MIND, BRAIN, AND EDUCATION

Brain-Anatomy Differences in the Commission of Reversal Errors during Algebraic Word Problem Solving

Noelia Ventura-Campos[1,2], Lara Ferrando Esteve[1,2], Anna Miró-Padilla[2], and César Ávila[2]

ABSTRACT— An important line of research related to the resolution of word problems is the study of the cognitive processes involved when subjects translate problems into the language of algebra. One of the most common errors in problem-solving is the reversal error (RE), which occurs when students reverse the relationship between two variables when translating equations from comparison word problems. The aim of this neuroeducational study is to investigate the brain anatomy differences between two groups, one group that commits RE and a second group that does not. Magnetic resonance images of 37 normal and healthy participants between the ages of 18–25 years were acquired. Differences in gray matter were assessed using voxel-based morphometry analysis. Our results show that the RE group has a larger volume in the putamen, suggesting that these subjects have to make a greater effort to solve problems.

A better understanding of the way the brain learns mathematics has led to a significant number of articles. Most of these studies focus on number sense, arithmetic learning, and difficulties in mathematical learning (Ansari, 2008; Butterworth, 2010; Butterworth, Varma, & Laurillard, 2011; Cantlon, Brannon, Carter, & Pelphrey, 2006; Feigenson, Dehaene, & Spelke, 2004; Zamarian, Ischebeck, & Delazer, 2009). In the specific case of problem-solving, although several behavioral studies have been carried out (Boonen, de Koning, Jolles, & van der Schoot, 2016; Clement, 1982; Lee, Ng, & Ng, 2009; Lee, Ng, Ng, & Lim, 2004; Marshall, 1995; Nathan, Kintsch, & Young, 1992; Puig, 1996; Puig & Cerdán, 1988; Riley, Greeno, & y Heller, J. L., 1983; Sweller, 1988), as far as we know, few studies have used neuroimaging techniques to find out more about the underlying neural processes or structural changes involved in mathematical problem-solving. These studies (Anderson, Fincham, Qin, & Stocco, 2008; Anderson, Betts, Ferris, & Fincham, 2012; Qin et al., 2004; Lee et al., 2007) could shed light on the transition from arithmetic to the symbolic language of algebra, where students have to develop abstract reasoning skills that allow them to generalize, model, and analyze mathematical equations and theorems.

Problem-solving is considered one of the most important components of the study of mathematical knowledge (Castro & Ruíz, 2015; Schoenfeld, 1992; Schoenfeld, 2013). According to Halmos (1980), there are several essential elements of mathematics, including theorems, demonstrations, formulas, theory, and so forth, but he suggests that the most important things in mathematics are problems and their solutions. In addition, Kleiner (1986) emphasizes that effort in solving certain problems leads to a good development of concepts and mathematical theories. The importance of problem-solving is widely accepted by researchers. As Castro (2008) states, problem-solving is not just a scientific activity; it is also a type of educational task that must be given a prominent position in the teaching and learning processes of children, adolescents, and students in general.

As part of the mathematical curriculum content in primary and secondary education, it is common to set out problems consisting of verbal statements that describe possible situations in the world (word problems). From a cognitive point of view, solving word problems involves an analytical

[1]Department of Education and Specific Didactics, Universitat Jaume I
[2]Neuropsychology and Functional Neuroimaging Group, Universitat Jaume I

Address correspondence to Lara Ferrando Esteve, Department of Education and Specific Didactics, Universitat Jaume I, Castellón, Spain; e-mail: lferrand@uji.es.

process that reduces the statement to mathematical information that identifies and connects quantities. During this analytical process, the resolver tries to match conceptual schemes that are evoked by the situation described in the problem (Marshall, 1995; Riley et al., 1983; Sweller, 1988). We draw on the idea that a mathematical problem can be considered a combination of subproblems, and we integrate it into the model of understanding and solving word arithmetic problems formulated by Kintsch and Greeno (1985) and generalized in Nathan et al. (1992) for algebra-word-problem comprehension. For each subproblem, the student must carry out a partial analysis where the activation of a conceptual scheme at the cognitive level will be attempted, in order to establish a relationship between quantities and later translate it from words to equations. Obviously, the difficulty of identifying some of the conceptual schemes can be an obstacle to defining the relationship between the quantities and, therefore, completing the resolution process. However, it is possible that, after the correct identification of the conceptual scheme, an incorrect formulation of the relationship may occur. An example of this situation would be the reversal error (RE), although not all the studies carried out agree that the creation of a problem model is necessary when committing RE.

Some of the investigations on problem-solving, such as those by Clement (1982), Clement, Lochhead, and Monk (1981), and Clement, Lochhead, and Soloway (1980), found that on some comparison problems, both additive and multiplicative, most of the students made mistakes when translating some sentences from natural language to algebraic language. The structure of the sentences these researchers used to find the error was similar to the following example: "Write an equation using the variables S and P to represent the following statement: *There are six times as many students as professors at this university*. Use S for the number of students and P for the number of professors" (Clement, 1982, p. 17). In these statements, it was observed that most of the wrong answers were in the form of $P = 6 \cdot \cdot \cdot S$, which they called RE because students reversed the order of the letters, compared to the correct answer $S = 6 \cdot \cdot \cdot P$. The initial research by Clement (1982) was followed by numerous behavioral studies on RE (Cohen & Kanim, 2005; Cooper, 1986; Fisher, 1988; González-Calero, Arnau, & Laserna-Belenguer, 2015; González-Calero, Berciano, & Arnau, 2019; Kim et al., 2014; Lopez-Real, 1995; Wollman, 1983). However, these studies did not consider the importance of human brain development in learning, which is the aim of this study.

