Method

2.1. Subjects

Subjects were eight right-handed, neurologically normal adults (five females, three males). The subjects' median age was 24 years (range = 20–34), their median education level was 16 years (range = 14–18), and their median scaled WAIS vocabulary score was 14 (range = 10–18). Each subject was screened for physical conditions that would preclude MRI scanning, gave informed, written consent to participate in MRI scanning, and was paid $70 for their participation.

2.2. Apparatus

Stimuli were presented by a Apple PowerMacintosh 8100 computer (Apple; Cupertino, CA, USA) using SuperLab (Cedrus; Wheaton, MD, USA) and were back projected with a magnetically shielded LCD video projector (Sharp; Mahwah, NJ, USA) onto a translucent screen placed at the subject's feet. Subjects were given earplugs and during scanning they were positioned supine with their head restrained by an adjustable metal band and foam padding. The subject looked upward and could see the screen by the use of a mirror in the head coil tilted at a 45° angle.

MRI images were obtained using a 1.5 T GE Signa (General Electric, Milwaukee, WI, USA) with gradient head coils designed for echo plannar imaging (EPI) [36]. An interleaved multislice gradient echo EPI scanning sequence was used to produce 18 axial sections, each 5 or 6 mm thick (depending on the size of the subject's head), with a 64×64 matrix and a 24 cm field of view (repetition time (T_R) = 3000 ms, echo time (T_E) = 40 ms, flip angle = 90°). In order to obtain T1 weighted anatomical images of each subject, 3 d SPGR structural images were acquired at the same locations as the echo-planar images ($T_E = 5.4$, flip angle = 20°, matrix = 256 × 192).

2.3. Materials

The stimuli for the detect ones task consisted of 60 unique strings of four digits (e.g., 4 1 8 7). Thirty of these strings included a single 1 (necessitating a true response) and 30 strings did not include a 1 (necessitating a false response). The number '1' occurred equally often in the four positions across items. Foil numbers were randomly determined with the constraint that each number, two through nine, occurred equally often. These 60 strings were presented in a different randomized order for each run.

Thirty-four single digit multiplication problems were used as stimuli for the runs using the verification task. All problems were presented twice during each run with the exception of all tie problems (e.g., 6 × 6) and one non-tie problem (i.e., 2 × 4), which were presented once. Each repeated problem was presented once with each of the two operand orders (e.g., 4 × 5 and 5 × 4). One randomly selected operand order was paired with a correct answer (necessitating a true response) and the other was paired with an incorrect answer (necessitating a false response). Across every four runs that included the verification task, each problem was presented twice with correct answers and twice with incorrect answers. The incorrect answers were also varied such that across repetition each problem was presented with two different incorrect answers. Furthermore, each of the false answers was related to the multiplication table of one of the operands and was a small distance away from the correct answer. Table related false answers that are a small distance from the correct answer facilitate a *retrieve–compare* strategy in which an answer is retrieved from memory and then compared to the answer given [32].

The stimuli for the magnitude comparison task consisted of 60 pairs of two digit numbers between 14 and 81. All numbers used were legal multiplication answers and corresponded to the answers to problems used in the verification task. Both small and large difference between the numbers were included in this set of number pairs. Thirty pairs for each run were presented with the larger number on the right and 30 were presented with the larger number on the left, randomly intermixed. Across every two runs that included the magnitude comparison task, each number pair was presented once with the larger number on the right and once with the larger number on the left.

2.4. Procedure

Before scanning, subjects were first trained in the multiplication verification condition. This training consisted of an experimenter verbally testing the subject on all single-digit number combinations between 2 and 9 in the multiplication table (e.g., '7 × 8 = ?'). Subjects were tested on each number combination until they reported that they were able to directly retrieve the answer from memory and could provide the correct answer within 2 s. Subjects were informed that, during the scanning session, they should work as quickly and accurately as possible and use a strategy of retrieving the correct answer from memory and then checking their result against the displayed answer. This second instruction was employed to help minimize possible usage of a magnitude estimation strategy to solve the verification problems [32].

Stimuli in all conditions were presented for 2100 ms each and there were 20 stimuli presented per 42 s condition. Subjects indicated yes (or true) responses and no (or false) responses by pressing buttons held, respectively, in their right and left hands. Response time (RT) and accuracy was recorded for each trial.

After being placed in the scanner, each subject participated in two practice runs (45 s each), in which no fMRI scanning was performed, followed by nine (six subjects) or 12 (two subjects) experimental fMRI runs (284 s each). All runs were preceded by an experimenter announcing the two conditions in which the subject would participate on that run. Each run then began with a 12 s display stating 'Prepare to begin.' The four scans collected over this 12 s interval at the beginning of each experimental run were included to eliminate saturation effects and were removed from the data sets prior to analysis. Within each experimental run, each block of trials started with a screen displayed for 1000 ms that identified that block's condition and the appropriate left–right button responses. Each run consisted of three interleaved alternations of the two task conditions, with 42 s per task condition. All three condition combinations (detect ones vs magnitude judgment; detect ones vs verification; magnitude judgment vs verification) were presented in an interleaved fashion across multiple runs for each subject. Depending on time constraints, subjects were tested for either three or four runs on each task pair. Prior to these fMRI runs, two practice runs were given. One of these involved five trials of the control task followed by five trials of the magnitude task, and then repeated this sequence. The second practice run was similar to the first, with the exception that the verification task replaced the magnitude task.

