**UNIVERSITY OF TRENTO**

*Doctoral School in Psychological and Education Science*

# A study on the representation of the arithmetic facts memory: cognitively speaking, is the commutativity a property of multiplications and additions?

**DANIELE DIDINO**

Advisor: Prof. Francesco Vespignani, University of Trento

Jury: Prof. Massimo Warglien, Università Ca'Foscari di Venezia
Prof. Cristina Cacciari, Università di Modena and Reggio Emilia
Prof. Konstantinos Priftis, Università di Padova

**2011**

## Table of Contents

# Chapter 1

## General Introduction

In this thesis we deal with the organization of the memory that encodes the knowledge about two arithmetic operations: multiplication and addition. The interest about how arithmetic operations are performed takes shape in the early history of experimental psychology. For example, in the "prehistory" of the experimental psychology, Francis Galton studied if "arithmetic may be performed by the sole medium of imaginary smells" (Galton, 1894). In the early $19^{th}$ century, most of the studies on arithmetic were performed in an education context, aimed to improve the performance of children that have to learn how to solve arithmetic problems (for an historical review see Zbrodoff & Logan, 2005). More recently, over the last decades, the researches mainly addressed the nature of the representation of the arithmetic knowledge and the cognitive architecture that allows its use. The arithmetic knowledge about the simple one-digit addition and multiplication problems (e.g., 3×7, 9×8, 4+7, 8+2, etc.) are considered to be encoded in memory structures known as "arithmetic facts" (Campbell & Epp, 2005). An arithmetic fact is the memory representation of a multiplication or addition problem, so that each problem is encoded in terms of a specific arithmetic fact. For example, the problem 7×3 is encoded as an arithmetic fact in which the operands (7 and 3) are associated trough the operation (x) with the result of the problem (21). In this thesis we use the idiomatic expression "multiplication facts" when we refer to the arithmetic facts that encode one-digit multiplication problems, and "addition facts" for the arithmetic facts that encode one-digit addition problems, whereas "arithmetic facts" is used to refer to memory that encodes the arithmetic facts without distinguishing between the operations.

The arithmetic facts memory is a fundamental component of the three main cognitive models that describe number processing. The main difference between these models are about the kind of representation adopted to store the arithmetic facts. The *abstract code model* (McCloskey, 1992; McCloskey & Macaruso, 1995; Sokol, McCloskey, Cohen, & Aliminosa, 1991) assumes that the arithmetic facts are represented and retrieved by means of an abstract amodal code. The *triple code model* (Dehaene, 1992; Dehaene & Cohen,

1995) states that the arithmetic facts are encoded and retrieved in a verbal/linguistic format. The *encoding complex model* (Campbell, 1992; 1994; Campbell and Clark, 1988; 1992) assumes that different formats (e.g., verbal, Arabic-visual, etc.) contribute to encode the arithmetic facts and that in the retrieval process this formats communicate interactively rather than additively.

The *triple code model* (Dehaene, 1992; Dehaene & Cohen, 1995) assumes that the arithmetic facts are represented in a linguistic format. This hypothesis implies that the arithmetic facts are learned by rote memory in a passive way. Obviously, during the memorization of the arithmetic facts the verbal repetition of the problem is very important. It is common learn multiplication by reciting the problems as a series of fixed expressions (phrasal frequent collocations such as book/film titles, poetry verses, idioms, proverbs, etc.) like “one time two is two”, “two times two is four”, “three times two is six”, “four times two is eight”[1], and so on. However, a sequence like “four times two is eight” is more than a simple meaningless expression in which the words “four” and “two” have to be associated with the word “eight”. These sequences of words, differently from idioms for which the meaning is conventional, includes a semantically complex literal meaning that concerns the “conceptual” arithmetic relations between the words (“times”, “is”) and the symbol (i.e., the Arabic format “1”, “2”, “3”, and so on) that represent the numbers. In the early stages of learning it is possible that the children learn multiplication as fixed expressions, like verses similarly to poetry, but this rote learning is supported also by the conceptual comprehension of the meaning of arithmetic operations. The children are taught what multiplication and addition are, which is the relation between multiplication and addition (4×2=4+4), which are their properties (e.g., commutative (4×2=2×4), associative (4×2)×3=4×(2×3), and distributive 4×(2+3)=(4×2)+(4×3)), and so on. Therefore, learning multiplication is more than a simple

[1] This is an example of the 2 table in the English language. In Italian the 2 table has the name of the table in first position: “due per uno due”, “due per due quattro”, “due per tre sei”, “due per quattro otto” (“two time one is two”, “two times two is four”, “two times three is six”, “two times four is eight ), and so on. This difference is more thoroughly discussed below in the thesis.

memorization of expressions or words/symbol association. This complex set of knowledge could contribute to organize the arithmetic facts memory, which therefore could be more than a simple verbal/linguistic storage.

According to the idea that the arithmetic facts memory is organized and shaped by the comprehension of the conceptual meaning of the operations, this thesis deals with the role that the commutative property of additions and multiplications has in the organization of the arithmetic facts memory. The commutative property states that by changing the order of the operands the result of the problem does not change. However, even if the result does not change, the two problems of a commuted pair (e.g., 7×3 and 3×7, 7+3 and 3+7) are not the same mathematical problem. It is not yet clear if in the arithmetic facts memory the commuted pairs are encoded as the same problem (in a single arithmetic fact representation) or as two different problems (in two separate representations).

In Italian 8×3 is learned as “otto per tre” (“eight *by* three”), whereas in English it is learned as “eight times three”. The meanings of the sentences in the two languages are different: “otto per tre” means “take eight and multiply it by 3” (8+8+8); “eight times three” means “take eight times the number 3” (3+3+3+3+3+3+3+3). Both languages have a way to express the multiplication in other way: in Italian “3 volte 8” that perfectly correspond to “3 times 8” can be used, as well as in English where “3 by 8” correspond to the Italian “3 per 8”. However the linguistic format that suggests the first operand is the base (“3 per 8”) is the preferred one in Italian educational system, while in English the preferred expression (“3 times 8”) suggests the second operand is the base. This linguistic difference is reflected by the name of the multiplication table. In Italian the 2 table is learned as “2 per 1”, “2 per 2”, “2 per 3”, “2 per 4”, and so on, In English the 2 table is “1 times 2”, “2 times 2”, “3 times 2”, “4 times 2”, and so on. Namely, in Italian the base is linguistically in first position (in “otto per tre” the base is “otto”) and the name of the table as well; in English the base is in second position (in “eight times three” the base is three) and the name of the table as well. Obviously, both Italian and English individuals solve the problems “otto per tre” and “eight times three” relying on the symbolic problem 8×3. However, this linguistic difference, reflected by the education system, may induce a difference in the way the symbolic equation 8x3 is semantically interpreted. Moreover the order in which the two commuted pairs are learned are different in Italian and English in terms of order of the operands. Namely, in Italy 2×9 is learned before 9×2 (because the former is in the 2 table and the latter in the 9 table), whereas in England 9×2 is learned before 2×9 (because the former is in the 2 table and the latter in the 9 table). The order in which the arithmetic problems are learned could influence the arithmetic facts memory. Since, for each commuted pair, one order of the operands (e.g., 2×9 in Italy) is learned before and then more practised than the inverse order (e.g., 9×2), the former could have an advantage. In the experiments reported in this thesis we test if such an advantage exists. Moreover, we investigate if order preferences are similar or different between Italian and English individuals.

The remainder of this chapter is organized in five sections. In the first section the models that describe the architecture of the arithmetic facts memory and that assume that both problems of each commuted pair are stored as arithmetic fact are briefly reviewed. In the second section the models that assume that only one arithmetic fact is represented for each commuted pair are introduced. In the third section empirical evidence showing that adults solve arithmetic problems not only by directly retrieving the result from memory, but also by adopting more complex procedures is illustrated. In the fourth section a review of the few studies showing that the order of the operands can affect the performance in the solution of arithmetical problems is reported. Finally, in the last section the aims of this thesis is presented.

## 1.1 MODELS ABOUT THE ARCHITECTURE OF THE ARITHMETIC FACTS MEMORY

### 1.1.1 Counting Models

The *counting models* (Groen and Parkman, 1972) constitute one of the first approach to the arithmetical cognition. These models were proposed to explain the performance of children with simple addition problems. Groen and Parkman (1972) proposed five *counting models*: *sum model*, *left model*, *right model*, *max model*, and *min model*. All these models assume a main mechanism which requires to add unit by unit the operands to a base, that can be zero or one of the two operands. The basic assumption of this models was that the children, when they do not yet memorized the arithmetic facts, use counting algorithms to solve simple addition problems. The models have been conceived to explain the size effect (i.e., the time required to solve a problem is proportional to the size of the problem; e.g., 7+8 is solved slower than 2+3). In this proposal the response times (RTs) are directly proportional to the number of increments needed to complete the counting procedure. The *sum model* states that the counting begins from zero and that the two operands are added unit by unit (e.g., 2+3=0(+1+1)(+1+1+1)=5). The RTs for this model are therefore directly related to the sum of the operands. The number of increments is in fact equal to the sum of the operands. The *left model* states that children start to count from the right operand and to add unit by unit the left operand (e.g., 2+3=3+1+1=5). The RTs in this model are related to the size of the left operand which determines the number of increments required to solve the problem. The *right model* is similar to the left model with the exception that the base is the left operand and the right operand is added (e.g., 2+3=2+1+1+1=5). In this case the RTs depend on the size of the right operand. The *max model* and the *min model* state that the base is the smaller operand (e.g., 2+3=2+1+1+1=5) and the larger operand (e.g., 2+3=3+1+1=5), respectively. Therefore, the RTs are associated with the size of the larger operand only (for the *max* *model*) or with the size of the smaller operand only (for the *min model*). The results of the empirical study of Groen and Parkman (1972) showed that the best fit of children RTs was given by the *min model*, which is the most efficient counting procedure. The Authors concluded that children naturally learn to use the most efficient procedure to solve addition problems, a two-stages procedure that requires a comparison between the sizes of the two operands prior to counting. Nevertheless, the *counting models* were challenged by two other facts (Zbrodoff & Logan, 2005; see also Parkman, 1972; Parkman & Groen, 1971). First, tie problems (problems in which the same number is repeated two times: 2+2, 3+3, 4+4, and so on) were solved faster than all the other problems with similar size (tie effect). If children adopt the counting procedures of the *min model* (e.g., 2+3=3+1+1) the time required to solve the tie problems should depend on the size of the smaller operand, but this was not the case. Groen and Parkman (1972) suggested that tie problems could be memorized by children before other problems and for this reason then were not affected by the size of the operands since these were solved by mean of a direct retrieval procedure (that is retrieving the result of the problem directly from the memory). A second observation is that the size effect was also been found in adults, who are assumed to rely mainly on retrieval procedures and not on the use of counting models. The Authors argued that sometimes adults also could use non-retrieval procedures that require more time to solve the problem. Despite the historical interest in this early attempt to describe the cognitive procedures used to solve arithmetic problems during the childhood, it is interesting to note for the present purposes that Groen and Parkman (1972) considered order and/or reordering of arithmetic problems as a central topic in numerical cognition and already introduced the idea that adults solve arithmetic problems not only by using retrieval procedures but also by adopting more complex non-retrieval procedures.