Knowledge about the way the brain learns could have a major impact on education. Understanding the brain mechanisms underlying learning and memory, as well as the effects of genetics, environment, emotion, and age on learning, could transform educational strategies and make it possible to devise programs that optimize learning for people of all ages and with diverse needs (Blakemore & Frith, 2005). Therefore, educational neuroscience could help to better understand the relationship between biological brain development and the development of the human capacity for mathematical cognition, mediated by educational experience (Royer, 2003).

The application of non-invasive neuroimaging methods, such as Magnetic Resonance Imaging (MRI), among others, has the potential to provide knowledge about cognitive processes related to learning at a more detailed level than behavioral studies alone (Ansari, De Smedt, & Grabner, 2012). Thus, it would be interesting to see the relationship between the brain and learning in the academic field. This relationship can be seen in studies that have used MRI, which can study even the most complex mental processes, such as problem-solving (Hanakawa, Honda, Okada, Fukuyama, & Shibasaki, 2003). Radford and Andre (2009) reported that little research has focused on the brain and advanced mathematical thinking, in particular the brain-algebra relationship during development. Among these studies, the longitudinal study by Qin et al. (2004) used functional MRI to see the differences in algebraic learning between adults and teenagers. The results showed that, after practice with solving word problems, both adolescents and adults presented reduced activation in the prefrontal cortex, which is involved in mathematical cognition and other higher-order processes that develop throughout childhood and adolescence. Moreover, a reduction in the activation of the left parietal cortex, which holds an image of the equation, and an increase in the left putamen were observed only in adolescents. These results suggest that the brain response in adolescents is more plastic in this neural stage of development and, therefore, undergoes more changes due to practice and learning effects. In conclusion, this increased brain response in adolescents due to practice seems to indicate that this period would be the most appropriate one for learning algebra (Qin et al., 2004). These observations would be consistent with neural changes due to development explored in several brain imaging studies (Blakemore, 2012; Giedd et al., 1999; Sowell et al., 1999, 2003; Sowell & Jernigan, 1998; Wierenga et al., 2014).

These MRI studies of children during their development have shown that, during the period of adolescence, the white matter (WM) volume continues to increase (there are even some local areas that change rapidly), and so the gray matter (GM) in some areas begins to lose volume (Sowell et al., 2003; Sowell, Thompson, Holmes, Batth, et al., 1999). In addition, Blakemore's (2012) extensive review of MRI studies performed during development provides more information about brain maturation. Blakemore (2012) states that brain development does not end at an early age, but rather it extends to adolescence, reaching its maximum volume around the age of 25 (Caviness Jr., Kennedy, Bates, &

Makris, 1996; Sowell & Jernigan, 1998). Changes take place in both the GM and the WM, and the WM has the greatest development during this stage (Ortiz, 2009). In addition, the study by Wierenga et al. (2014) reveals that the GM volume in the striatum (dorsal: the putamen, caudate; and ventral: nucleus accumbens) decreases linearly with age from seven to 23 years old (see Figure 2 from Wierenga et al., 2014). The development of the striatum structures indicates that there are dynamic changes during development that are related to changes in cognitive development, experience, and behavior.

This literature on the areas of algebra learning and brain maturation highlights the importance of introducing algebra at the age when the brain is ready. The brain is prepared when the structures related to learning are in the process of maturation and there is greater brain plasticity, which supports learning, as Qin et al. (2004) indicated. In this period, students are prepared to absorb and assimilate the new mathematical knowledge. In this regard, it seems necessary to obtain more knowledge about the anatomical and functional neural bases of academic competences through neuroeducation studies, in order to better structure learning environments that favor the acquisition of these competences, given that they are crucial in modern society (Ansari et al., 2012).

The goal of this study is to investigate the anatomical brain differences among subjects while performing a problem-solving task with associated RE, thus exploring the effects of learning problem-solving skills and brain maturation. To achieve this objective, we will observe the differences in GM volume between participants who commit RE and those who do not (non-RE), using the Voxel-Based Morphometry (VBM) method.

MATERIALS AND METHODS

The data (structural MRI) (Ventura-Campos, Ferrando, Miró-Padilla, & Ávila, 2022) are available for reproducibility purposes.

Participants

In this study, participants were 37 students at the *University Jaume I* with ages ranging from 18 to 26 years. Informed consent was obtained from each participant before the study. For structural analysis, we used 33 participants divided into two groups: the non-RE group (18 subjects, 9 female; mean age: 22.5, SD: 2.53), who responded correctly 100%, and the RE group (15 subjects, 11 female, mean age: 21.466, SD: 2.1), who failed on more than 50% of the answers on the task. The other four participants were excluded because they did not perform the task correctly.

The study exclusion criteria were the presence of neurological and medical illness, trauma with loss of consciousness lasting more than 1 hr, and the typical resonance exclusion criteria such as iron prosthesis and dental implants.

Informed Consent

All participants received remuneration for completing the study. The Ethical Committee of Universitat Jaume I approved the research project. All participants gave informed written consent prior to participation.

Experimental Paradigm

First, the structural image of each subject was obtained with the MRI. Second, outside of the MRI, to obtain the classification of subjects who committed RE versus non-RE, we used an application similar to González-Calero et al. (2015), where the participants had to build a mathematical equation for each of the statements presented to them.

The task contained a total of 16 statements ($2 \times 2 \times 2 \times 2$) in Spanish, focusing on the following: (a) whether the comparisons are multiplicative or additive; (b) whether the comparisons are increasing (times more than, more than) or decreasing (times less than, less than); (c) contextual or non-contextual clues; and (d) discrete or continuous amounts. Thus, subjects could use only multiplication/division or addition/subtraction to express the equation. To construct the equation, we made it easier by giving them the variables and the amounts they had to use (see Figure 1). By clicking on each of the variables and operation signs, they built the equation and later validated it. It should be noted that the response time was unlimited.