2.5. Analysis

All data transformation and statistical analyses were performed using SPM96b (Wellcome Department of Cognitive Neurology, UK) implemented in Matlab

(Mathworks Inc. Sherborn MA, USA). First, functional images from all runs for each subject were coregistered to the first functional image of the first run for that subject. Second, within each run values for each voxel were adjusted by removing any component correlated with a function of movement estimates, obtained at the time of the current and the previous scan [11]. Third, the SPGR anatomical image for each subject was spatially normalized to the Talaiarch atlas [34]. The affine transformation for the anatomical image was then applied on all realigned functional images for each subject. Functional images were not coregistered to anatomical images. However, given the very secure head restraints used, the small interval of time between collection of the anatomicals and the first functional, and the minimal subject movement across functional runs (as indicated by the SPM functional coregistration output, which typically indicated a maximum deviation from the starting location of all functional runs of around 1 mm), we can reasonably infer that movement between collection of the anatomical image and the first functional image was typically less than 1 mm.

Next, Gaussian spatial smoothing with full width at half maximum of two voxel widths in each dimension ($7.5 \times 7.5 \times 10$ for most subjects, and $7.5 \times 7.5 \times 12$ for subjects with 6 mm slice thickness) was applied to all functional images to assure that Gaussian random field assumptions of the SPM cluster analysis were met [42].

To estimate the statistical relation of each voxel to the task paradigm, a delayed box car function corresponding to the 'on-off' switching of the task comparison was fitted simultaneously to all run replications of a given task comparison for each subject, adjusted to account for expected hemodynamic lag and dispersion (Gaussian kernel of $t = 2.82$ s). Several other SPM statistical adjustments were included as additional covariates at this stage of analysis, including: (a) high frequency filtering (with maximum frequency of two task alternation periods, or 168 s), (b) adjustment for global signal variation, (c) temporal smoothing, and (d) adjustment of degrees of freedom to account for temporal autocorrelation [13]. Each run of a given task comparison was treated as a separate 'subject'. Parameters for each covariate were estimated for each and every voxel according to the general linear model. The estimates for the covariate modeling the signal response to the tasks were then used in subsequent analyses described below.

Before computing and displaying the statistical parametric maps [12], we applied the conjunction analysis option in SPM96b. This analysis allows results from two different task comparisons which are hypothesized to isolate the same process to be combined, thus increasing statistical power while also eliminating some sources of potential artifact [28]. Consider the verification vs detect ones comparison and the verification vs magnitude judgment comparisons. In both comparisons, we expect to observe areas activated uniquely by verification. There is some danger, however, that some activated areas for verification might in fact reflect deactivation caused by the reference task (either detect ones or magnitude comparison). Conjunction analysis allowed us to combine these two task comparisons to look for areas activated by verification in both comparisons while simultaneously eliminating areas activated in only one of the comparisons. Thus, this approach both increases statistical power (relative to only looking at, for example, verification vs detect ones), while also eliminating comparison specific activations which may reflect idiosyncratic influences of one of the reference tasks. Conjunction analysis was performed for verification as described above, and for magnitude judgment by combining the magnitude vs detect ones task comparison with the magnitude judgment vs verification task comparison. SPM cluster analysis was then performed using a voxel level P value of 0.001, and a global cluster level correction of 0.01. This analysis identified the set of clusters such that one or more clusters would be expected to occur by chance only about 1% of the time (Friston et al., 1994).

Table 1
Behavioral results for each subject[a]

Subject		Control	Magnitude	Verification
1	Mean RT	588	745	958
	Error	0 (0)	8 (1)	8 (0)
2	Mean RT	712	946	1186
	Error	9 (1)	34 (13)	50 (17)
3	Mean RT	619	895	1090
	Error	2 (0)	16 (0)	27 (1)
4	Mean RT	597	848	1049
	Error	9 (0)	13 (0)	36 (10)
5	Mean RT	755	888	1078
	Error	15 (6)	9 (2)	45 (22)
6	Mean RT	529	727	855
	Error	5 (2)	12 (0)	5 (0)
7	Mean RT	600	841	1044
	Error	83 (82)	58 (52)	65 (58)
8	Mean RT	567	806	969
	Error	1 (1)	7 (1)	15 (3)

[a] RT in milliseconds. Error data indicate total number of omission plus commission errors. Data in parentheses indicate number of omission errors.