### 1.1.2 Table Search Model

Ashcraft and Battaglia (1978) proposed the *table search model* in which they delineated a cognitive architecture for the representation of addition problems in adults based on a structure similar to an array. Namely, each row and column of the array represents an operand and the result of the problem is the intersection of the row and the column representing the operands, see figure 1.1. In this model each operand activates the first node of the corresponding row or column and then the activation spreads from this node to the others, following the order of the nodes in the structure. For example, the operand 7 activates the first node (i.e., 8) and then the activation spreads from the node 8 to the node 9, from 9 to 10, from 10 to 11, and so on. The result node is the only one activated at the same time by both a row and a column. The RT required to identify the result is therefore directly proportional to the number of steps needed to reach the result node. Contrary to the *counting models* (Groen and Parkman, 1972), the *table search model* takes into account only a memory process. In fact, like the other models discussed below in this section, this model does not involve any non-retrieval procedure.

The main challenge of this model comes from the neuropsychological studies. In fact, if a node is damaged the following nodes cannot receive activation from it and then cannot be activated, and the result cannot be identified. For example, the problem 7+4 implies that the activation spreads from the node 7 to the node 11 by using the intermediate nodes (i.e., the nodes 8, 9, and 10). If the node 9 is damaged the activation cannot be spread to the nodes 10 and 11. Therefore, a patient unable to solve 7+2 (for which the result node is 9) should not be able to solve the problems 7+3, 7+4, 7+5, and so on, since the memory architecture has no way to activate the correspondent result nodes (i.e., 10, 11, 12). Neuropsychological evidence about patients that can solve a large problem even if he/she is not able to solve a smaller problem for which the result is in the same row or column (see for example Sokol, McCloskey, Cohen, and Aliminosa, 1991), are very difficult to be explained within the frame of the *table search model*. Moreover, the tie effect cannot be explained by this model, since the tie problems require a number of steps proportional to the size of the operands, and thus they could not be less time consuming than other problems with similar size.

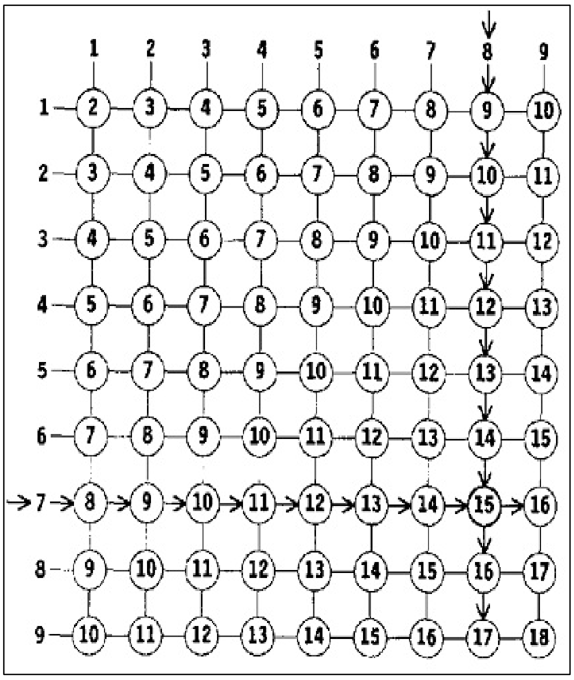

**Figure 1.1: table search model (modified by McCloskey et al., 1991).**

### 1.1.3 Network Retrieval Model

The *network retrieval model* (Ashcraft, 1987) assumes an architecture in which the arithmetic facts memory is a network with three sets of interconnected nodes (figure 12). This model has been proposed in order to describe the cognitive processing of both multiplications and additions. Each of the two operands of a problem activates the corresponding operand node in the corresponding set. For example, the problem 7×3 (or 7+3) activates the node 7 in the set of the first operands and the node 3 in the set of the second operands. The third set includes the nodes for the results of the problems. In the previous example, the problem 7×3 (7+3) activates in the result set the node 21 (10). A relevant features of this model is that a same result associated with different problems is represented by different result nodes. For example, even if both the problems 6×4 and 8×3 have as result 24, in the set of the result nodes there are two different 24 nodes, one for each problem.

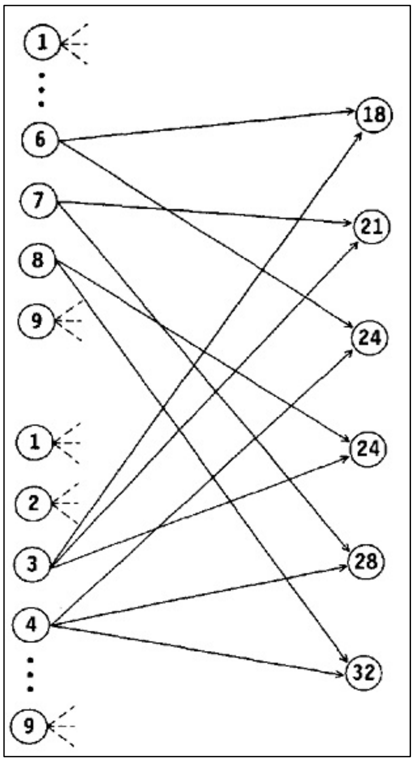


**Figure 1.2: network retrieval model (modified by McCloskey et al., 1991).**

Differently to the *table search model*, in the *network retrieval model* each operand is directly associated with all its sums or multiples without the mediation of intermediate sums or multiples. In the *table search models* the operand activates the first result node and then each result node activates the following one, whereas in the *network retrieval model* the operand activates at the same time all its sums or multiples. For example, given the problem 7+3, in the *table search model* the operand 7 activates the node 8, then the node 8 activates the node 9, and so on. In the *network retrieval model*, on the contrary, the operand 7 activates at the same time the result nodes 8, 9, 10, 11, and so on. Nevertheless, the result nodes of the *network retrieval model* are also associated one to each other, that is a result node spreads activation to all its neighbourhood (i.e., the result nodes that share an operand). For example, the problem 7×3 activates the node 21 and the latter spreads its activation to all its neighbourhood, that is the result nodes of the problems like 7×N and N×3 (i.e., 7, 14, 28, 35, etc. and 3, 6, 9, 12, etc., for the operand 7 and 3 respectively). However, a result node (e.g., 21) spreads more activation to the closer neighbourhood (e.g., 14, 28) than to the more distant ones (e.g., 35, 42). After the activation spreads from the operands to the results and between the results, the most activated result node is selected as the result of the problem. Therefore, the RT required to solve a problem depends on the strength of the association between the operands and the result nodes. Stronger is the association between operands and result, the higher is the activation of the result node and the faster the identification of the result is.

The size and the tie effects are both explained by means of the different strength of association between the operands and the results. The strength of association between the operands and the larger results is weaker than the association with the smaller result (e.g., the operand node 7 is mode associated with the result node 14 (7×2) than with 56 (7×8)). Therefore, since once presented the operands the larger results are less activated than the smaller ones, the former require longer RTs than the latter to be selected as the correct result of the problem. In the tie problems the association of the operands (e.g., 7×7, 7+7) with the result (e.g., 49, 14) is stronger than for the non-tie problems, therefore the result nodes of the tie problems are more easily activated and require shorter RTs. The fact that some operand–result associations are stronger than others is explained by means of the frequency with which the problems occur. Namely, according to the Author (Ashcraft, 1987), when the children are taught to solve addition and multiplication, in the textbook small problems are more frequent than large problems and tie problems are more frequent than non-tie problems. However, the frequency explanation encounter same difficulties if it is considered through the entire life of an adult. In fact, even though the frequency could bias the children during the first years of school, it is not clear if a different frequency across problems also occurs in everyday life for adults (see McCloskey et al., 1991).

The *network retrieval model* introduces two innovative aspects about the arithmetic facts memory. First of all this model states that the arithmetic facts memory is an highly interconnected complex memory network. Which node is selected as the result of a sum or of a product depends on the dynamic of a spreading of activation inside the memory network. Second, some elements of the architecture of the model are shaped by the learning process. The strength of the connections between the operands and the result nodes is assumed to be modulated by the frequency at which problems occur, and the connections may be reinforced by practicing the problems.

### 1.1.4 Distribution of Association Model

The *distribution of association model* (Siegler and Shrager, 1984; Siegler, 1988) states that the operands of a multiplication problem are represented together in a single representation, which is associated with different result nodes, representing the correct and the incorrect results (figure 1.3). However, the strength of the association between the problem nodes and the correct results is stronger than the association between problem nodes and the incorrect results. For example, if the problem 7×3 is presented the corresponding problem node (which include both operands: {7×3}) is activated. This problem node is associated with different result nodes (e.g., 14, 21, 28, 18, 24), but the strength of association is higher between the problem node and the correct result (e.g., 21) than between the problem node and the incorrect results (e.g., 14, 28, 18, 24). The result node (correct or incorrect) that exceeds a given threshold level is selected to be produced as the result of the problem. However, if no result nodes exceed the threshold level non-retrieval procedures are adopted to solve the problem.

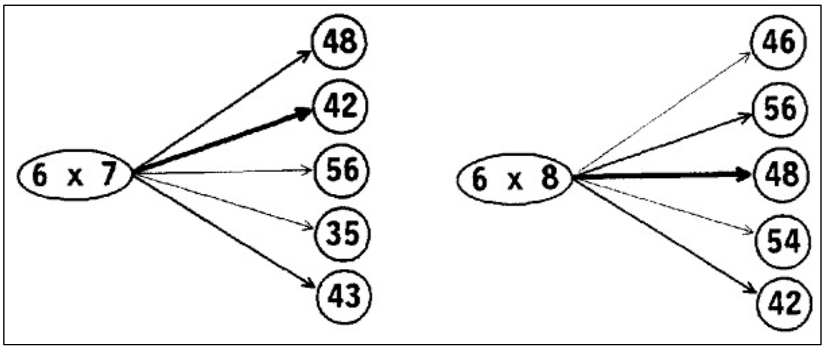


**Figure 1.3: distribution of association model (modified by McCloskey et al., 1991).**

This architecture is motivated by the assumption that children initially adopt non-retrieval procedures (e.g., repeated addition: 7×3=7+7+7) to solve multiplication problems. Each time the result is calculated by using a non-retrieval procedure the association between problem node and computed result is reinforced, even if the calculation is wrong. For example, when a child uses a repeated addition procedure (7+7+7) to solve the problem 7×3, if the calculation is correct the result 21 is associated to the problem node, whereas if he/she makes an error in the procedures then the association with a wrong result (e.g., 23 or 24) increases. When the association between the problems and the correct (and incorrect) results reaches a sufficient unbalance the problem can be solved by means of direct retrieval of the result from the memory. In adults the association between problem nodes and result nodes is assumed to be so strong that the problems are nearly exclusively solved by means of retrieval.

The size effect is produced by the different association strength that exists between problem nodes and correct/incorrect result nodes. In fact, the non-retrieval procedures are more prone to error with large problems than with small ones, and then incorrect results are produced more often with large problems than with small problems. Therefore, the association between problems and the incorrect result nodes is stronger for large problems than for smalle problems. Since the strength of the association between problem nodes and incorrect result nodes is higher for large problems, the time required to overcome a threshold and select the correct result is longer for the large problems. On the contrary, the tie effect is explained having recourse to the assumption that they are more frequent than non-tie problems and then the association between problems and correct result nodes can be more strongly reinforced by practice. However, this frequency explanation falls into the same criticism discussed above for the network retrieval model.