Neuroimaging Data Acquisition

Structural MRI data were acquired using a 1.5T Siemens Symphony scanner and a 3T Philips scanner. The parameters of the MPRAGE sequence were as follows: (a) The Siemens Symphony scanner used a high-resolution T1-weighted, repetition time = 2,200 ms, echo time = 3 ms, flip angle = 90°, matrix size = $256 \times 256 \times 160$, voxel size = $1 \times 1 \times 1$ mm; and (b) The Philips scanner used a high-resolution T1-weighted, repetition time = 8.4 ms, echo time = 3.8 ms, matrix size = $320 \times 320 \times 250$, voxel size = $0.75 \times 0.5 \times 0.8$ mm. All the scanner acquisitions were performed in parallel to the anterior commissure-posterior commissure plane (AC-PC), and they covered the entire brain. Participants were placed in a supine position in the MRI scanner. Their heads were immobilized with cushions to reduce motion degradation, and they were asked to minimize their head movement.

Behavioral Analysis

The responses collected from the app were classified as *correct*, *reversal error*, and *other errors*. That is, using the

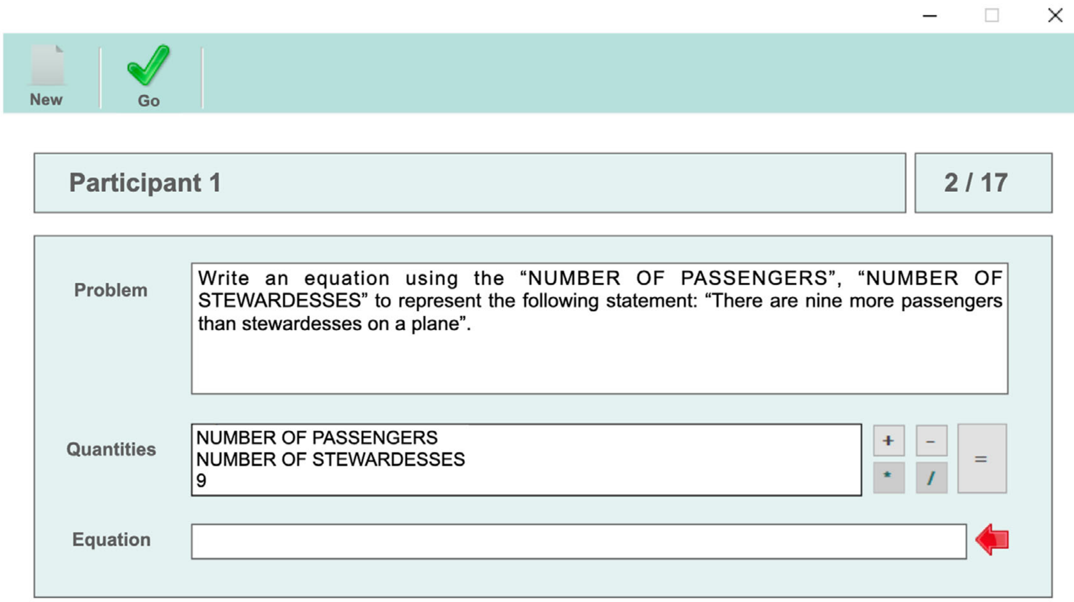

Fig. 1. Application implemented for data collection in studies on the RE (González-Calero et al., 2015). By clicking on each of the variables ("Cantidades") and with operation signs (+, −, *, / and =; right to "Cantidades"), participants had to write the equation that corresponded to the statement ("Problema": *Statement*: Write an equation using the "number of passengers," "number of stewardesses" to represent the following statement: "There are nine more passengers than stewardesses on a plane"). The rectangle called "Ecuación" contains the variables they clicked. After that, they validated their equation. An equation without RE would be "number of passengers = 9 * number of stewardesses" or "number of stewardesses = number of passengers/9"; and an equation with RE would be "number of passengers * 9 = number of stewardesses" or "number of passengers = number of stewardesses/9." Any other equation would be considered another type of error, and the participant would be excluded from the sample.

example in Figure 1, an equation that could be reduced to S = P * 9 or P = S/9 would be classified as a reversal error. Any other equation would be considered *another error*. and the participant who made more than 25% of other errors would be excluded from the sample. In the case of additive comparisons, an analogous criterion was used.

In terms of RTs, participants' performance was processed with the IBM SPSS Statistics software (Version 22, Armonk, NY). A two-sample t-test was conducted to show the differences in RTs between groups in building the equation using the app.

Neuroimaging Analysis: Voxel-Based Morphometry (VBM) Analysis

The VBM analysis was performed using the Computational Anatomy Toolbox 12 (CAT12.5, http://dbm.neuro.uni-jena.de/cat/), a toolbox from the software: Statistical Parametric Mapping (SPM12 [v7219], Wellcome Trust Centre for Neuroimaging, Londres, UK, http://www.fil.ion.ucl.ac.uk/spm/software/spm12). We used the standard procedure suggested by CAT12, which included: (a) normalization and segmentation of the images into GM, WM, and cerebrospinal fluid; (b) alignment of the GM and WM between the images; (c) using the DARTEL-template in the VBM analysis to obtain the GM and WM tissues, which were normalized to MNI standard space; (d) modulation by the "affine + nonlinear" components derived from spatial normalization; (e) estimation of the total intracranial volume (TIV) for each subject; (f) quality check of the images; no outliers were identified; and (g) Finally, the images were spatially smoothed using a Gaussian filter (8 mm Full-Width Half-Maximum, FWHM).