3. Results

3.1. Behavioral

Behavioral results for each subject are summarized

Table 2
Activation loci by subject and region

Subject		Parietal	Fusiform	Brodmann's area 44	Prefrontal
1	Left	−34, −52, 45	−49, −49, −20	−41, 4, 30	−34, 45, 25
	Right	26, −49, 55	3, −79, −15	38, 8, 25	34, 45, 20
2	Left	−26, −68, 35	−22, −79, −10	−41, 8, 25	−34, 45, 25
	Right	38, −60, 30	30, −79, −15	38, 8, 25	34, 45, 20
3	Left	−22, −56, 50	−41, −68, −10	−41, 4, 35	−38, 34, 26
	Right	22, −71, 35	34, −79, −10	41, 0, 30	34, 41, 25
4	Left	−11, −79, 55	−45, −56, −10	−45, 30, 40	−34, 56, 25
	Right	22, −75, 60	30, −82, −10	56, 22, 5	45, 49, 15
5	Left	−25, −49, 40	−49, −52, −12	−49, 11, 36	−38, 52, 30
	Right	39, −41, 54	26, −86, 0	45, 8, 30	49, 56, 10
6	Left	−26, −56, 30	−38, −41, −6	−30, 15, 30	−19, 52, 12
	Right	26, −60, 36	38, −52, −18	52, 19, 24	49, 45, 30
7	Left	−22, −71, 48	−41, −49, −18	−30, 19, 24	−41, 41, 18
	Right	26, −75, 48	26, −82, −18	34, 22, 18	41, 41, 30
8	Left	−34, −52, 35	−45, −45, −10	−45, 11, 35	−30, 52, 20
	Right	38, −60, 35	30, −79, 0	56, 26, 30	36, 60, 15

in Table 1. Error rates during the fMRI runs were low. Overall, subjects failed to respond in the time available (omission errors) on 3.38, 2.76, and 4.86% of the detect ones, magnitude judgment, and verification trials, respectively. Subjects made an incorrect response in the time available on 0.78, 2.83, and 3.83% of detect ones, magnitude comparison, and verification trials. Mean RT for correctly solved problems, averaged over subjects, was 620, 837, and 1028 ms for the detect ones, magnitude, and verification conditions, respectively. These differences were statistically significant in a one-way within subjects analysis of variance, $F(2,21)=276.4$, $P < 0.0001$.

3.2. fMRI

For each of the eight subjects, four general areas exhibited strong bilateral activation in the verification conjunction analysis: (1) a large contiguous activation area consisting of fusiform, lingual and inferior occipital gyri, (2) an inferior and superior parietal cluster (including cuneus and precuneus), (3) tissue proximal to Brodmann's area 44, forming a column extending upward to as high as Talairach $z = 40$, and downward to as low as Talairach $z = 0$, and 4) dorsolateral prefrontal areas 9 and 10. These activation clusters were statistically significant bilaterally in nearly all subjects, with exception of one subject who showed no reliable prefrontal activation, one subject who exhibited only left prefrontal activation, and a third subject who exhibited only left fusiform-lingual activation (but note that there were trends toward activation in these areas in each of these cases). Table 2 shows, for each subject, the Talairach coordinates of the maximum activated pixel for each of the four major areas in each hemisphere. Fig. 1 is an activation map for a typical subject, with both the structural and functional images normalized to Talairach coordinates. The image labeled '$Z=-10$' shows the bilateral fusiform and lingual gyrus activation. The image labeled '$Z = 30$' shows the Brodmann's area 44 cluster, the prefrontal cluster, and the inferior end of the parietal cluster. Also shown in that image are a few additional and less intense activation clusters. Such clusters occurred in apparently unsystematic patterns in some subjects but did not replicate even for two subjects. These clusters are thus ignored in the reported results. The image labeled '$Z = 45$' showed the superior end of the parietal and 'Brodmann's 44' clusters, along with anterior cingulate activation which was present for five subjects. Table 3 shows the mean Talairach coordinate location over subjects of the most intensely activated voxel in each of the four major areas. The distance between each subject's maximum activation coordinate in each area and the group maximum activation coordinate were also computed and averaged across subjects. This value, which provides an index of the variability of the maxima across subjects, is shown for each location under *Mean Deviation* in Table 3. These results reveal a striking degree of stability in the activation patterns across subjects and also highlights the rough bilateral symmetry of the activation maxima. In addition to these four primary areas, the anterior cingulate was significantly activated in five subjects, and the bilateral thalamus was significantly activated for two subjects.

In a separate non-conjunction analysis, areas deactivated for verification relative to the detect ones control condition were explored. For all subjects, deactivation was observed in bilateral superior and inferior temporal gyri, and a bilateral (but stronger left than right) area centered between the supramarginal and angular gyri (approximate Talairach coordinate of the centroid: $\times = +(-)$ 50, $y = 62$, $z = 32$). This cluster extended about 1 centimeter forward and backward in the Talairach y direction, and about 1 cm inward from the cortical surface (in the Talairach x direction). It had a Talairach z coordinate range of about $+25$ to $+40$. Note that the activated inferior parietal area for each subject in the verification conjunction analysis described previously included the areas immediately medial, superior, and posterior to this area. Analogous results were obtained in a non-conjunction analysis of the verification vs magnitude runs.