The distribution of association model introduces two interesting ideas. First, the model explicitly assumes that adults, when the retrieval procedure fails (i.e., no result nodes exceed the threshold level), switch to non-retrieval procedures. The non-retrieval procedures are considered to be very infrequent in adults, but if the system fails to identify a result these procedures are supposed to support the solving process. In children, instead, the retrieval fails more often and then the use of non-retrieval procedures is more common. Second, the Authors (Siegler and Shrager, 1982; Siegler, 1988) introduced the idea that the non-retrieval procedures adopted during the acquisition of the arithmetic knowledge shape the arithmetic facts memory system. In fact, unlike the *network retrieval model* which assumes that the associations between operands and result are simply shaped by the frequency of occurrence of problems, the *distribution of association model* assumes that the organization of the arithmetic facts memory is determined by the use of non-retrieval procedure during the acquisition of the problems. The strength of association between problem and (correct or incorrect) result nodes is in fact established by the outcomes of the non-retrieval procedures, that is each time a result is produced as result of a problem, the association between that result (correct or incorrect) and the problem node is reinforced.

### 1.1.5 Network Interference Model

The *network interference model* (Campbell, 1987a; 1987b; 1987c; 1995; Campbell & Clark, 1989; Campbell & Graham, 1985; Campbell & Oliphant, 1992; Graham, 1987) is a very complex architecture for multiplication, which assumes different kinds of highly interconnected representations (figure 1.4).

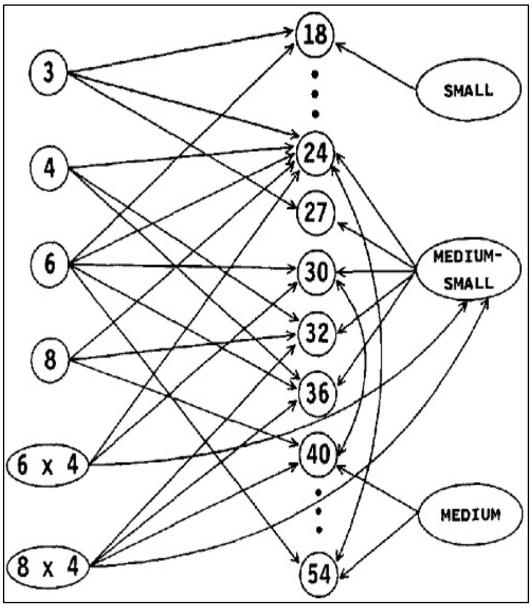


**Figure 1.4: network interference model (modified by McCloskey et al., 1991).**

Within this architecture each operand node is associated with all the result nodes representing its multiples but, unlike the *network retrieval model* (Ashcraft, 1987), the operands of different problems that produce the same result are associated with the same result node. Namely, the operands that share a multiple are connected with the same result node that represent that multiple. For example, the same result node 24 can receive activation from the operands 8 and 3 (3×8=24), and the operands 6 and 4 (6×4=24).

Therefore, if the problem 6×8 is presented, both the operand node 6 and the operand node 8 are at the same time associated with 48 (6×8=48) and 24 (6×4=24, 8×3=24). Considering only the connections between operand and result nodes, this architecture cannot distinguish between the result nodes that a result of the problem presented (e.g., the result 48 for the problem 6×8) and the result nodes that simply are a multiple of both operands (e.g., 24 is a multiple of both 6 and 8 even if it is not the result of the problem 6×8). The selection of the correct result cannot be to performed only on the basis of the activation that spreads from the operands nodes to the result nodes. However, the *network interference model* assumes that, in addition to the connections between the operands and the results, the result nodes are also connected to nodes that represent the whole problem. For example, when the problem 6×8 is presented, besides the nodes that represent each of the two operands, a node representing the whole problem ({6×8}) is also activated, and then the result node 48 receives the activation from both the two operand nodes (e.g., 6 and 8) and the whole problem node (e.g., {6×8}). Therefore, the result node 48 can be selected since it receive more activation than the result 24 that receives activation only from two nodes (the operand nodes 6 and 8). The whole problem nodes assure that the correct result receives more activation than the other nodes representing multiples of the operands but that are not the result of the problem. The *network interference model* also assumes that a problem node can be connected with both correct and incorrect results (e.g., the problem 6×8 activates the correct result 48 but can also activate other results, like 24, 42, etc.). The architecture of the *network interference model* assumes other two kinds of connections. The first kind of connections is from the whole problem nodes to magnitude nodes (representing the approximate size of the problems) and from the magnitude nodes to the result nodes that are included within that size level. For example, the presentation of the problem 6×8 activate the corresponding problem node (i.e., {6×8}) and it activates a magnitude node representing the size of the problem (e.g., {large}); whereas the problem 3×2 activates the problem node ({3×2}) and then a different magnitude node (e.g., {small}). The magnitude nodes in turn spread activation to the result nodes and contribute then in the selection of the correct result of the problem. The second kind of connections is between the result nodes themselves, like in the *network retrieval model* (Ashcraft, 1987). For example, results that share a digit (e.g., 24 and 28 share the decade digit “2”) could be associated each other and then could mutually spread activation.

The selection of the result node to select is therefore determined by the interaction of various connections and weights within this complex architecture. Namely, the presentation of a problem (e.g., 6×8) activates the operand nodes (e.g., 6 and 8) and the problem nodes (e.g., {6×8}), which in turn spread activation to the result nodes (e.g., 48, 24, etc.). Moreover, the problem nodes (e.g., {6×8}) spread activation also to the magnitude nodes (e.g., {large}), which in turn spread activation to the result nodes (e.g., 48, 54, 56, etc.). Therefore, the result nodes receive activation from operands, whole problem, and magnitude nodes. Moreover, once activated the result nodes, the activation spreads also between the result nodes (e.g., 48 spread activation to 42 because of they share the decade digit “4”), and the incorrect result nodes can interfere with and slow down the selection of the correct result.

Campbell & Graham (1985) attributed the size effect to the higher frequency of the small problems during the childhood. The frequency explanation falls however into the same criticism discussed above for the *network retrieval model*. More recently Campbell (1995) proposed another explanation of the size effect based on the activation of the magnitude nodes. Starting from the view that the representation of the numbers becomes more compressed as their magnitude increases (Dehaene, 1992), Campbell (1995) suggested that the magnitude nodes representing large problems could spread activation to more result nodes than what the magnitude nodes representing smalle problems do. Namely, the magnitude node {large} activates more result nodes than the magnitude node {small}, and then the activation of the competitors (incorrect result) and their interferences are higher for the large problems than for the small problems. The tie effect is explained by assuming that the tie problems have less competitors (incorrect result that are activated) and then less interferences from non-tie problems because the tie problems are stored in a separate storage with respect to non-tie problems (Campbell 1995; Campbell, Dowd, Frick, McCallum, & Metcalfe, 2011).

To resume, the *network interference model* states that the arithmetic facts memory is an highly interconnected system in which the result selected as correct answer depends on the integration of different connections (operands to results, whole problems to results, magnitude nodes to results, and results to other results) that spread activation each other. However, the high complexity of the architecture could be seen as a weakness of the model. In fact, since the activation level of the result node that will be selected depends on the activation spreading from many other nodes it could be theoretically and computationally difficult to determine which connections or node sets affect the performance in the various experimental tasks and procedures.

## 1.2 MODELS ASSUMING ONLY ONE ARITHMETIC FACT FOR EACH COMMUTED PAIR

The *network retrieval model*, the *distribution of association model*, and the *network interference model* described above share some common assumption about the arithmetic facts memory. First, the retrieval of the arithmetic facts is mediated by an associative network in which nodes representing the operands (that can be represented individually, like in the *network retrieval model* or in the *network interference model*, or together, like in the *distribution of association model* or in the *network interference model*) are connected and spread activation to nodes representing the results. Second, both the *network retrieval model* and the *network interference model* assume that the result nodes are interconnected. In the *network retrieval model*, the result nodes spread activation to all the other result nodes that share an operand with it (e.g., the result node 24 (=6×4) spreads activation to the result nodes of the problem 6×N (12, 18, 30, and so on) and N×4 (8, 12, 16, and so on). However, the closer the results are the higher the amount of activation spreading is (e.g., the result node 24 spreads more activation to 18, 30, 20, and 28 than to 12, 36, 16, and 32). In the *network interference model* the activation spreads between the result nodes that share some structural features, like digits in the same decade or unit position (e.g., the result node 24 spreads activation to the result 28 because of they share the digit "2" in the decade position). Third, both the *distribution of association model* and the *network interference model* assume that the problem nodes can also be associated with incorrect results (e.g., the problem 6×8 can be associated with the incorrect result 42). Fourth, all the models described above assume that the commutative property of both multiplication and addition is not used by the cognitive system to organize the representation of the arithmetic facts. In fact these models assume that for each commuted pair (e.g., 7×3 and 3×7, 7+3 and 3+7) there are two different representations, one for each order of the operands. The three models that will be described in this section integrate the commutative property within the arithmetic facts memory. Unlike the models of the previous section, these models assume that for each commuted pair there is only a single arithmetic fact encoded in memory. However, these models approach commutativity in two different way. The *identical elements model* (Rickard, 2005; Rickard & Bourne, 1996; Rickard, Healy, & Bourne, 1994) assumes that *"for commutative operations, the order of the numbers is not represented. Thus, for example, the two operand orders of a multiplication problem map on to the same unitary representation within the cognitive stage"* (Rickard & Bourne, 1996, p. 1281). The assumption that the order of the numbers (operands) is not represented means that any kind of information about the order and the position of the operands of the problems is not encoded in the arithmetic facts memory. On the contrary, the *interacting neighbors model* (Verguts & Fias, 2005) and the *COMP model* (Butterworth, Zorzi, Girelli, & Jonckheere, 2001) assume that the arithmetic facts encode information about the order of the operands. Both models assume that the arithmetic facts are represented in a format that specifies the order of the operands.

### 1.2.1 Identical Elements Model

The *identical elements model* (Rickard, 2005; Rickard & Bourne, 1996; Rickard et al., 1994; see also Campbell, 1999; Campbell, Fuchs-Lacelle, & Phenix, 2006) was proposed to explain the positive practise transfer that occurs between commutative operations (multiplication and addition), but not between non commutative operations (division and subtraction) or between commutative and non commutative operations (multiplication to division, addition to subtraction). Namely, by practicing a multiplication or an addition problem (e.g., 7×3, 7+3) there is a positive transfer on the commuted problems (e.g., 3×7, 3+7). After a training period on a problem (e.g., 7×3 or 7+3), the RTs needed to solve that problem (e.g., 7×3 or 7+3) and its commuted (e.g., 3×7 or 3+7) became faster to the same extent. Whereas, the practice on a division or subtraction problem (e.g., 21÷7, 10–7) does not have any positive transfer on the associated division or subtraction (e.g., 21÷3, 10–3). To explain this pattern of data, Rickard and Colleagues (Rickard, 2005; Rickard & Bourne, 1996; Rickard et al., 1994) proposed that for multiplication and addition there is only a single representation for each commuted pair, whereas for division and subtraction there are two different representations for the associated pair. For example, the problem 7×3 (7+3) and 3×7 (3+7) are represented as a single arithmetic fact, whereas the problem 21÷7 (10–7) and 21÷3 (10–3) are represented by two distinct arithmetic facts. The *identical elements model* does not make any assumption about the architecture of the arithmetic facts memory because it has only the aim to describe which kind of representation are encoded as arithmetic facts. For the purpose of this thesis, only the implications of the model for the commutative operation (multiplication and addition) are further discussed. For a more exhaustive description of the aspects of model that regard the non commutative operations see the original articles of Rickard and Colleagues (Rickard, 2005; Rickard & Bourne, 1996; Rickard et al., 1994).