Statistical Analysis

The smoothed GM images were entered into a statistical analysis using the General Linear Model in SPM12. A two-sample t-test was performed to obtain the differences in GM volume between groups using two covariates of non-interest: (a) the TIV to correct different brain sizes; and (b) because the data were obtained from two machines, we created a covariate to remove the effect of obtaining the MRI from different machines. Finally, we defined the contrast to obtain the results of the analysis. The statistical criterion was set at $p < .05$, using Family-wise error (FWE) cluster-corrected for multiple comparisons (voxel-level uncorrected threshold of $p < .005$ with a critical cluster size).

Table 1
Task Results for the RE Group

Participant	1	2	3	4	5	6	7	8	9	10	11	12	13	14	*15*
RE	12	11	11	15	14	6	14	12	13	16	7	10	14	8	16
Correct	4	2	2	1	2	9	2	2	3	0	9	6	2	8	0
Machine	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1

Note: The "Participant" row contains the ID of each subject. The "RE" row contains the number of incorrect answers on the task. The "Correct" row is the opposite of the "RE" row. The "Machine" row is a covariate to remove the effect of getting the images from different machines, where 1 is the Philips machine and −1 is the Siemens machine. Participants 2, 3, 6, and 8 have fewer than the 16 results of the other participants because they made errors in the equations belonging to *another type of error*, but not more than 25%.

Table 2
Task Results for the Non-RE Group

Participant	16	17	18	*19*	20	21	22	23	24	25	26	27	28	29	30	31	32	33
RE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Correct	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
Machine	1	1	1	1	1	1	1	1	1	−1	−1	−1	−1	−1	−1	−1	−1	−1

Note: The "Participant" row contains the ID of each subject. The "RE" row contains the number of incorrect answers on the task. The "Correct" row is the opposite of the "RE" row. The "Machine" row is a covariate to remove the effect of getting the images from different machines, where 1 is the Philips machine and −1 is the Siemens machine.

RESULTS

Behavioral Results

The response time mean was 29.80 s (RE group = 30.39 s; non-RE group = 29.32 s). The results of the t-test showed no significant differences in time between groups. Tables 1 and 2 provide the performance on the task by the RE group and the non-RE group, respectively.

MRI Results

To study the difference in GM volume between groups, a two-sample t-test was performed. When studying the RE versus non-RE contrast, we observed an increase in GM volume in the bilateral putamen in the participants with RE. The MNI coordinates for the left putamen were $x = -23$, $y = -12$, $z = 0$, with a Z-value = 3.96 ($k = 938$), and the MNI coordinates for the right putamen were $x = 26$, $y = -8$, $z = 3$, with a Z-value = 3.41 ($k = 813$) (see Figure 2). The opposite contrast did not yield any significant differences.

DISCUSSION

The present educational neuroscience study focuses on the brain anatomy differences in young people who have performed an algebra problem-solving task that can produce the so-called RE (Clement, 1982).

The algebra problem-solving process often involves multiple cognitive components to reach a solution. Much of the relevant mathematical and pedagogical literature on problem-solving was influenced by Polya's (1945) four phases. The first phase in solving a problem is to read and understand the problem and identify the data involved. The second phase is to devise a plan, that is, to find the connection between the data and the unknown. This phase requires the skill of choosing an appropriate problem-solving heuristic. The third phase is to carry out the plan, which involves solving the problem by executing the heuristic. And the final phase, looking back, basically involves checking whether all the information was used and whether the answer makes sense.

If you look at Polya's first phase, reading comprehension is important in order to know what the problem is asking. Therefore, it is to be expected that the better the reading comprehension, the greater the problem-solving accuracy. The working memory has been shown to predict individual differences in reading comprehension (Daneman & Merikle, 1996). Furthermore, several studies have shown correlations between working memory and problem-solving, depending on reading comprehension (Fuchs et al., 2006; Swanson, Cooney, & Brock, 1993). However, in the study by Lee et al. (2004), the working memory contributed independently to problem-solving when reading comprehension was controlled. Therefore, this study suggests that the working memory cannot be attributed solely to its relationship with reading comprehension. This finding agrees with previous studies that relate working memory to mathematical problem-solving (Lee et al., 2004; Passolunghi & Pazzaglia, 2004; Swanson, 2006) and predict the performance on a problem-solving task in fifth grade (Lee et al., 2009). The findings by Lee et al. (2009) showed that the working memory plays an important role in text decoding and in constructing a schematic representation of the problem. Moreover, this study suggests that working memory

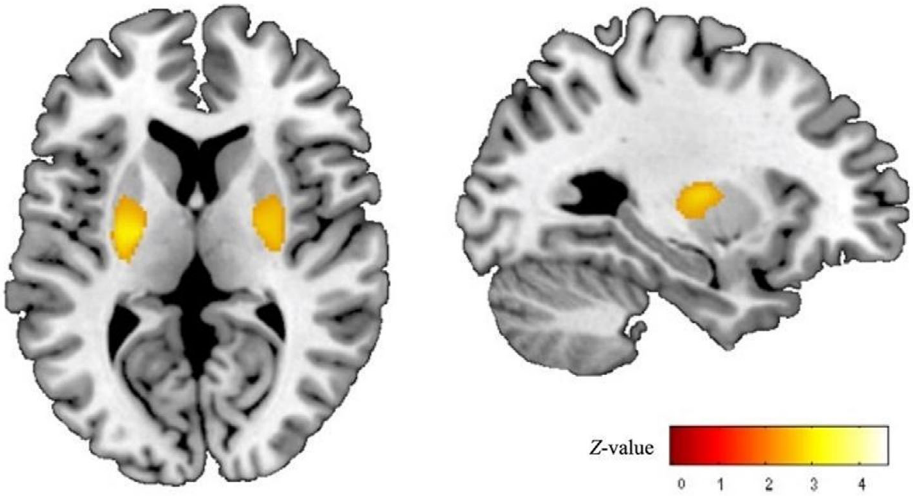

Fig. 2. Neuroimaging results of the two-sample t-test performed between groups. The figure represents the contrast: RE group versus non-RE group in our study (FWE cluster-corrected $p < .05$ for multiple comparisons at the whole-brain level, voxel $p < .005$, and a cluster size of $k = 813$ voxels). Results show a higher volume in the bilateral putamen in subjects who commit RE

was strongly associated with the selection of the mathematical operations the participants needed to compute the solution. Furthermore, some studies have claimed that a low working memory capacity would lead to a low ability to solve mathematical problems (Díaz, 2010; Ruiz, Escotto, & Sánchez, 2012). Thus, the evidence seems to suggest that the working memory is involved in the problem-solving process and, therefore, in all of Polya's phases.