The conjunction analysis results for the magnitude judgment task revealed far less activity and much more variability across subjects. Areas activated included the right cuneus (two subjects), the bilateral medial frontal gyrus (Brodmann's area 6; three subjects), posterior cingulate (two subjects), and left or bilateral angular gyrus (two subjects). Three subjects

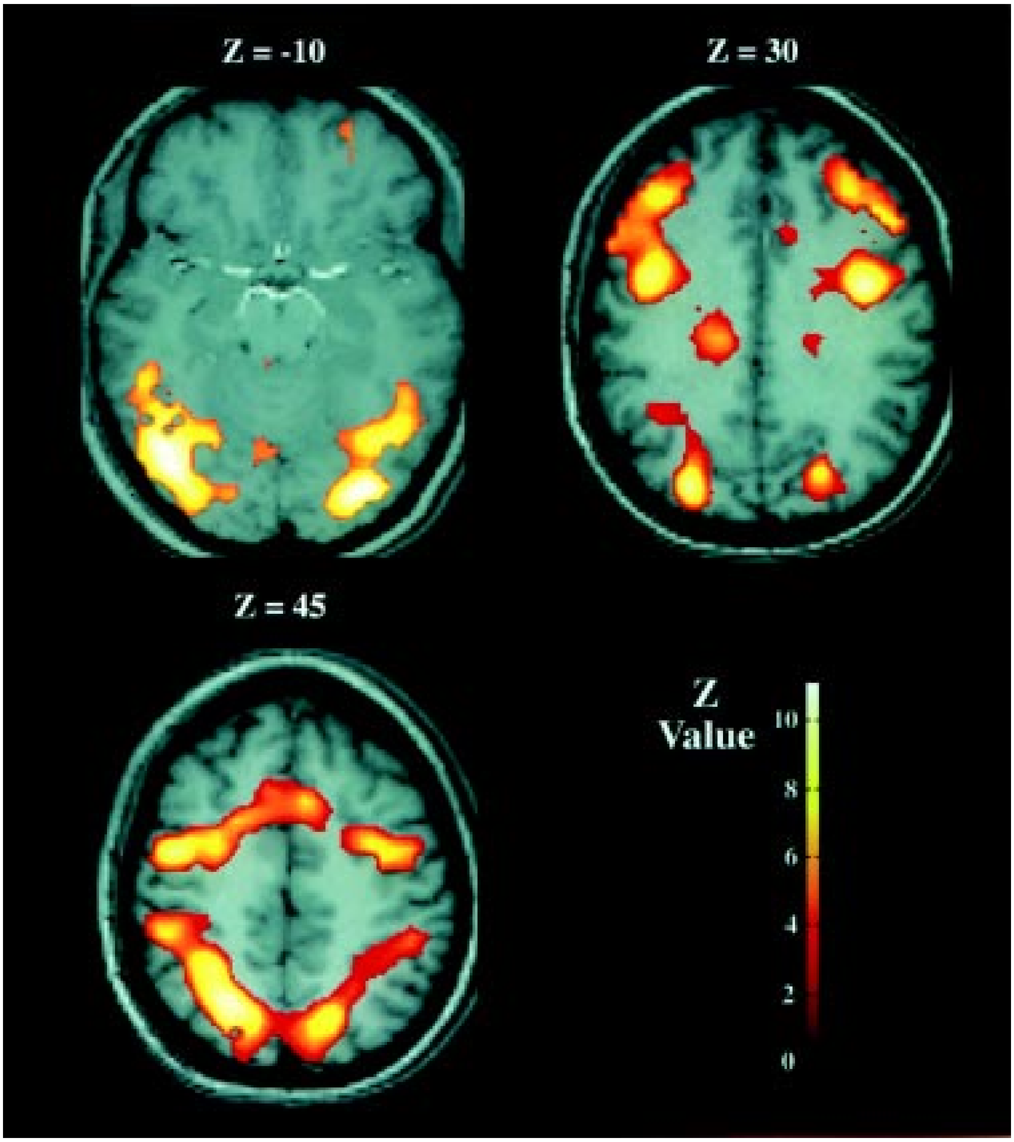

Fig. 1. Activation in the verification conjunction analysis for an example subject at three axial slices with Talairach *z* coordinates of −10, +30, +45.

exhibited no reliable activation in any area in this analysis.

It is possible that both the verification and magnitude judgment tasks activate largely the same brain areas and that verification produces more intense activation. This possibility is suggested by previous PET results, which revealed inferior parietal activation for both multiplication and magnitude judgment relative to rest [10]. If this is the case, then the conjunction analysis for magnitude judgment would not reveal these areas because activation there would be masked by the verification task. A supplementary analysis was thus performed limited to the magnitude judgment vs detect ones task comparison. In this analysis, two subjects showed reliable bilateral inferior parietal activation (this area was roughly the same as the inferior parietal area activated by the verification task), and three additional subjects exhibited trends toward this

Table 3
Activation loci collapsed over subjects

Brain area	Talairach coordinate			Mean deviation
	X	Y	Z	mm
Left Parietal	−25	−60	42	13.8
Right Parietal	29	−61	44	16.4
Left Fusiform	−41	−55	−12	13.0
Right Fusiform	31	−77	−10	9.9
Left Brodmann's 44	−40	13	32	10.9
Right Brodmann's 44	45	16	22	14.6
Left Brodmann's 9 and 10	−34	48	22	9.4
Right Brodmann's 9 and 10	41	48	22	10.6

same pattern of bilateral inferior parietal activation in an analysis which relaxed the significance criteria (voxel level $P < 0.05$). In addition, two subjects showed activation in the bilateral inferior frontal gyrus (near Brodmann's area 44; roughly the same area as activated in the verification task for these subjects), and two subjects exhibited bilateral fusiform activation (again, overlapping strongly with the same activation areas in verification). No subject exhibited the prefrontal activation (areas 9 and 10) which was observed for the verification task.