Even if the model states that there is only one arithmetic fact for each commuted pair, it does not assume that the order of the operands is encoded within the memory system. The arithmetic facts encode only the operands, the “conceptual arithmetic operation” (that is the kind of operation that has to be performed), and the result. For example, given the problems 7×3 and 3×7, in the arithmetic facts memory there is a single representation, that is {3,7,×,21}[2]. Therefore, the model predicts that there should be no difference in the RTs for the two problems of each commuted pair (e.g., 7×3 vs 3×7). In fact, since the arithmetic facts representations do not contain any information about the order of the operands, the process that encodes the operands in the representation format used to access to the arithmetic facts should not preserve any information about the order in which the operands are presented. Any effect of the order of the operands on the performance on solving or verification tasks has thus to be explained within this model as due to non-retrieval procedures or at least as due to non-retrieval processing stages within of the whole cognitive process that allows to perform the task.

### 1.2.2 Interacting Neighbors Model

Like the *identical elements model*, the *interacting neighbors model* (Verguts & Fias, 2005) assumes that for each commuted pair there is a single representation in the arithmetic facts memory. However, unlike the *identical elements model*, the *interacting neighbors model* states that the arithmetic facts representations preserve the information about the order of the operands. This model has been developed for multiplication and it involves four fields: input field, semantic field, decomposition field, and response field (figure 1.5).

---

[2] In this example the order in which the operands are written inside the brace brackets is only explicative and do not represent any information about a order of the operands. Moreover the operation sign (×) does not represent the symbol associated with the operation but the “operation” (multiplication in this case) in the mathematical sense.

Like in the *network retrieval model*, the input field is organized with two separated sets of nodes, one for each operand. However, unlike the *network retrieval model*, the operands are selected with respect to their size before accessing and activating the node inside the input field. Namely, there is one set for the larger operand and one set for the smaller one. For example, the presentation of the problem 7×3 activates the node 7 in the larger operand set and the node 3 in the smaller operand set. The nodes of the input field activate then the semantic field, where the results of the problems are represented.

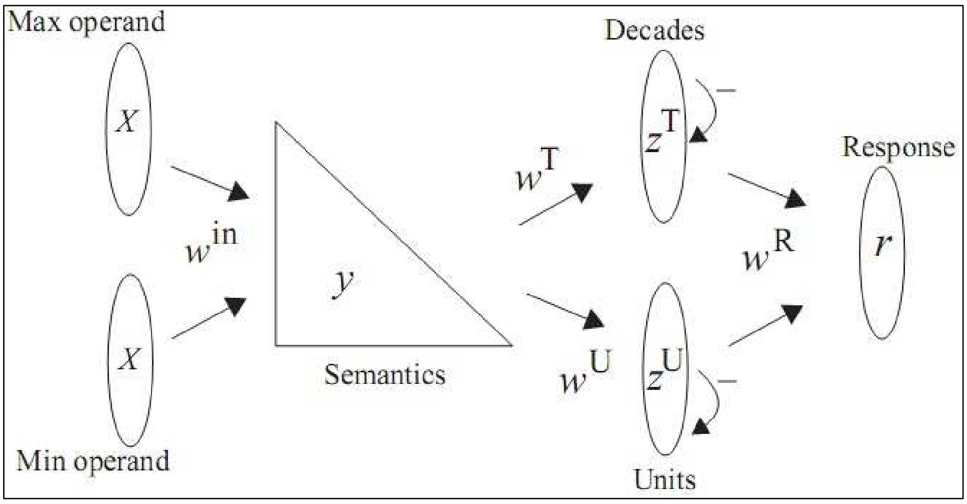


**Figure 1.5: interacting neighbors model (modified by Verguts & Fias, 2005). "Max operand" and "Min operand" sets of nodes represent in the input field the nodes for the larger and the smaller operand, respectively.**

The *interacting neighbors model* states that in the semantic field only one half of the multiplication table is represented. Namely, only the results of the problem in the L×s[3] order are represented. Moreover, the model assumes that the organization of the semantic field is based on the operands table, similarly to the organization described for the *table search model*. For example, if the problem 7×3 is presented the result 21 is activated, but also the

3 Hereafter the string "L×s" means that the multiplication problems where the first operand is larger than the second one (e.g., 7×3); "s×L" means that the first operand is the smaller one; "L+s" means that in the addition problem the first operand is the larger one; "s+L" means that the first operand is the smaller one.

result 14 and 28 (14=7×2, 28=7x4; these result are close to the correct result in the 7 table) are activated. Moreover, it is assumed that the size of the result also affects the activation in the semantic field. In the previous example, other possible results of the operation, member of other tables, that are close to the result of the specific problem are activated, e.g., 24 (=6×4) can be activated even if it is not a multiple of neither 7 nor 3. Nevertheless, the activation of the correct result (21 in the example) is higher of the activation of the other possible results (e.g., 14, 28, and 24). Then, the activation of the result nodes in the semantic field spreads to the decomposition field, in which the decade and the unit of the result are represented separately. Each result node spreads activation to the digit that constitute the result. For example, the result node 21 spreads activation to the node 20 in the decades set and to the node 1 in the units set, the node 24 spreads to the node 20 in the decades set and to the node 4 in the units set, the node 14 spreads to the node 10 in the decades set and to the node 4 in the units set, and so on. Since more results are activated in the semantic fields, different decade and unit nodes are activated in the decomposition field. Two different processes act in the decomposition field: cooperation and competition. Namely, the result nodes 21 and 28 "cooperate" in the decades set because they activate the same decade node (20), whereas the result node 14 and 21 "compete" because they activate different decade nodes (10 and 20, respectively). Finally, the activation is spread to the response field, where the results are holistically represented and the highest activated response node (in the previous example the highest activated response node should be 21) is selected as the result of the presented problem.

With respect to the aim of this thesis, it is relevant to underline two features of the model. First, in the semantic field only one half of the problems are represented, that is only one arithmetic fact is represented for each commuted pair. Second, in the input field the smaller and the larger operands are represented separately, that is the smaller operand of a presented problem activates the corresponding node in the smaller operand set, whereas the larger operand activates the corresponding node in the larger operand set. According to this architecture, the order and the size of the operands have to be processed before accessing to both the input field. In other words, if the problem is presented in the stored order it can directly access to the input fields, whereas if it is not in the stored order it need to be reordered before to access to the input field. Verguts and Fias (2005) assumed that the multiplication facts are represented in the L×s order according to the result of a study of Butterworth and Colleagues (Butterworth, Marchesini, & Girelli, 2003), in which the order L×s order was solved faster than the s×L order (e.g., the problem 5×2 was solved faster than 2×5). However, for the *interacting neighbors model* is not relevant which order is stored, but only that each commuted pair correspond to a single arithmetic fact. In fact, the prediction of the model are identical regardless to which the stored order is. Despite the assumption than only one order is stored and that the operands have to be reordered (when presented in the non-stored order) before accessing to the arithmetic fact, the Authors do not assume that the order of the operands can affect the performance. In fact, the reordering process *"takes some time, but the time it takes cannot contribute substantially to the differences between problems with different operands"* (Verguts & Fias, 2005, p. 5). Nevertheless, given the architecture of the model, it must necessarily be assumed that a reordering process exists and that this process takes (even if little) some time.

### 1.2.3 COMP Model

Butterworth and Colleagues (Butterworth et al., 2001) proposed the *COMP model* (figure 1.6) to describe how the addition facts are represented and organized in memory. This model is based on the observation that children use non-retrieval procedures before being able to solve arithmetic problem with the direct retrieval. According to the Authors, the use of non-retrieval procedures during the acquisition of the arithmetic knowledge shapes the addition facts memory so that only the L+s order (e.g., 8+4) is stored as addition fact. In fact, like in the *interacting neighbors model*, the *COMP model* assumes that for each commuted pair only the L+s order is represented in memory. The L+s order becomes privileged because of, "*as experience of addition increases, counting on from the larger addend could serve as the basis of the organization of facts in memory*" (Butterworth et al., 2001, p. 1009).

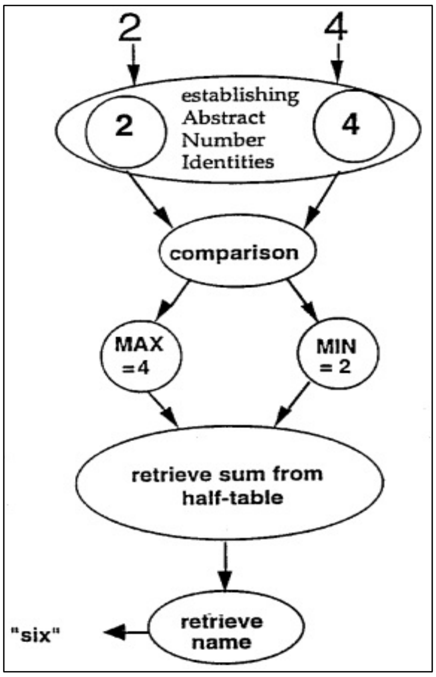


**Figure 1.6: COMP model (modified by Butterworth et al., 2001).**

The procedure of starting to count from the larger operand is the most efficient way to solve an addition problem when the direct retrieval is not yet available as already proposed by the Groen and Parkma (1972) within the frame of *counting models*, discussed at the beginning of this chapter. The L+s order when read left to right is the abstract formal structure that in likely to represents the procedure of counting from the larger and thus could have a meaning like "take the operand L and add s to it" (despite formally L+s does not deny the possibility to interpret it as the converse). Similarly the s+L order is likely to represent the inverse procedure for left to right readers: "take the operand s and add L to it". Hence, the use of the procedure to count from the larger should privilege the L+s order, which could be stored in memory as an addition fact within models that assumes only one order for each commuted pair is stored.

The *COMP model* states that the retrieval of addition facts is based on a comparison process that identify the relative size of the operands (but for evidence contrary to an comparison process see Robert & Campbell, 2008). The architecture of the model assumes in fact four stages. In the first stage the operands of the problem are identified and the "abstract number identities" are activated. These abstract number identities are representations of the cardinal magnitudes of the operands. Once abstract number identities are activated, the comparison stage compares them to identify which is the larger and the smaller operand (the tie are assumed to be in the L+L order). The output of the comparison stage is the representation of the problem in the L+s order (reordered if necessary). In the third stage this representation is used to retrieve the result of the problem. Finally, in the fourth stage the selected result is used to retrieve the form for which the result has to be produced (e.g., the spoken name or the arabic representation).

This model does not make any assumption about the internal organization of the arithmetic facts system, as for example if the result are associated one to each other or if a problem can also activate incorrect results. However, two aspects of this model are relevant for the aims of this thesis. First, it assumes that only the L+s order is stored as addition arithmetic facts, like for the *interacting neighbors model.* Since only one order is stored as arithmetic fact, it is necessary a reordering mechanism before accessing the result. Second, this model assumes that the addition facts memory is organized by the use of non-retrieval procedures during the childhood. Namely, the use of non-retrieval procedures reorganizes the addition facts memory so that the easier to solve order is the privileged. Therefore, like for the *interacting neighbors model*, the performance should be affected by the order of the operands, that is the stored order should be solved faster.