In addition to working memory, other executive functioning domains may also contribute to performance on algebraic problems. Switching is an executive functioning domain that is often examined in the mathematics literature (Clements, Sarama, & Germeroth, 2016). To complete Polya's second phase, the information extracted from a sentence in the algebraic problem has to be integrated with pre-existing knowledge to form a conceptual problem model (Kintsch, 1998). Hence, switching may be important in the representation of the problem and, therefore, in the ability to move between alternative sets of mental operations to choose an appropriate heuristic to solve the problem. This executive function is also a feature that will play an important role in Polya's third phase, where the subjects switch from a problem-solving heuristic to letter-symbolic algebra. Finally, to solve the algebraic problem, it is necessary to have good mathematical skills and knowledge of the mathematical concepts of the problem.

In order to study the brain anatomy of the subjects who performed the algebra problem-solving task with RE, we compared the GM brain volume of the RE group and the volume of the non-RE group using the VBM method. An increase in the GM volume in the bilateral putamen was found in the RE group, compared to the non-RE group. This follows along the lines of the study by Qin et al. (2004), where both adolescents and adults performed a word problem-solving task; the results showed that, after performing a practice session, only the adolescents showed increased activation in the left putamen ($x = -26$, $y = -9$, $z = 5$). The authors suggested that the adolescents may have required more effort in performing complex calculations, thus making use of brain regions that are not necessary for adult activity. Furthermore, they also suggest that this might be related to the striatal region changes observed in adolescents during maturation (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999).

Moreover, in different fMRI studies (Amalric & Dehaene, 2016; Hsu & Goh, 2016), a bilateral activation of the putamen was found after performing mathematical tasks, which leads us to think that the putamen is an important structure in performing mathematical tasks. The study by Arsalidou and Taylor (2011) showed the significant contributions of the left putamen on number tasks. They suggested that the putamen's involvement in number tasks was related to integrating information by pacing the coordination of top-down and bottom-up items. They also proposed that the putamen may play a role in assigning priority values or sequencing information that needs to be processed during number tasks.

Furthermore, regarding the putamen's association with executive functions, which can help to perform algebraic problems, Constantinidis and Klingberg (2016) carried out a review investigating activation during working memory tasks. The authors indicated that, although the putamen (included in the dorsal striatum) has no main function in working memory, activity has been found in this structure, indicating its participation during this cognitive process, with the striatum being one of the brain structures related to working memory. Moreover, the putamen is associated with anticipations of high versus low working memory updating (Yu, FitzGerald, & Friston, 2013). Therefore, the evidence seems to indicate that the putamen also plays a relevant role in working memory.

In addition, investigations on the role of the putamen in executive functions suggest that the striatum also appears to be able to perform switching. Sowell, Thompson, Holmes, Jernigan, and Toga (1999) indicated that the putamen has been involved in the cognitive function of task switching, as in learning of stimulus–response associations (Packard & Knowlton, 2002), which is linked to frontal system function (Rolls, 1994) and improves throughout adolescence (Levin et al., 1991). This suggests that normal brain development has temporal and functional relationships between simultaneous post-adolescent reductions in GM density in frontal and striatal regions (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). Thus, Sowell, Thompson, Holmes, Jernigan, and Toga (1999) highlighted the potential importance of frontal/striatal maturation in adult cognition in achieving a temporal and spatial progression of post-adolescent maturation into the frontal lobes. In addition, neuroimaging evidence suggests that the fronto-striatal network is activated when subjects have to choose from many possible responses (Desmond, Gabrieli, & Glover, 1998), and so the striatum may play a role in selecting between alternative responses. Thus, this ability to readily switch between stimulus–response would seem to be a necessary component of many complex executive tasks, such as planning, problem-solving, and strategizing. Although these executive functions rely on the prefrontal cortex, the task-switching functions of the striatum may make a critical contribution to executive abilities via the fronto-striatal network (Packard & Knowlton, 2002).

Finally, we understand that the effects of low math skills and problem-solving difficulties in post-adolescent subjects may be due to the type of learning obtained at school during brain development. Therefore, to find out more about the brain development of the putamen during learning, this structure must be examined in detail in preadolescent children. The structural study by Sandman et al. (2014) with 50 preadolescent children using three different methods of analysis, including the VBM method, was conducted to determine and locate areas where the GM volume was associated with poor cognitive performance. The participants were assessed with the WISC-IV test and declarative memory, motor, and executive functioning tasks. Their results showed that the larger the GM volume in the bilateral putamen, the worse the cognitive performance, especially on working memory tasks. Sandman et al. (2014) concluded that larger GM volume in the putamen was associated with impaired cognitive function in typically developing young children.

This inverse relationship between GM volume and learning might be explained from a developmental perspective. According to Kanai and Rees (2011), the reduction in GM volume is thought to reflect pruning, as a process of removing inefficient synapses and neurons underlying brain maturation. Thus, a smaller GM volume may be a consequence of synaptic pruning, which leads to more efficient processing. This idea is supported by Hartzell et al. (2016). They studied a group of professional Vedic Sanskrit Pandits in India who trained from childhood for about 10 years in an ancient, formalized tradition of oral Sanskrit text memorization and recitation, mastering the exact pronunciation and invariant content of multiple 40.000–100.000 word oral texts. The mean age of the participants was 22 years, and the control group consisted of members of India's National Brain Research Center community or students from a nearby technical college. The authors found that Pandits showed smaller GM volume than controls in subcortical regions, including the putamen. Although these results were unexpected, they proposed the explanation that they indicate faster maturation of these subcortical regions in Pandits, based on the study by Wierenga et al. (2014). However, we can now add the assumption that greater learning leads to a smaller GM volume in the putamen.