4. Discussion

The results extend neuropsychological and neuroimaging evidence regarding brain areas involved in basic numerical processing. We will focus our discussion first on each of the four major areas activated during the verification task. First consider the parietal activation. As expected, there was clear left hemisphere parietal activity in all eight subjects. There was also homologous parietal right hemisphere activity, although to a lesser degree. However, the focus of this activity in both hemispheres was roughly at the sulcus between the inferior parietal lobe and the precuneus (Talairch coordinate of the centroid: $x = +(-)\ 27$, $y = 61$, $z = 43$). This area is substantially posterior and superior to the angular and supramarignal gyri, which would have constituted our best guess about the focal point of inferior parietal activation based on the neuropsychological literature. In fact, there was actually a robust deactivation around the supramarginal and angular gyri for all eight subjects. Interestingly, an earlier PET study by Deheane et al. produced deactivation in supramarginal gyrus, and no reliable activation in the angular gyrus, for a multiplication production task [10]. Assuming that deactivation can be interpreted to reflect non-involvement of an area in verification, then these results appear to contradict neuropsychological results which implicate angular gyrus in arithmetic [15,17,18,20,21,39,40]. One possible explanation of these conflicting is that, although there is deactivation in the angular gyrus area itself in this task, there is activation for inferior parietal tissue which is proximal to the cortical surface of the angular and supramarginal gyri. Perhaps the previous neuropsychological results reflect brain damage not only to the angular and (or) supramarginal gyri but also to the juxtaposed structures which were activated in our study. In this scenario, it would be natural for researchers to interpret their findings in terms of the surface structures, but equally valid to have interpreted them in terms of adjacent and more medial structures that are activated in out study. This account does not directly explain the Rueckert et al. fMRI report of angular gyrus activity in the counting backward by sevens task. However, Rueckert et al. may have interpreted activation similar to ours as reflecting angular gyrus activation. This account seems plausible since there is no direct Talairach labeling of the inferior parietal area medial to the angular gyrus in which part of our activation occurred. Alternatively, other researchers have proposed that the area around the angular gyrus is not involved in simple arithmetic, but rather is critical to the execution of calculation procedures such as borrowing and carrying operations [19]. This account would potentially explain the activation in that area in the Rueckert et al. serial sevens task, along with the lack of activation in that area in our study and the Dehaene et al. PET study.

Next consider the bilateral activation in the fusiform and lingual gyri. Activation was also observed in these areas by Deheane et al. using PET, although only in the left hemisphere [10]. They offered the interpretation that the fusiform gyrus is involved in processing visual number forms. However, in our verification conjunction analysis, verification was compared jointly against detect ones, which involved a string of four single-digit numbers, and against the magnitude task, which involved a string of two two-digit numbers. Thus, it is unlikely that the fusiform activation for verification in our study reflects processing of visual numbers per se. One plausible elaboration of the Dehaene et al. proposal suggested by our results is that the fusiform gyrus stores configural visual representations, such as whole multiplication facts (e.g., '4 × 7'), either in addition to or instead of more simple visual representations, such as single digit numbers.

A third and very intense area of activation for all subjects during the verification task was in an area proximal to Brodmann's area 44. Although not observed or predicted by previous work on arithmetic, this activation pattern is sensible given the information processing requirements of the task. Subjects were instructed to employ a produce-compare strategy, in which they first produced the answer, and then compared it to the candidate answer presented. Answer production most likely reflects subvocal articulation of the verbal form of the answer. One interpretation is that the Brodmann's area 44 activation reflects this processing. Alternatively, activation in this area may reflect syntactic processing required to encode and comprehend the arithmetic problem itself.

Finally, the verification task elicited clear prefrontal (Brodmann's areas 9 and 10) activation in at least seven subjects. This finding falsifies the plausible hypothesis that the activation in similar areas observed by Rueckert et al. [30] reflected solely the complex demands of the serial sevens task. Most neuropsychological evidence suggests that prefrontal areas are not directly involved in single digit multiplication. In agreement with this evidence, Dehaene et al. [10] did not observe prefrontal activation in their multiplication production task. Nevertheless, our results, combined with two reports of patients with left prefrontal damage who exhibited mild to moderate deficits with arithmetic [24,35], indicate that it would be premature to conclusively dismiss the possibility of prefrontal involvement in that task.