## 1.3 THE ROLE OF THE NON-RETRIEVAL PROCEDURES IN THE ARITHMETICAL COGNITION

The models described so far assume that adults solve multiplication and addition problems mainly (or even exclusively) by means of direct retrieval. In fact, these models states that only during the acquisition (when the children are learning multiplications and additions) the use of non-retrieval procedures is common, whereas skilled children and adults use retrieval only to solve the multiplication and addition problems. However, recent evidence showed that the percent of use of non-retrieval procedures is wider than it was thought previously (Campbell & Austin, 2002; Campbell & Xue, 2001; Grabner et al., 2009; Hecht, 1999; LeFevre, Bisanz, Daley, Buffone, Greenham, & Sadesky, 1996a; LeFevre, Sadesky, & Bisanz, 1996b; Smith-Chant & LeFevre, 2003; Thevenot, Fanget & Fayol, 2007; Zbrodoff & Logan, 2005).

LeFevre and Colleagues (Lefevre et al., 1996a) reported evidence that adults solve simple one-digit multiplications with a mixture of retrieval and non-retrieval procedures (for similar result with addition see LeFevre, Sadesky, & Bisanz, 1996b). The participants in the study had to solve one-digit multiplication problems producing verbally the answer. After each trial, the participants had to report the procedure they supposed to use to solve the problems. The results showed that a noteworthy percent of multiplication problems were solved by means of non-retrieval procedures (the percent of use of non-retrieval procedures was 12% and 18.6% in the experiment 1 and experiment 2, respectively). The analysis of the non-retrieval procedures also showed that the use of these procedures varied according to individual differences and problem properties. More skilled adults tended to base their solution more on retrieval than less skilled adults (for similar results see LeFevre et al., 1996b; Thevenot et al., 2007). Moreover, the selection of the procedure was influenced by the kind of problem that has to be solved. For example, when solved by using non-retrieval procedures, problems with 2 as one of the two operands were often solved by means of rephrasing the problem in terms of a tie addition (e.g., 5×2=5+5=10), the 5-problems were sometimes solved by means of table sequence procedures (e.g., 3×5=5, 10, 15), and the large problems were frequently solved by means of derived-fact procedures (e.g., 6×9=(6×10)−6=60−6=54). See also Campbell & Penner-Wilger (2006) for similar result on the differences in the use of non-retrieval procedures between large and small problems.

LeFevre and Colleagues (LeFevre et al., 1996a) explained the results of their study within the context of the *Adaptive Strategy Choice Model* (ASCM) (Siegler & Shipley, 1995). The ASCM model was developed in order to explain the performance of children when they solve addition problems. According to this model, the procedure selected to solve a problem depends on the probability of success of that procedure has and the strength of association between the problem (the operands) and the result. Stronger is the association between the operands and the result higher is the probability to use retrieval procedures to solve the problem. Furthermore, if the strength of association is weak the use of non-retrieval strategies is more likely, and the strategy that is selected depends on the probability of success of that specific strategy with respect to others. LeFevre and Colleagues (LeFevre et al., 1996a) suggested to extend the ASCM to the adults. In fact, even though the original model assumes that the children use a variety of strategies, the prediction of the ASCM is that during the development the children switch from a mixture of retrieval and non-retrieval strategies to a pure retrieval procedure in the adulthood. However, the assumption that adults use exclusively retrieval procedures to solve one-digit problems does not fit with evidence supported by the studies in which the participants are required to report the strategies they used (Campbell & Austin, 2002; Campbell & Xue, 2001; Hecht, 1999; LeFevre et al., 1996a; LeFevre et al., 1996b; Smith-Chant & LeFevre, 2003; Thevenot et al., 2007). Surely, the retrieval is a very common procedure used by the adults to solve simple arithmetic problems. Other procedures, however, are available and can be used in case the retrieval fails to identify the result or the individuals habitually rely on non-retrieval procedure to solve a particular problem.

## 1.4 THE OPERANDS-ORDER EFFECT

In the literature there is few evidence that order of operands can affect the way in which or the speed at which the cognitive system processes problems of commutative operations. This evidence comes mainly from studies on multiplications solution in Chinese population (LeFevre & Liu 1997; Zhou, Chen, Zhang, Chen, Zhou, & Dong, 2007). In Western countries, multiplication is usually learned by studying the whole table, which typically includes the problems from 1×1 to 9×9 in both orders. In the Chinese arithmetic teaching system only one half of the table is learned, pupils in fact learn only the problems in the s×L order (e.g., the problem 3×7 is learned, 7×3 is not). This peculiarity of the educational system is assumed to be the cause of an operands-order effect found in the adult Chinese population (LeFevre & Liu, 1997; Zhou et al., 2007). Studies showed that there are behavioural (LeFevre & Liu, 1997) difference between the two orders of the operands with an advantage for the s×L order (e.g., 3×7) compared to the L×s order (e.g., 7×3); in an ERPs study Zhou et al. (2007) have shown that non-privileged order elicit a long lasting frontal negativity with respect to the privileged one with a very early onset (120 ms). LeFevre & Liu (1997), comparing Chinese and Canadian participants, found a clear operands-order effect in the formers and only weaker effect in the latters. In fact, the results showed that the Chinese participants solved the multiplication problems in the s×L order (948 ms) faster than in the L×s order (978 ms). Furthermore, at the end of the experiment when the participants were interviewed, the 60% of the Chinese participants (12 on 18) "*spontaneously*[4] reported using a procedure of reversing the digits" (LeFevre & Liu, 1997, p. 56) to solve the problems in the s×L order (e.g., they solved the problem 9×6 transforming it in the problem 6×9). The operands-order effect was stronger in the reverse group (participants that reported to use reverse strategy) than in the non-reverse group (participants that did not reported to use reverse strategy), 66

[4] Italic of the Authors (LeFevre & Liu, 1997).

ms and 14 ms, respectively. On the contrary, the Canadian participants showed an operands-order effect that varied in size and direction across the various problems. The effect was significant only with the problems were one of the operands was 4, 5, or 9 and consisted in a vantage for the problems in the L×s order (the mean RTs of the L×s and s×L orders were 1.186 ms vs 1.224 ms, 1.093 ms vs 1.152 ms, and 1.369 ms vs 1.418 ms, for the 4, 5, and 9 problems respectively). The results of the interviews to Canadians participants showed that the operands-order effect was significant only for participants that reported to use both retrieval and non-retrieval procedures to solve the multiplication. These results cannot discriminate however whether the operands-order effect has to be attributed to retrieval procedures or to non-retrieval procedures. In fact, for Chinese participants the effect can be due to both the fact that they learn only one half of the multiplication table (i.e., the influence of the learning experience on the retrieval process) and the fact that some of them explicitly adopt a reversing procedure before retrieval to solve the problems (i.e., the L×s problems are solved by reversing the order before retrieval, a procedure that takes some time). Similarly, it is not clear whether the effect of the order of the operands in Canadian participants was due to the non-retrieval procedure (and the selection of this specific procedure among possible others), to the retrieval procedures, or both. The operands-order effect could be more evident in the Chinese population because of they are “forced” by the educational system to base the acquisition of the one-digit multiplications knowledge mainly on the s×L order, and hence the effect can be easily found simple averaging across the two operands orders. On the contrary, since the Western populations learn the whole table, any putative operands-order effect could be more difficult to detect because the development of a preferred order can differ both across problems and participants and thus, by comparing s×L and L×s order overall, across problems and/or participants, can easily lead to small or, even worst, null results. Despite this, some evidence that the order of the operands can affect the performance also in Western populations are provided by a study of Butterworth and Colleagues (Butterworth, Marchesini, & Girelli, 2003). In this study, Italian children of the third, four, and fifth grades (8, 9, and 10 years old, respectively) had to perform multiplication problems between 2×2 and 6×5. The Authors aimed to verify if one of the two orders of the operands were privileged in terms they are solved faster by children. In the Italian educational system, unlike England or USA, the name of the table is in first position (e.g., the 2 table is 2×1, 2×2, 2×3, 2×4, and so on; the 3 table is 3×1, 3×2, 3×3, 3×4, and so on), then the s×L order is taught before the L×s order (e.g., 3×6 is learned before 6×3). The Authors taken into account two hypotheses about the privileged order. The first hypothesis is that the s×L order could be solved faster due to the larger learning experience. In fact, the s×L order is learned before and more practised than the L×s order, in according to models that assume frequency as a key factor in shaping activation weights within the arithmetic fact memory (*distribution of association model* and network retrieval model, see sections 1.1.3 and 1.1.4 of this chapter). The second hypothesis is that the privileged order depends on a reorganization of the multiplication facts memory when new facts that correspond to the commutate of an already learned problem is learned. Moreover, based on the observation that the children often use procedures like repeated addition to solve multiplications, the privileged order should be the easier to solve with this procedure, that is the L×s order. To transform in a efficient way a multiplication in a repeated addition problem, the children must consider both the order and the size of the operands. In fact, the repeated addition procedure could be faster and easier to solve when the larger operand is in first position. For example, given the problems 3×6 and 6×3, the most efficient way to transform them in an repeated addition is repeat 3 time the larger operand (6+6+6) independently from the order of the operands, and the problem 6×3 could more often suggest the use of the larger operand than the problem 3×6. The results showed that the L×s order was solved faster than the s×L order (e.g., the problem 5×2 was solved faster than 2×5). Moreover, this operands-order effect was significant only for the fourth and the fifth grade, and it was significant for the 2 table, marginally significant for the 3 and 4 table, and non significant for the 5 table. In other words, the effect emerged only for the older children (the ones for which the Authors assume the reorganization had have more time to shape the memory system) and for the earliest learned multiplications (that is, the problems "most susceptible to reorganization"). Butterworth et al. (2003) interpreted the effect as produced by a reorganization of the arithmetic facts memory due to both the use of the repeated addition procedure and the comprehension by the children of the commutative property.

## 1.5 AIMS OF THIS THESIS

In this chapter we have briefly reviewed the main models that describe the architecture of the arithmetic facts memory and the retrieval processes. It was also showed that a noteworthy percent of the multiplication and addition problems are solved by the adults by means of non-retrieval procedures. Finally, some evidences have been reported that order of the operands can affect performance in arithmetic problem solving.

Despite the relevance of the commutative property in arithmetic and the important role of order assumed by early models of arithmetical cognition (Groen and Parkman, 1972), the following development of the field largely neglected the problem by assuming a symmetry between commuted pairs. Only in more recent models (i.e., *the identical element model*, the *interacting neighbors model*, and the *COMP model*) the problem of order of operands and the fact that the commutativity can contribute to shape the arithmetic facts memory is considered. Despite this renewed interest, even models like that of Verguts and Fias (2005, *interacting neighbors model*) claim that an possible reordering process has not effect on the performance since the time it requires is very brief. Thus, despite their model assume only one order is store for each commuted pair, it is invariant with respect to which of the orders is stored. It is possible that this theoretical underestimation of operand order and of the cognitive processing of commutative property in general is due to the lack of strong empirical evidences. Nevertheless, there are both behavioural (LeFevre & Liu, 1997) and electrophysiological (Zhou et al., 2007) evidences that the order of the operands can contribute to the organization of the multiplication facts memory in Chinese population. Some evidence show that the order of the operands can affect as well the performance in populations that learn the whole multiplication table, that is Canadian adults (LeFevre & Liu, 1997) and Italian children (Butterworth et al., 2003).