Limitations of the Current Study

This study has a few limitations. We did not collect behavioral tests related to executive functions involved in the problem-solving process to control the differences between groups. Future studies should replicate these results in bigger samples and use behavioral tests related to working memory, IQ, reading comprehension, problem-solving, and decision making.

In any case, we think it is relevant that this is the first study to investigate brain differences in the commission of reversal errors during algebraic word problem-solving.

CONCLUSIONS

Based on our results showing a higher GM volume in the bilateral putamen in the RE group, we conclude that the RE group requires a large executive function capacity to solve an algebra problem, with the putamen playing a role in this process during adolescence. We tentatively propose that subjects who commit RE have a greater demand for the putamen in algebra problem-solving during their learning and development. They need to make a greater effort on problem-solving, resulting in the putamen not maturing properly and, thus, not decreasing its GM volume as it should. Overall, our results demonstrate a strong effect of brain maturation on algebra learning and problem-solving during development.

To our knowledge, this is the first brain anatomy report showing the structure that supports the learning of problem-solving and algebraic competence in the phenomenon of reversal error. Moreover, we have studied the use of SPHARM representation (local shape analysis) of the bilateral putamen for classification methods (Ferrando, Ventura-Campos, & Epifanio, 2020). This study showed that the shape of the putamen is a good biomarker to identify competent resolvers.

This result provides us with a first view of the anatomical differences between people who make reversal errors and those who do not. This gives us the opportunity to propose interventions to help solve this type of problem correctly and observe whether these structural differences in the putamen are modified by training.

In future work, we will continue with this educational neuroscience study by collecting data from participants through functional MRI while they perform a word problem-solving task with associated RE inside the machine. Our aim is to investigate the underlying neural basis of RE.

AUTHOR CONTRIBUTIONS

Noelia Ventura-Campos devised the project and the main conceptual ideas. Noelia Ventura-Campos, Lara Ferrando Esteve and A. Miró-Padilla contributed to the design and implementation of the research and to the analysis of the results. Noelia Ventura-Campos and Lara Ferrando Esteve wrote the paper with input from all authors. C. Ávila aided in interpreting the results and worked on the manuscript.

Acknowledgments—Project PID2020-118763GA-100 financed by MCIN/AEI/10.13039/501100011033 and Project UJI-A2021-04 financed by Universitat Jaume I Grant. Noelia Ventura-Campos was supported by the Requalification program grant REQ/2021/02(UP2021-021) financed by the European Union - NextGenerationEU. Lara Ferrando Esteve was supported by the Margarita Salas program grant MGS/2021/26 financed by the European Union - NextGenerationEU.

Conflict of interest

The authors declare that they have no conflicts of interest.

REFERENCES

Amalric, M., & Dehaene, S. (2016). Origins of the brain networks for advanced mathematics in expert mathematicians. *PNAS*, *113*(18), 4909–4491. https://doi.org/10.1073/pnas.1603205113

Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2012). Tracking children's mental states while solving algebra equations. *Human Brain Mapping*, *33*(11), 2650–2665. https://doi.org/10.1002/hbm.21391

Anderson, J. R., Fincham, J. M., Qin, Y., & Stocco, A. (2008). A central circuit of the mind. *Trends in Cognitive Sciences*, *12*, 136–143. https://doi.org/10.1016/j.tics.2008.01.006

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, *9*, 278–291. https://doi.org/10.1038/nrn2334

Ansari, D., De Smedt, B., & Grabner, R. H. (2012). Neuroeducation – A critical overview of an emerging field. *Neuroethics*, *5*(2), 105–117.

Arsalidou, M., & Taylor, M. J. (2011). Is 2+2=4? Meta-analyses of brain areas needed for numbers and calculations. *NeuroImage*, *54*, 2382–2393. https://doi.org/10.1016/j.neuroimage.2010.10.009

Blakemore, S. J. (2012). Imaging brain development: The adolescent brain. *NeuroImage*, *61*(2), 397–406. https://doi.org/10.1016/j.neuroimage.2011.11.080

Blakemore, S. J., & Frith, U. (2005) *The learning brain: Lessons for education*. Oxford: Wiley-Blackwell.

Boonen, A. J., de Koning, B. B., Jolles, J., & van der Schoot, M. (2016). Word problem solving in contemporary math education: A plea for reading comprehension skills training. *Frontiers in Psychology*, *7*, 191. https://doi.org/10.3389/fpsyg.2016.00191

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Science*, *14*, 534–541. https://doi.org/10.1016/j.tics.2010.09.007

Butterworth, B., Varma, S., & Laurillard, D. (2011). Dyscalculia: From brain to education. *Science*, *332*, 1049–1053. https://doi.org/10.1126/science.1201536

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biology*, *4*, e125. https://doi.org/10.1371/journal.pbio.0040125

Castro, E. (2008). Investigación en Educación Matemática XII. In R. Luengo, B. Gómez, M. Camacho & L. J. Blanco (Eds.), *Resolución de problemas. Ideas, tendencias e influencias*. (pp. 113–140). Badajoz: SEIEM. España.

Castro, E., & Ruíz, J. F. (2015). Matemáticas y resolución de problemas. In P. Flores, & L. Rico (Eds.), *Enseñanza y Aprendizaje de las Matemáticas en Educación Primaria*. (pp. 89–107). Madrid: Pirámide.