On the other hand, perhaps a more likely account of the prefrontal activation is that it reflects some unique property of the verification task. Two specific hypotheses warrant consideration. First, verification requires that the answer is first retrieved and then compared to the candidate answer given. The prefrontal activity may reflect this comparison process (which may involve working memory), rather than the retrieval of the multiplication fact itself. Second, the candidate answers used in the verification task were by design highly plausible answers. For example, one false candidate answer for 8 × 4 was 24. It has been clearly established in behavioral studies that highly plausible candidate answers generate interference and slower RTs relative to less plausible answers [4,32]. The prefrontal activation may reflect processing needed to overcome this interference.

As expected, there was greater left than right side activation in both the parietal area and Brodmann's area 44 for all eight subjects. In contrast, no consistent hemispheric bias was observed in the prefrontal or fusiform/lingual gyrus areas. One interesting albeit speculative account of this finding is that the parietal and Brodmanns' area 44 activation reflect a network involved in attending to, retrieving, and articulating linguistic information, which is known to have a strong left hemisphere bias. In contrast, fusiform-lingual and prefrontal areas may store and process visual representations and execute abstract executive and working memory processes, respectively. Neither of these processes would be directly related to language and it is possible that they are more symmetrically localized in the brain.

In addition to the four major areas discussed above, bilateral thalamic activity was observed for two subjects. Thalamic involvement in arithmetic is also suggested by direct electrical stimulation results [27]. However, there was the sharp contrast between these two subjects and the other six subjects, who exhibited absolutely no thalamic activity even with individual voxel level thresholds set to very low values. We have no strong hypothesis regarding these contrasting results across subjects. However, it is worth noting that the two subjects who exhibited thalamic activation were first and third in overall intensity of activation (averaged over both hemispheres and over the four major activation areas observed).

Basal ganglia have also been implicated for arithmetic in three neuropsychological reports [5,19,41]. Nevertheless, we observed no activation in these areas for any of the eight subjects. One of the subjects with thalamic activation did exhibit some bilateral activity in the caudate nucleus and putamen, but this appeared to reflect 'bleeding of activation' from the adjacent thalamus and from the lower end of the column of activation extending downward from areas proximal to Brodmann's area 44. The spatial smoothing performed prior to analysis provides a quite plausible account of this result. Activation was also not observed in this area in a recent PET study [10]. The fact that arithmetic deficits have not been identified as characteristic of Parkinson's and Huntington's patients furthers suggests that these areas may not be central to arithmetic. However, a resolution to the contrasting conclusions reached via these different approaches awaits further research.

It is important to note that although the verification task has some advantages for fMRI, it also two disadvantages relative to a pure production task. First, in the verification task subjects may in some cases bypass answer retrieval (i.e., production) in favor of other strategies like magnitude estimation [32]. As noted in the introduction, we took a number of measures to maximize use of a produce-compare strategy by our subjects. Nevertheless, these measures do not guarantee that subjects used produce-compare during scanning for all items. A second complication of the verification task is that it requires a comparison stage which is not of direct interest in this study. However, any fMRI study of arithmetic in which accuracy data is collected will likely require some type of comparison stage by means of which subjects can report their answer through a dichotomous response. Obtaining such responses is important where possible, because if accuracy cannot be determined, there is no assurance that subjects were performing the required tasks. Ultimately, however, converging evidence from tasks like verification that allow for behavioral data collection, as well as from 'pure' tasks like production that more exclusively measure the process of interest, will be needed to fully understand the localization of basic arithmetic skills. Our results contribute to this goal by revealing a set of areas activated during verification that are strikingly consistent across each of eight subjects. Some significant subset of these areas is almost surely responsible for arithmetic production itself.

The results did not conclusively identify any brain areas uniquely involved in magnitude judgment but not in verification. Some subjects exhibited reliable activation in bilateral superior parietal areas during magnitude comparison vs the detect one's task, but this activity pattern was not evident for all subjects. Other studies have also failed to conclusively isolate magnitude processing using neuroimaging [10]. We offer three candidate explanations. First, magnitude processing may be highly distributed and may elicit only very weak signal. This account is consistent with the neuropsychological finding that although selective deficits in arithmetic are fairly common, selective deficits in magnitude processing are rare at best. Second, there may be pronounced individual differences in areas involved in magnitude processing. This account would potentially explain the high variability of activated areas across subjects in our study, as well as the negative findings for magnitude comparison in the Deheane et al. PET study (i.e., individual differences may have washed out the activation in their group analysis). Third, magnitude judgment may activate a subset of areas also activated by other tasks involving digits. In particular, we suspect it may not be possible to detect areas unique to the magnitude task in direct comparison to arithmetic, which may elicit automatic magnitude processing. Even the control task may have masked some of the magnitude effect in the magnitude vs detect ones comparison.

One reasonable approach in future investigations of magnitude processing would be to use a non-numerical control task and thus eliminate the possibility that magnitude processing is present in the control stimuli. However, as noted in the introduction, a non-numerical control task is subject to the equally problematic criticism that it would differ from the magnitude judgment task not only with respect to the involvement of numerical magnitude, but also with respect to other aspects of the stimuli (i.e., digits vs non-digits). Thus, any activation obtained in number comparison relative to a non-numerical control might reflect magnitude processing itself, or other non-magnitude aspects of processing numerical stimuli. In our view, both types of designs (i.e., use of both numerical and non-numerical control tasks), and perhaps also parametric designs that manipulate the difficulty of the magnitude comparison task, will likely be needed to conclusively isolate brain areas uniquely involved in magnitude processing.