The aim of this thesis is to provide new and clear empirical evidences of operand order effects. The lack of strong evidence of order effect in processing commutative problems by Western population that learn both orders of a commutative operation, can be due to several factors: a) the two orders are processed more or less in the same way (e.g. inversion is costless as Verguts and Fias assumes); b) the preferred order can differ from a problem to the other and the preferred order for each pair can be idiosyncratic (e.g. different for different individuals); c) there are systematic preferences in the population but not all problems in the table share the same preference of order; d) the effect emerges only with explicit production task and not with other simpler tasks used to study arithmetic cognition (e.g. verification task, implicit automatic activation of multiples and dividends, etc.).

For all these reasons this research started with an experiment where the paradigm was chosen in order to maximize the possibility to find any preference of order, both at group level and individual level. We chosen a production task (in which the participants have to produce the result of presented problems) with a sequential presentation of the operands of the problems. The sequential presentation should emphasize the order in which the operands are presented. Moreover, we asked to the participants to report the procedures they supposed to use when they solved the problems. The procedures reported could provide interesting information about the role of the order of the operands in the procedures selection.

Five experiments were conducted by using various methodologies and by collecting data from two different mother language groups. In the first experiment we tested the performance on both multiplication and addition of Italian participants. This experiment involved two tasks:

a chronometric production task, in which the participants were required to product the result of the presented problems; and a self-report production task, in which the participants were required to both product the result and to report the procedures they supposed to use to solve the problems. The result of this experiment has been used to expand the two hypotheses of this thesis and to schedule the following experiment. In the second experiment English participants performed exactly the same tasks used in the first experiment. This second experiment was conducted to test the possibility that linguistic and cultural differences could affect the result we found in the first experiment. In the third experiment we adopted a matching task and a new task expressly created to test our second hypothesis about the asymmetric activation spreading between the result nodes. However, due to methodological troubles we could not test our prediction. In the fourth experiment, we adopted a verification task in which the participants had to verify if the presented result of a multiplication problem was correct or incorrect. The result of this experiment have been used to improve the interpretation of the result of the first experiment and to test the asymmetric activation spreading hypothesis. Finally, in the fifth experiment, we adopted the event related potentials (ERPs) methodology to investigate the electrophysiological correlated of the operands-order effect we found in the previous experiments.

Despite the focus of the project is mainly empirical, to efficiently plan and interpret the results the models described in this introduction have to be ranked, at least in terms of working hypothesis. Clearly, given that our aim is to study possible asymmetries in the processing of two problems that constitute a commuted pair, models that consider order has to be chosen to describe the phenomena under study. Both the *COMP model* and the *interacting neighbors model* are considered to develop hypothesis and discuss the results. The results of this thesis are discussed in the last chapter in terms of new constrains to models of arithmetical cognition, coming from order effects in the processing of one digit commutative arithmetical problems.

# Chapter 2

## The reorganization of arithmetic facts in memory affects the speed of resolution of arithmetic problems: a study of operand order effects in Italian and English

## 2.1 INTRODUCTION

As we have seen in the first chapter a fundamental property of arithmetic problems like multiplications and additions is the commutativity, which means that the product is the same regardless of the order of the operands. Despite this, the two orders of a commutative pairs are mathematically two distinct problems and may be processed differently. Within empirical studies the selection and statistical analysis of the experimental stimuli typically consider the commuted pairs as the same problem (see for example, Metcalfe, & Campbell, 2011; Campbell, & Austin, 2002; Kirk, & Ashcraft, 2001; Smith-Chant & LeFevre, 2003). This practice correspond to the assumption, either explicit or implicit, that the order of the operands does not affect the performance.

Both the *interacting neighbors model* (Verguts and Fias, 2005) and the *COMP model* (Butterworth et al., 2001) assume that only one order is represented as arithmetic fact for multiplication and addition, respectively (see chapter 1). The *interacting neighbors model* assumes that the multiplication facts are represented in the L×s order (e.g., 7×3) and the *COMP model* assumes that the arithmetic facts are encoded in the L+s order (e.g., 7+3). Therefore, the operands reordering process is a fundamental component for both models that allows to access to the arithmetic facts memory. This reordering process should affect in different way the performance on the problems of the commuted pairs. In fact, the retrieval of the result for the problem presented in the non-stored order should be more time consuming than for the problem in stored order due to the additional reordering process needed to access to the arithmetic fact. However, despite the theoretical relevance of the reordering process, only very few evidences have been so far collected that the order of the operands can affect the performance in multiplication and addition problems (see LeFevre & Liu, 1997; Zhou et al., 2007; and chapter 1).

The *interacting neighbors model* and *COMP model* assume that one-digit additions and multiplications are solved by means of retrieval. This very likely to happen for persons that practice arithmetic in everyday life, however, the arithmetic problems can also be solved by means of non-retrieval procedure (see chapter 1). Non-retrieval procedures involve the use of abstract mathematical procedures, like rules (e.g., $N\times1=N$), derived-fact ($7\times9=(7\times10)-7$), counting ($13+2=13+1+1$), repeated additions ($5\times3=5+5+5$), etc. Retrieval and non-retrieval procedures are assumed to contribute to produce one of the most known effect in arithmetical cognition, that is the problem-size effect. This effect refers to the fact that small[5] problems (e.g., $3\times4$) are solved faster and are less prone to errors than the larger problems (e.g., $7\times8$). In fact, the problem size effect can be explained by three factors (Campbell and Xue, 2001): 1) retrieval procedures are more efficient for small problems than for large problems; 2) non-retrieval procedures are more efficient for small than for large problems; 3) the use of retrieval procedures (that are generally faster than non-retrieval procedures) may be more common for small than for large problems. Clearly the three explanations are not mutually exclusive. These factors could influence operands-order effects as well. Namely, like for the size effect, retrieval and non-retrieval procedures could concur into generating an asymmetry in the performance between the two orders of the commuted pairs. Therefore, the operands-order effect might also depend on an asymmetry in the selection of different procedures to solve the two orders (one of the two orders could rely on procedures that are slower than those used to solve the inverse order).

The use of non-retrieval procedures can vary as a function of the structural characteristics of a problem (LeFevre et al., 1996a; LeFevre et al., 1996b), that is the size of the problem influences the procedure adopted to solve it. In fact, retrieval procedures are generally more often adopted to solve small problems, whereas non-retrieval procedures are more often used with large problems (see for example LeFevre et al., 1996a; LeFevre et al., 1996b; Campbell & Austin, 2002). Therefore, the order of the operands (a structural characteristic of the problems) could also influence the selection of the procedures adopted to solve the

[5] Here small problems, differently from divisions or subtractions, refers to both small operands and small results, given the intrinsic correlation between size of operands and results for additions and multiplications.

problems. In other words, a problem (7×3) and its commuted (3×7) could urge or suggest individuals on to select different solving procedures. The different procedures selected to solve a problem and its commuted could generate a difference in the latency of response times (RTs) between the two orders because of one of the two procedures could be more time consuming than the other.

Our purpose is to evaluate if the order of the operands can also affect the selection of the procedures. For example, the order could suggest to adopt a particular derived-fact procedure, that is the inversion of the operands (e.g., to transform the problem 3×7 into the problem 7×3). The probability of success of a procedure and the strength of association between the operands and the result could in fact be affected by the order of the operands. Namely, the probability of success of a procedure and the strength of association could be different for the two problems of a commuted pair. For example, one of the two orders could make easier to access to the corresponding arithmetic fact than the inverse order. Therefore, one order could be stronger associated with retrieval procedure, whereas the other order could be urge the selection of slower non-retrieval procedures.

The aims of the experiment presented in this Chapter are: 1) to test if the order of the operands can affect the performance in a result production task with addition and multiplication problems; 2) to evaluate if the order of the operands can affect the selection of the procedures (retrieval and non-retrieval) that allow to solve these arithmetic operations. The second goal is not independent from the first one since it could aid in interpreting possible operands-order effects found in the RTs. Differences found in the selection of the solving procedures across the two operands order could in fact inform whether the operands-order effects in the chronometric data is due to retrieval or to non-retrieval procedures. To this end, we decided to adopt two production tasks: 1) a chronometric production task (under the instruction of both speed and accuracy); 2) a self-report in which the participants had to product the result of the problem and report the procedures they used (under only the instruction of accuracy).

In a production task participants are usually presented with a problem and have to produce the associated result. For example, a participant has to produce the number 42 when presented with the stimulus “7×6”. The production tasks, unlike the verification tasks (that requires to judge whether a presented equation, e.g. “7×6=42”, is true or false), involve the full identification of the result of the arithmetic problem (Zbrodoff & Logan, 1990; Zbrodoff & Logan, 2000). The verification task could be performed by solely retrieval procedures or by recognising of the whole presented equation, given that if the proposed result fits with one of the more active result representation a yes/no choice could be done, achieving a rather good accuracy, without having selected the result between the more active representations in advance to the presentation of the proposed result. Differently, in a production task, the explicit selection of the response needs the full identification of the result and this can make more likely the implementation of non-retrieval procedures in case the retrieval procedure takes too much time. Since our aim was to force the experimental paradigm in order to find operands-order effects, that could be due to the selection and/or implementation of non-retrieval procedures the choice of a production task is favourite with respect to a verification task. Therefore, it would be interesting to understand if any asymmetry between the two orders, if it exists, is reflected not only on RTs but also on a difference on the procedures (retrieval and non-retrieval) that are self-reported by participants after having solved the problem. This would indicate that they are aware of a qualitative difference in the solution of a problem and its commutate. The procedures used by the participants during a production task are in fact typically collected by using a self-report (see for example, Campbell & Xue, 2001; Hecht, 1999; LeFevre et al., 1996a; LeFevre et al., 1996b). In these studies the participants had to product the result of a problem and just after they had to immediately report the procedure they suppose they have used to solve that problem. However, this methodology has been criticized by Kirk & Ashcraft (2001). In this study the Authors found that instructions can significantly affect the procedures that are reported by the participants. Retrieval biased instructions induced participants to report retrieval more often, whereas non-retrieval biased instruction induced participants to report non-retrieval procedures more often. Moreover, Kirk and Ashcraft (2001) found that RTs were also affected by the instruction bias, that is participants non-retrieval biased solved the problems slower overall. However, some other studies (Campbell & Austin, 2002; Campbell & Penner-Wilger, 2006; Smith-Chant & LeFevre, 2003) have showed that, despite any possible biases, the self-report holds some validity and thus it can be used to collect information about which procedures the participants used. One possibility to avoid the problems associated with the bias in the on-line self-report is the use an off-line self-report, that is to divide the chronometric and the self-report production tasks into two different experimental sessions. In the first session (chronometric task) the participants have to product the result of the problems just giving the standard instruction of a RTs task, e.g. to be "as quick and accurate as possible", and without asking them any report about procedures used; whereas in a second subsequent session (self-report task) the same participants have to solve again the problems without any time pressure and report, after the solution, the procedure they think they have used. This off-line self-report task has the disadvantage that the participants could use a different procedure in the chronometric task and in the self-report one, especially given that time pressure can induce a strategic difference in the way a problem is solved. On the other side the advantage of splitting the two tasks are that both task are more ecological, the chronometric task in fact is unbiased with respect to both the experiment aims and instruction biases, covering possible criticisms similar to that by Kirk & Ashcraft (2001). The second task (self-report) is also more ecological since, independently from how the problem was solved in the chronometric task, it  allows the participant, in absence of time pressure, to have time to think about which is the way he/she typically solves this problem in everyday life where calculations are typically performed without much time pressure and with more attention to accuracy. To be more precise, the self-report allows to collect indications about how the participants perceive they prefer to solve a given problem in order to gain maximal accuracy. For example, a possible use of the self-report data in aiding the interpretation of RTs data would be the following rationale: an absence of asymmetry in the reported procedures across the two orders of a problem would be useful to exclude that possible operands-order effects found in the RTs can depend on the selection of procedures. If no differences would be found in the second session (self-report), possible effects on the first one (chronometric task) would be much likely attributed to the speed of retrieval only (i.e. the dynamic of activation of the arithmetic facts in long term memory).