Caviness, V. S., Jr., Kennedy, D. N., Bates, J. F., & Makris, N. (1996). The developing human brain: A morphometric profile. In R. W. Thatcher, G. R. Lyon, J. Rumsey & N. Krasnegor (Eds.), *Developmental neuroimaging: Mapping the development of brain and behavior*. (pp. 3–14). New York: Academic Press.

Clement, J. (1982). Algebra word problem solutions: Thought processes underlying a common misconception. *Journal for Research in Mathematics Education*, *13*(1), 16–30. https://doi.org/10.2307/748434

Clement, J., Lochhead, J., & Monk, G. S. (1981). Translation difficulties in learning mathematics. *The American Mathematical Monthly*, *88*, 286–290. https://doi.org/10.2307/2320560

Clement, J., Lochhead, J., & Soloway, E. (1980). Positive effects of computer programming on students' understanding of variables and equations. In *Proceedings of the ACM 1980 annual conference*. (pp. 467–474). Nashville, TN: ACM.

Clements, D. H., Sarama, J., & Germeroth, C. (2016). Learning executive function and early mathematics: Directions of causal relations. *Early Childhood Research Quarterly*, *36*, 79–90. https://doi.org/10.1016/j.ecresq.2015.12.009

Cohen, E., & Kanim, S. E. (2005). Factors influencing the algebra "reversal error". *American Journal of Physics*, *73*(11), 1072–1078. https://doi.org/10.1119/1.2063048

Constantinidis, C., & Klingberg, T. (2016). The neuroscience of working memory capacity and training. *Nature Reviews Neuroscience*, *17*(7), 438–449. https://doi.org/10.1038/nrn.2016.43

Cooper, M. (1986). The dependence of multiplicative reversal on equation format. *Journal of Mathematical Behavior*, *5*(2), 115–120.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin and Review*, *3*, 422–433. https://doi.org/10.3758/BF03214546

Desmond, J. E., Gabrieli, J. D. E., & Glover, G. H. (1998). Dissociation of frontal and cerebellar activity in a cognitive task: Evidence for a distinction between selection and search. *NeuroImage*, *7*, 368–378. https://doi.org/10.1006/nimg.1998.0340

Díaz, R. (2010). La memoria de trabajo y su relación con habilidad numérica y el rendimiento en el cálculo aritmético elemental [Thesis, Universidad Pedagógica Nacional Francisco Morazán, Honduras]. Retrieved from https://www.cervantesvirtual.com/nd/ark:/59851/bmcv71b1

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, *8*, 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Ferrando, L., Ventura-Campos, N., & Epifanio, I. (2020). Detecting and visualizing differences in brain structures with SPHARM and functional data analysis. *NeuroImage*, *222*, 117209. https://doi.org/10.1016/j.neuroimage.2020.117209

Fisher, K. M. (1988). The students-and-professors problem revisited. *Journal for Research in Mathematics Education*, *19*(3), 260–262. https://doi.org/10.2307/749069

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., … Fletcher, J. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*, 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Giedd, J., Blumenthal, J., Jeffries, N., Castellanos, F., Liu, H., Zijdenbos, A., … Rapoport, J. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, *2*, 861–863. https://doi.org/10.1038/13158

González-Calero, J. A., Arnau, D., & Laserna-Belenguer, B. (2015). Influence of additive and multiplicative structure and direction of comparison on the reversal error. *Educational Studies in Mathematics*, *89*(1), 133–147. https://doi.org/10.1007/s10649-015-9596-0

González-Calero, J. A., Berciano, A., & Arnau, D. (2019). The role of language on the reversal error. A study with bilingual Basque-Spanish students. *Mathematical Thinking and Learning*, *22*(3), 214–232. https://doi.org/10.1080/10986065.2020.1681100

Halmos, P. R. (1980). The heart of mathematics. *The American Mathematical Monthly*, *87*(7), 519. https://doi.org/10.2307/2321415

Hanakawa, T., Honda, M., Okada, T., Fukuyama, H., & Shibasaki, H. (2003). Neural correlates underlying mental calculation in abacus experts: A functional magnetic resonance imaging study. *NeuroImage*, *19*(2), 296–307. https://doi.org/10.1016/S1053-8119(03)00050-8

Hartzell, J. F., Davis, B., Melcher, D., Miceli, G., Jovicich, J., Nath, T., … Hasson, U. (2016). Brains of verbal memory specialists show anatomical differences in language, memory and visual systems. *NeuroImage*, *131*, 181–192. https://doi.org/10.1016/j.neuroimage.2015.07.027

Hsu, C. W., & Goh, J. O. (2016). Distinct and overlapping brain areas engaged during value-based, mathematical, and emotional decision processing. *Frontiers in Human Neuroscience*, *10*, 275. https://doi.org/10.3389/fnhum.2016.00275

Kanai, R., & Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience*, *12*, 231–242. https://doi.org/10.1038/nrn3000

Kim, S. H., Phang, D., An, T., Yi, J. S., Kenny, R., & Uhan, N. A. (2014). POETIC: Interactive solutions to alleviate the reversal error in student-professor type problems. *International Journal of Human Computer Studies*, *72*(1), 12–22. https://doi.org/10.1016/j.ijhcs.2013.09.010

Kintsch, W. (1998) *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*(1), 109–129.

Kleiner, I. (1986). Famous problems in mathematics: An outline of a course. *For the Learning of Mathematics*, *6*(1), 31–38.