Despite the empirical limitations of our conclusions regarding magnitude judgment, the findings do converge with previous results [7,10] suggesting bilateral parietal involvement. Roughly the same inferior parietal areas were also activated by the verification task. In combination, these findings provide some preliminary support for cognitive models that assume that representations for arithmetic facts and for numerical magnitude are closely integrated, perhaps as part of a single system. This interpretation is consistent with some computational models of arithmetic which assume that arithmetic knowledge is stored (in part) within a network which employs a form of representation based solely on numerical magnitude [26,37]. However, this interpretation is far from conclusive for at least two reasons: (1) there may be unique anatomical localization for these two processes which we did not detect, perhaps because of limitations in spatial resolution, and (2) these two forms of representation may be functionally distinct yet distributed in closely overlapping networks. In general, a finding that separate areas are activated by arithmetic and magnitude tasks would constitute important evidence against a single network model, but a finding that the same areas are activated by these tasks does not provide conclusive evidence against a dual network model.

Finally, there was a striking and surprising degree of hemispheric symmetry in the activation patterns for the verification task. All subjects exhibited roughly symmetric bilateral activation in parietal, fusiform, lingual, speech-language (Brodmann's area 44), and prefrontal areas. Further, thalamic activation was bilateral in the two subjects who exhibited it. This finding appears to be inconsistent with the default assumption in much of the neuropsychological literature that many cognitive functions, including certain aspects of language and simple arithmetic, reside almost exclusively in the left hemisphere in the normal brain. The results also suggest that previous neuroimaging conclusions of unihemispheric activation may in fact reflect the following underlying pattern: (1) activation is typically hemispherically symmetrical, (2) the degree of activation may, however, be greater in one hemisphere (typically the left), and (3) statistical thresholds for detecting activation in many previous experiments may have been such that only the left hemisphere activation was detected [21]. The net result of this scenario would be interpretation of a brain response as hemispherically asymmetric when in fact it is symmetric, at least in terms of the spatial areas engaged in the task (functional asymmetry may of course still exist even if this hypothesis proves correct). This perspective suggests that caution should be used when inferring that a given brain region is activated in only one hemisphere.

To conclude, our findings, combined with those of previous neuropsychological and neuroimaging studies, clearly demonstrate that parietal areas are activated during numerical processing, and also implicate involvement of several other areas, including fusiform and lingual gyri, Brodmann's area 44, prefrontal cortex, and perhaps the thalamus. The results also shed some new light on the possible roles played by these structures. Many important questions remain to be resolved regarding both the functions of the various activated regions. However, it appears likely that converging evidence from incisive new fMRI task comparisons and neuropsychological studies will provide answers to these questions in the foreseeable future.

References

[1] Ashcraft MH, Yamashita TS, Aram DM. Mathematical performance in left and right brain-lesioned children and adolescents. Brain & Cognition 1992;19:208–52.

[2] Boller F, Grafman J. Acalculia: Historical development and current significance. Brain & Cognition 1983;2:205–23.

[3] Burbaud P, Degreze P, Lafon P, Jean-Michel F, et al. Lateralization of prefrontal activation during internal mental calculation: A functional magnetic resonance imaging study. Journal of Neurophysiology 1995;74:2194–200.

[4] Campbell JID. Production, verification, and priming of multiplication facts. Memory & Cognition 1987;15:349–64.

[5] Corbett AJ, McCusker EA, Davidson OR. Acalculia following a dominant-hemisphere subcortical infarct. Archives of Neurology 1986;43:964–6.

[6] Dehaene S. Varieties of numerical abilities. Cognition 1992;44:1–42.

[7] Dehaene S. The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. Journal of Cognitive Neuroscience 1996;8:47–68.

[8] Dehaene S, Cohen L. Two mental calculation systems: A case study of severe acalculia with preserved approximation. Neuropsychologia 1991;29:1054–74.

[9] Dehaene S, Cohen L. Towards a functional and anatomical theory of numerical processing. Mathematical Cognition 1995;1:83–120.

[10] Dehaene S, Tzourio N, Frak V, Raynaud L, Cohen L, Mehler J, Mazoyer B. Cerebral activations during number multiplication and comparison: A PET study. Neuropsychologia 1996;34:1097–106.

[11] Friston K, Williams S, Howard R, Frackowiak RSJ, Turner R. Movement related effects in fMRI time series. Magnetic Resonance Medicine 1996;35:346–55.

[12] Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. Human Brain Mapping 1994;1:214–20.

[13] Friston KJ, Holmes AP, Worsley KJ, Poline JB, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general approach. Human Brain Mapping 1995;2:189–210.

[14] Grafman J. Acalculia. In: Boller F, Grafman J, editors. Handbook of Neuropsychology, vol. 1. Holland: Elsevier, 1988. p. 415–31.