A second possible interplay between self-report and chronometric data would be that in absence on an overall order effect on RTs it is nevertheless possible that idiosyncratic order differences exists (e.g. some order for some specific problems are preferred in the population or in single participant). In this case the self-reports can be used to perform explorative analyses in order to select cells or groups of problems that are reported to be solved with specific non-retrieval procedure in an unspeeded, maximal accuracy, task and see if the same problems also show effects in the chronometric task. This kind of strategy has been already used to define experimental conditions in an fMRI study where physiological data were correlated with off-line self reports, showing a greater activation of the left angular gyrus when participants reported fact retrieval whereas the non-retrieval procedures were associated with a broad activation in the frontal and parietal areas (Grabner et al. 2009).

In this study we aimed to test the hypothesis that the order of the operands can affect the RTs and/or the selection of procedures. To this end, two experiments were conducted on two groups of participants both belonging to populations that learn the whole multiplication table. In fact, until now the stronger evidences found regard the Chinese population, that learn only one half of the multiplication table. In the first experiment Italian participants performed a chronometric task followed by a self-report one, whereas in the second experiment the same tasks were performed by English participants. These two groups differ in the order in which the multiplication table is learned. In the Italian learning system the s×L problems (e.g., 2×7) are taught before the L×s problems (e.g., 7×2), whereas in the English system the L×s problem are taught before the s×L problems. In fact, the Italian children learn the 2-table in the order 2×1, 2×2, 2×3, 2×4, ..., 2×9, whereas English children learn it in the order 1×2, 2×2, 3×2, 4×2, ..., 9×2.

## 2.2 EXPERIMENT 1: ITALIAN PARTICIPANTS

The aim of the experiment 1 was to study if the order of the operands in commutative arithmetic operations (multiplication and addition) can affect the RTs and the selection of the procedures. Moreover, we would like to evaluate if the selection of procedures can aid us to explain an operands-order effect found in RTs. To this end we used two production tasks: a chronometric task (in which the participants had to simply product the result of the presented problems) and a self-report task (in which the participants had to product the result and report the procedures they supposed to use to solve the problems).

According to the *interacting neighbors model* (Verguts & Fias, 2005) and he *COMP model* (Butterworth et al., 2001) behavioural advantages in solving the L×s order (in multiplication) and the L+s order (in addition) are expected in the chronometric task. Moreover, in the self-report we expect as well as that participants report more often retrieval for the orders that are supposed to be stored (i.e., L×s and L+s) than for the inverse orders (i.e., s×L and s+L).

Nevertheless, since we do not exclude in principle that different orders can be preferred for different problems and given that size of the problem is the main variable that typically affects RTs and thus reflects problem difficulty, the size factor have clearly to be taken in account within our experimental design. The arithmetic problems are usually divided into two categories: small and large problems (see for example, Campbell & Austin, 2002; Campbell & Penner-Wilger, 2006; Jost, Beinhoff, Hennighausen, & Rosler, 2004; Zhou et al., 2006). However, we divided the multiplication table into 3 partitions, according to the size of the single operands (for a similar division in 3 levels see Pauli et al., 1994; for a division in 4 levels see Kirk & Ashcraft, 2001; Smith-Chant & LeFevre 2003). Namely, we distinguished each operand in small and large: the operands between 2 and 5 (included) were considered small; whereas the operands between 6 and 9 (included) were considered large. In this way each problem is associated with two labels (one for each operand), which produce three possible combination: small-small, small-large, large-large. For example, the problem 2×4, where both operands are small, falls within the small-small category (hereafter small problems); the problem 3×8, where one operand is small and the other large, falls within the small-large category (hereafter medium problems); the problem 7×8, where both operands are large, falls within the large-large category (hereafter large problems). This 3-level classification of size was done since models that assume order asymmetries in arithmetic facts memory (*interacting neighbors model* and COMP model) assume an important role of operand sizes comparison as a preliminary stage to access to the nodes in the associative network that allow to retrieve the result of arithmetical problems.

### 2.2.1 Method

**Participants.** Twenty-four native Italian-speaking students (12 females; mean age: 28, *sd*: 5.49) from the University of Trento participated in the experiment as volunteers. All participants had normal or corrected-to-normal vision. This experiment, like all the experiments reported in this thesis (excepted the experiment carried out in London), was approved by the Ethical committee of the University of Trento.

**Material.** The participants had to perform two tasks: a chronometric task and a self-report task. In both tasks the stimuli were the same. Single-digit multiplication and addition problems were used during the experiment. The problems with 0 or 1 (e.g., 0×3, 0+5, 1×3, etc.) were excluded because they are likely solved by means of rules (LeFevre et al., 1996a).

Since we are studying the operand-order effect, the tie problems (e.g., 2×2, 3×3, 2+2, 3+3, etc.) were also excluded. Due to these constraints, there were 56 problems for each operation (a total of 112 problems). Each problem was presented once.

**Procedure.** The two task were subsequently performed with the self report following the chronometric one, in a same experimental session with a small break between the two. In the chronometric task each operation (multiplication and addition) was presented in separated blocks (2 blocks of 56 problems each). In order to familiarize with the experimental procedure, before each block the participants performed some practice trials with problems with 0 and 1 as operands (e.g., 0×3, 0+5, 1×3, etc.). The order in which the addition and multiplication blocks were presented was counterbalanced across the participants. The problems were sequentially presented at the centre of a monitor of a PC: the first operand was presented for 300 ms, followed by the sign of the operation ("+" or "×") for 300 ms, and finally the second operand for 300 ms. The second operand remained on the screen until the participants responded. However, if the participants did not responded within 9 seconds the second operand disappeared and the next trials started. The operands and the operation signs had a dimension of about 1 cm and the participants were at about 60 cm from the monitor. Each trial started with a briefly blinking fixation point ("#"). Participants were required to respond when the second operand was displayed with the right hand, by using the numeric keypad on the right of the PC keyboard, and they were instructed to be as quick and as accurate as possible. The participants had to press the keys corresponding to the digit of result of the problem (one key if the result was with one digit; two keys if the result was with two digits). The RTs and the accuracy of the keys pressed (one or two according to the number of digit of the result) were recorded. Between the two blocks the participants could take a little break.

After the chronometric task, the participants had to perform the self-report task on a notebook computer, in which they had to solve the same arithmetic problems. The order in which the operations blocks were presented was the same as in the chronometric tasks. In this second task, the participants had to report for each problem the result, the procedure used to solve the problem, and the perceived difficulty. In this task the participants were required to be as accurate as possible without time pressure (they might take all time they need to solve the problem and report the difficulty and the strategy). Before starting the task a sheet with the description of the procedures was given to the participants, who could take the sheet during the task to remember the procedures description. There were 5 procedures among which the participants could choose: retrieval, transformation, counting, inversion, and other. On the sheet given to the participants the procedures were described as following:

- Retrieval: “you remembered the solution of the problem, that is you retrieve the result directly from memory”
- Transformation: “you solved the problem by using other problems that can be members of the same arithmetical operation or of another operation (e.g., you solve the problem 9×9=? by using 9×10=90–9 = 81)”.
- Counting: “you solved the problem counting (maybe in a quiet voice) a certain number of times until you obtain the result of the problem (e.g., you solve the problem 4×4 by counting 4..8..12..16; or you solve the problem 13+4 by counting 13..14..15..16..17)”.
- Inversion: “you reversed the two operands to be able to find the result of the problem (e.g., you solve the problem N1×N2 by using the problem N2×N1)”.
- Other: “you solved the problem by using another procedure or you are not sure about the procedure used”.

The perceived difficulty was classified with a Likert scale from 1 (very easy) to 5 (very difficult). The problem, the procedures to select and the likert scale were presented together in the same screen. Therefore, unlike in the chronometric task, the operands and the sign of the problem were simultaneously presented, and they remained on the screen until the participant reported the result, the strategy and the perceived difficulty. The participants were required to solve first the problem and then to select the used procedure and the perceived difficulty. The participants had to use the numeric key on the notebook keyboard to report the result and the mouse to select the strategies and the difficulty. The problems were presented on the screen with on the right a white space in which the participants had to report the result of the operation (the white space had been selected with the mouse before write the result). Below the problem there was the Likert scale of difficulty (form 1 to 5), and below this scale there was the strategies (5 options). Once the participant filled in all the information required they could go to the next trials by pressing the “Enter key”. If the participant forgot to fill in one or more information a message dialog appeared on the screen asking to complete all the sections. The participants were asked to report the procedure and the perceived difficulty associated to the problem solved during the self-report and not trying to remember how they solved the task during the chronometric experiment.

**Data analysis.** We used the same statistical analysis for both multiplication and addition. The two operations have been analysed separately. For the chronometric task, for both RTs and accuracy (proportion of correct answers) a two-way repeated measures ANOVA was performed with size and order as within subject factors. The size factor included three levels: the problems with both operands larger than 5 were coded as “large” (e.g. 7×8); the problems with one operand larger and one smaller than 5 were coded as “medium” (e.g. 7×3); the problems with both operands smaller than 5 were coded as “small”. Both orders of the problems 6×5, 7×5, 8×5, and 9×5 were coded as “medium”, whereas both orders of the problems 2×5, 3×5 and 4×5 were coded as “small”. The order factor had two levels: L×s (or L+s for addition) and s×L (or s+L). For each participant we calculated the mean RTs and the proportion of correct answers in the six experimental cells (order X size). In the analysis of the RTs the ANOVA was performed on the correct trials and for each participant outliers have been removed using the outlier procedure described in Van Selst and Jolicoeur (1994).

This procedure recursively remove the data points that beyond 3.5 standard deviation from the mean RTs of each participant (for technical detail refer to the paper of Van Selst and Jolicoeur). In order to interpret ANOVAs significant main effect or the interaction t-test corrected with the FDR method were performed between different cells of the design. ANOVAs were Greenhouse-Geisser corrected when the degrees of freedom of a factor exceeded one (uncorrected degrees of freedom and epsilon values are reported).

The self-report results have not been statistically analysed, but descriptive statistics are used to aid the interpretation of the effects found in the chronometric task. We aggregated the procedures reported by the participants in the same cells used for the ANOVAs. The results of the difficulty likert scale have not been used because of the participants reported almost only the value 1 (easy).

### 2.2.2 Results

#### *Multiplications*

Each of the 24 participants had to solve 56 multiplications (from 2×2 to 9×9, tie problems excluded). Participants made errors on 8% of the trials, 109 errors on 1344 trials (56 problem × 24 participants). A two-way repeated measures ANOVA with size (small, medium, and large) and order (L×s and s×L) as within factors was conducted on the accuracy for the multiplication problems. The ANOVA revealed only a main effect of the size, $F(2,46)=22$, $\varepsilon_{GG}=0.61$, $p<0.001$. Post-hoc analysis revealed that the participants made more errors in the large condition (77% of correct answer) than in both medium (94%) and small (98%) condition, $t(23)=-4.49$, $p<0.001$ and $t(23)=-5.09$, $p<0.001$ respectively. Moreover, the participants made more errors in the medium condition than in the small condition, $t(23)=-2.19$, $p<0.05$. Neither the order factor nor the interaction reached the significance level.