Lee, K., Ng, E. L., & Ng, S. F. (2009). The contributions of working memory and executive functioning to problem representation and solution generation in algebraic word problems. *Journal of Educational Psychology*, *101*(2), 373–387. https://doi.org/10.1037/a0013843

Lee, K., Ng, S. F., Ng, E. L., & Lim, Z. Y. (2004). Working memory and literacy as predictors of performance on algebraic word problems. *Journal of Experimental Child Psychology*, *89*, 140–158. https://doi.org/10.1016/j.jecp.2004.07.001

Lee, K., Lim, Z. Y., Yeong, S. H., Ng, S. F., Venkatraman, V., & Chee, M. W. (2007). Strategic differences in algebraic problem solving: neuroanatomical correlates. *Brain Research*, (1155), 163–171.

Levin, H. S., Culhane, K. A., Hartmann, J., Evankovich, K., Mattson, A. J., Harward, H., … Fletcher, J. M. (1991). Developmental changes in performance on tests of purported frontal lobe functioning. *Developmental Neuropsychology*, *7*(3), 377–395. https://doi.org/10.1080/87565649109540499

Lopez-Real, F. (1995). How important is the reversal error in algebra? In S. Flavel et al. (Eds.), *GALTHA. Proceedings of the 18th annual conference of the mathematics education research Group of Australasia*. (pp. 390–396). Darwin, Australia: MERGA.

Marshall, S. P. (1995) *Schemas in problem solving*. New York, NY: Cambridge University Press.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implications for the design of computer learning environments. *Cognition and Instruction*, *9*(4), 329–389.

Ortiz, T. (2009) *Neurociencia y educación*. Madrid, España: Alianza Editorial.

Packard, M. G., & Knowlton, P. J. (2002). Learning and memory functions of the basal ganglia. *Annual Review of Neuroscience*, *25*, 563–593. https://doi.org/10.1146/annurev.neuro.25.112701.142937

Passolunghi, M. C., & Pazzaglia, F. (2004). Individual differences in memory updating in relation to arithmetic problem solving. *Learning and Individual Differences*, *14*, 219–230. https://doi.org/10.1016/j.lindif.2004.03.001

Polya, G. (1945) *How to solve it*. Princeton, NJ: Princeton University Press.

Puig, L. (1996) *Elementos de resolución de problemas*. Comares: Granada.

Puig, L., & Cerdán, F. (1988) *Problemas aritméticos escolares*. Madrid: Síntesis.

Qin, Y., Carter, C. S., Silk, E. M., Stenger, A., Fissell, K., Goode, A., & Anderson, J. R. (2004). The change of the brain activation patterns as children learn algebra equation solving. *PNAS*, *101*, 5686–5691. https://doi.org/10.1073/pnas.0401227101

Radford, L., & Andre, É. (2009). Cerebro, cognición y Matemáticas. *Revista Latinoamericana de Investigación En Matemática Educativa*, *12*(2), 215–250.

Riley, M. S., Greeno, J. G., & Heller, J. L. (1983). Development of children's problem solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking*. (pp. 153–196). New York, NY: Academic Press.

Rolls, E. T. (1994). Neurophysiology and cognitive functions of the striatum. *Revue Neurologique*, *150*(8–9), 648–660.

Royer, J. M. (2003) *Mathematical cognition*. Portland, OR: Book News, Inc.

Ruiz, M. G., Escotto, E. A., & Sánchez, J. G. (2012). Memoria de Trabajo y Resolución de Problemas matemáticos. *Revista Electrónica de Psicología de la FES Zaragoza-UNAM*, *2*(2), 43–51.

Sandman, C. A., Head, K., Muftuler, L. T., Su, L., Buss, C., & Davis, E. P. (2014). Shape of the basal ganglia in preadolescent children is associated with cognitive performance. *NeuroImage*, *99*, 93–102. https://doi.org/10.1016/j.neuroimage.2014.05.020

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. (pp. 334–370). New York: Macmillan Publishing Co, Inc.

Schoenfeld, A. H. (2013). Reflections on problem solving theory and practice. *The Mathematics Enthusiast*, *10*(1), 9–34. https://doi.org/10.54870/1551-3440.1258

Sowell, E. R., & Jernigan, T. L. (1998). Further MRI evidence of late brain maturation: Limbic volume increases and changing asymmetries during childhood and adolescence. *Developmental Neuropsychology*, *14*(4), 599–617. https://doi.org/10.1080/87565649809540731

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, *6*(3), 491–315. https://doi.org/10.1038/nn1008

Sowell, E. R., Thompson, P. M., Holmes, C. J., Batth, R., Jernigan, T. L., & Toga, A. W. (1999). Localizing age-related changes in brain structure between childhood and adolescence using statistical parametric mapping. *NeuroImage*, *9*(6), 587–597. https://doi.org/10.1006/nimg.1999.0436

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, *2*, 859–861. https://doi.org/10.1038/13154

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*(2), 265–281. https://doi.org/10.1037/0022-0663.98.2.265

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children's word problem solution. *Journal of Experimental Child Psychology*, *55*, 374–395. https://doi.org/10.1006/jecp.1993.1021

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Ventura-Campos, N., Ferrando, L., Miró-Padilla, A., & Ávila, C. (2022). Brain-anatomy images data set of problem solving associated with reversal error: Volumetric data. Data in brief (submitted).

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67–72. https://doi.org/10.1016/j.neuroimage.2014.03.072

Wollman, W. (1983). Determining the sources of error in a translation from sentence to equation. *Journal for Research in Mathematics Education*, *14*(3), 169–181. https://doi.org/10.2307/748380

Yu, Y., FitzGerald, T. H., & Friston, K. J. (2013). Working memory and anticipatory set modulate midbrain and putamen activity. *The Journal of Neuroscience*, *33*(35), 14040–14047. https://doi.org/10.1523/JNEUROSCI.1176-13.2013

Zamarian, L., Ischebeck, A., & Delazer, M. (2009). Neuroscience of learning arithmetic-evidence from brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *33*(6), 909–925. https://doi.org/10.1016/j.neubiorev.2009.03.005