[15] Grafman J, Passafiume D, Faglioni P, Boller F. Calculation disturbances in adults with focal hemispheric damage. Cortex 1982;18:37–50.

[16] Grafman J, Rickard TC. Acalculia. In: Feinberg TE, Farah MJ, editors. Behavioral Neurology and Neuropsychology. New York: McGraw-Hill, 1996.

[17] Hecaen H, Angelergues R, Houillier S. Les varietes cliniques des acalculias au cours des lesions retrolandiques: Approchestatstique du probleme. Revue Neurologique 1961;2:85–103.

[18] Henschen SE. Klinische und anatomische Beitrag zur Pathologie des Gehirns. Stockholm: Nordiske Bokhandein, 1920.

[19] Hittair-Delazer M, Semenza C, Denes G. Concepts and facts in calculation. Brain 1994;117:715–28.

[20] Jackson M, Warrington EK. Arithmetic Skills in Patients with Unilateral Cerebral lesions. Cortex 1986;22:610–20.

[21] Jonides J, Schumacher E, Smith E, Lauber E, Awh E, Minoshima S, Koeppe R. Verbal Working Memory Load Affects Regional Brain Activation as Measured by PET. Journal of Cognitive Neuroscience 1997;9:462–75.

[22] Kosslyn SM, Koenig O, Barrett A, Cave CB, Tang J, Gabrieli JDE. Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. Journal of Experimental Psychology: Human Perception and Performance 1989;15:723–35.

[23] Levin HS, Scheller J, Rickard T, Grafman J, Martinkowski K, Winslow M, Mirvis S. Dyscalculia and Dyslexia after right hemisphere injury in infancy. Archives of Neurology 1996;53:88–96.

[24] Lucchelli F, De Renzi E. Primary dyscalculia after a medial frontal lesion in the left hemisphere. Journal of Neurology, Neurosurgery and Psychiatry 1993;56:304–7.

[25] McCloskey M, Caramazza A, Basili A. Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. Brain & Cognition 1985;4:171–96.

[26] McCloskey M, Linndemann AM. MATHNET: Preliminary results from a distributed model of arithmetic fact retrieval. In: Campbell JID, editor. The nature and origins of mathematical skills. Amsterdam: Elsevier, 1992. p. 365–409.

[27] Ojemann G. Mental arithmetic during human thalamic stimulation. Neuropsychologia 1974;27:607–17.

[28] Price CJ, Friston KJ. Cognitive Conjunction: A New Approach to Brain Activation Experiments. Neuroimage 1996;5:261–70.

[29] Rosselli M, Ardila A. Calculation deficits in patients with right and left hemisphere damage. Neuropsychologia 1989;27:607–17.

[30] Rueckert L, Lange N, Partiot A, Appollonio I, Litvan I, Le Bihan D, Grafman J. Visualizing Cortical Activation During Mental Calculation with Functional MRI. Neuroimage 1996;3:97–103.

[31] Roland PE, Friberg L. Localization of cortical areas activated by thinking. Journal of Neuropsychiology 1992;53:1219–43.

[32] Romero SG, Rickard TC, Bourne LE Jr. On verification of multiplication facts: An investigation using retrospective protocols. Manuscript submitted for publication.

[33] Sakurai Y, Momose T, Iwata M, Sasaki Y, Kanazawa I. Activation of prefrontal and posterior superior temporal areas in visual calculation. Journal of Neurological Science 1996;139:89–94.

[34] Talairach J, Tournoux P. Co-Planar Stereotaxic Atlas of the Human Brain. New York: Thieme Medical Publisher, 1988.

[35] Tohgi H, Saitoh K, Takahashi S, Takahashi H, Utsugisawa K, Yonezawa H, Hatano K, Sasaki T. Agraphia and acalculia after a left prefrontal infarction. Journal of Neurology, Neurosurgery & Psychiatry 1995;58:629–32.

[36] Turner R. Gradient coil design: A review of methods. Magnetic Resonance Imaging 1993;11:903–20.

[37] Viscuso SR, Anderson JA, Spoehr KT. Representing simple arithmetic in neural networks. In: Tiberghien G, editor. Advances in Cognitive Science. Volume 2: Theory and Applications. New York: Wiley, 1989. p. 141–64.

[38] Warrington EK. The fractionation of arithmetic skills: A single case study. Quarterly Journal of Experimental Psychology 1982;34A:31–51.

[39] Warrington EK, James M, Maciejewski C. The WAIS as a laterlising and localising diagnostic instrument: A study of 656 patients with unilateral cerebral lesions. Neuropsychologia 1986;24:223–39.

[40] Whalen J, McCloskey M, Lesser RP, Gordon B. Localizing arithmetic processes in the brain: Evidence from a transient deficit during cortical stimulation. Journal of Cognitive Neuroscience 1997;9:409–17.

[41] Whitaker H, Habiger T, Ivers R. Acalculia from a lenticular-caudate infarction. Neurology 1985;35(Supplement 1):161.

[42] Worsley KJ, Evans AC, Marrett S, Neelin P. A three-dimensional statistical analysis for rCBF activation studies in human brain. Journal of Cerebral Blood Flow Metabolism 1992;12:900–18.