In the chronometric task the participants had to press one or two keys according to the number of digits of the results. The analysis of the RTs of the two keys showed a correlation of 0.98. Therefore, we analysed only the RTs associated with the first key pressed. A two-way repeated measures ANOVA with size (small, medium, and large) and order (L×s and s×L) as within factors was conducted on the RTs of the first key for the multiplication problems (see figure 2.1).

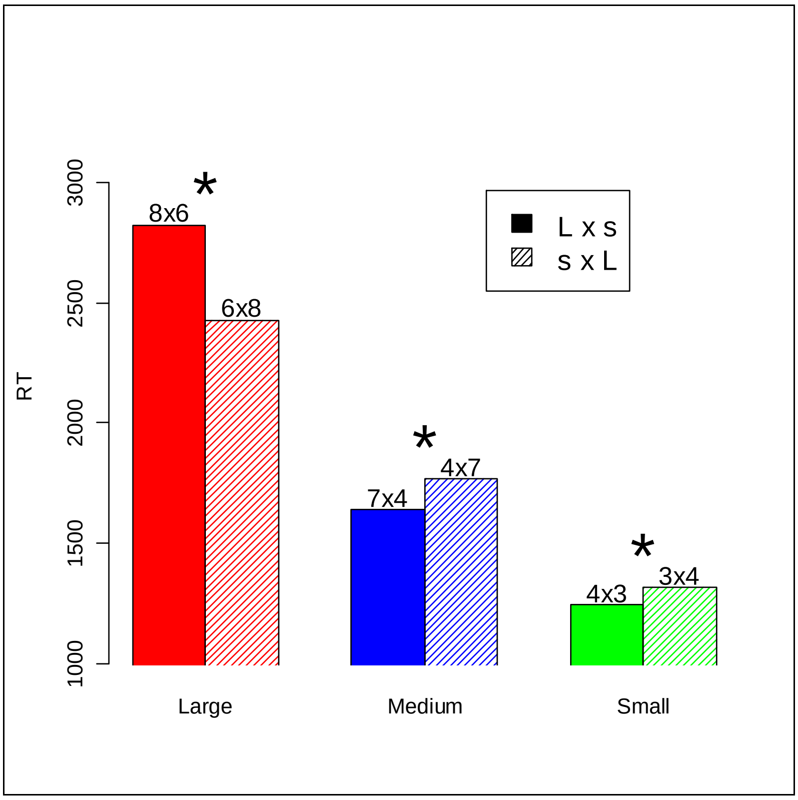


**Figure 2.1: Reaction times as function of size and order of the operands for the multiplication problems (Italians group). * *p*<0.05.**

The ANOVA revealed the a significant main effect of the size, F(2,46)=51.65, $\varepsilon_{GG}$=0.56, p<0.001. Post-hoc comparison revealed that the participants responded faster in the small condition (1276 ms) than in both medium condition (1696 ms) (t(23)=8; p<0.001) and in large condition (2602 ms) (t(23)=7.72; p<0.001); and that they responded faster medium condition than in the large condition (t(23)=6.15; p<0.001). The ANOVA revealed also a significant order by size interaction, $F(2,46)=10.21$, $\varepsilon_{GG}=0.6$, $p<0.01$. Post-hoc comparison revealed that when both operands were larger than 5 (large condition): the problems in the s×L order (2415 ms; e.g., 7×8) were solved faster than the problems in the L×s order (2799 ms; e.g., 8×7), $t(23)= 2.87$, $p<0.01$. When one operand was larger than 5 and one smaller than 5 (medium condition): the problems in the L×s order (1631 ms; e.g., 7×3) were solved faster than the problem in the s×L order (1761 ms; e.g., 3×7), $t(23)=-2.63$, $p<0.05$. When both operands were smaller 5 (small condition): the problems in the L×s order (1237 ms; e.g., 4×2) were solved faster than the problem in the s×L order (1315 ms; e.g., 2×4), $t(23)=-2.65$, $p<0.05$.

For the multiplication trials, we qualitatively analysed the self-report data by aggregating in the same cells as in the ANOVAs the percent of use of each procedure (figure 2.2 and table 2.1).

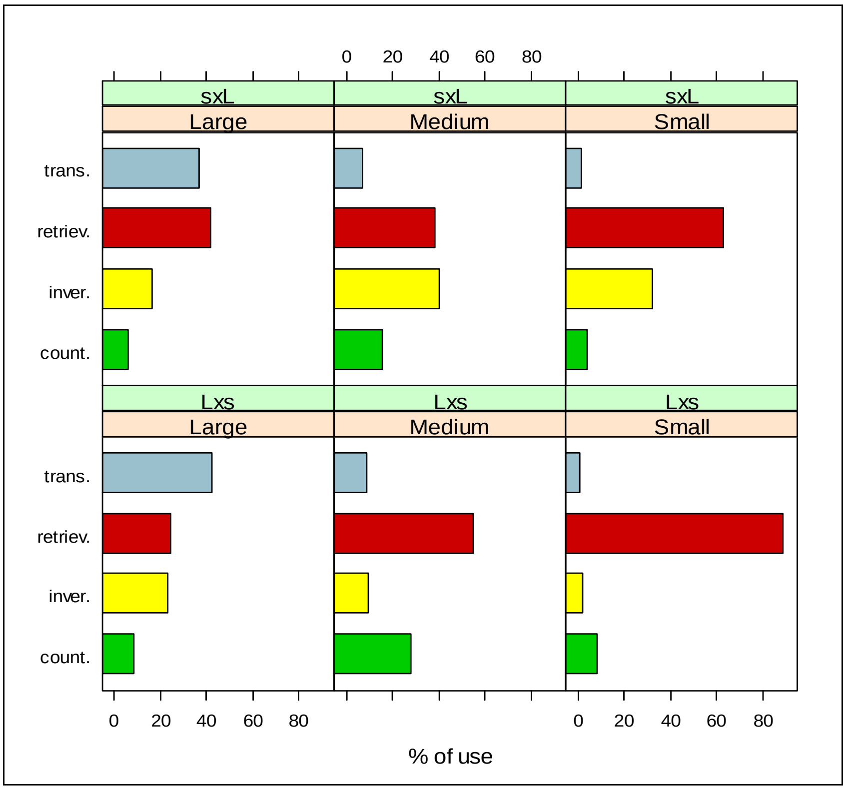


**Figure 2.2: percent of use of the procedures for multiplication in each experimental cell; trans: transformation, retriev: retrieval, inver: inversion, count: counting (Italians group).**

For each procedure ("other" excluded) we calculated, in each experimental cell given by the two factors size (large, medium, small) and order (L×s, s×L), the percent of use reported by the participants. For each participants, we used only the problems that were solved correctly in the chronometric task (92%) to the end to have a measure of the procedure in the same set of problems. We also excluded the problems in which the participants made an error in the self-report (only 9 errors on 1344 multiplication) and we did not considered the "other" strategy (the participants reported "other" only 7 times over 1344). The self-report results mirrored the interaction between size and order emerged in the RTs data.

| | % of retrieval | | | | % of transformation | | |
|---|---|---|---|---|---|---|---|
| | large | medium | small | | large | medium | small |
| Lxs | 23.9 | 54.9 | 88.9 | Lxs | 41.7 | 8.7 | 0.7 |
| sxL | 41.6 | 38.0 | 63.2 | sxL | 36.6 | 6.7 | 1.5 |
| | % of counting | | | | % of inversion | | |
| | large | medium | small | | large | medium | small |
| Lxs | 8.3 | 27.6 | 8.1 | Lxs | 22.9 | 9.3 | 2.2 |
| sxL | 5.9 | 15.2 | 3.8 | sxL | 15.8 | 40.0 | 32.3 |

**Table 2.1: the percent of use of the procedures in each experimental cell for multiplication (Italians group).**

As showed in table 2.1, for the large problems the participants reported more often retrieval in the s×L order with respect to L×s order, whereas in medium and small problems they reported to use more often retrieval in the L×s order with respect to s×L order. In other words, participants reported to use more often retrieval in the order they solved faster in the chronometric task. Transformation was mainly used to solve the large problems and did not show any strong difference between the two orders. The counting procedure was mainly used in the medium problems, and it was more used in the L×s order than for the s×L order, that is the order solved faster by the participants. Given the aim of the present study the most interesting debriefing variable is inversion and it relation with retrieval. The participants reported more often inversion for the order in which they reported less often retrieval. Moreover, inversion showed strong differences between the two orders in the medium and small problems than in the large problems. The participants reported to use much more often inversion in the s×L order than L×s order in small and medium problems, that is the participants reported more often inversion for the orders solved slower in the chronometric task. The difference for the large problems was smaller than for the medium and small problems, and it showed that this procedure was reported more often for the order solved slower in the chronometric task (that is the L×s order).

***Additions***

Each of the 24 participants had to solve 56 additions (from 2+2 to 9+9, tie problems excluded). Participants made errors on 3% of the trials, 40 errors on 1344 trials (56 problem × 24 participants). A two-way repeated measures ANOVA with size (small, medium, and large) and order (L+s and s+L) as within factors was conducted on the accuracy for the addition problems. The ANOVA revealed only a main effect of the size, $F(2,46)=7.72$, $\varepsilon_{GG}=0.96$, $p<0.01$. Post-hoc analysis revealed that the participants made more errors in the large condition (97% of correct answer) than in small condition (99%), $t(23)=-2.91$, $p<0.01$. Moreover, the participants made more errors in the medium condition (96%) than in the small condition, $t(23)=-4.09$, $p<0.001$. Neither the order factor nor the interaction reached the significance level.

The analysis of the RTs of the two keys showed a correlation of 0.97. Therefore, we analysed only the RTs associated with the first key pressed. A two-way repeated measures ANOVA with size (small, medium, and large) and order (L+s and s+L) as within factors was conducted on the RTs of the first key for the addition problems (see figure 2.3). The ANOVA revealed the significant main effect of the size, $F(2,46)=52.4$, $\varepsilon_{GG}=0.63$, $p<0.001$. Post-hoc comparison revealed that the participants responded faster in the small condition (1062 ms) than in both medium condition (1234 ms) ($t(23)=4.93$; $p<0.001$) and in large condition (1601 ms) ($t(23)= 7.82$; $p<0.001$); and that they responded faster medium condition than in the large condition, $t(23)= 7.32$; $p<0.001$. The ANOVA revealed also a main effect of order,

$F(1,23)=8.79$, $p<0.01$. Post-hoc comparison revealed that the participants responded faster to the L+s order (1238 ms) than the s+L order (1307 ms), $t(23)= -2.78$, $p<0.05$. Moreover, the ANOVA revealed also a significant order by size interaction, $F(2,46)=7.2$, $\varepsilon_{GG}=0.78$, $p<0.05$. Post-hoc comparison revealed that the L+s order was solved faster than the s+L order in large and medium problems: with large problems the L+s order (1521 ms; e.g., 8+7) was solved faster than the s+L order (1683 ms; e.g., 7+8), $t(23)= -3.11$, $p<0.01$; with medium problems the L+s order (1199 ms; e.g., 7+3) was solved faster than the s+L order (1268 ms; e.g., 3+7), $t(23)=-2.42$, $p<0.05$.

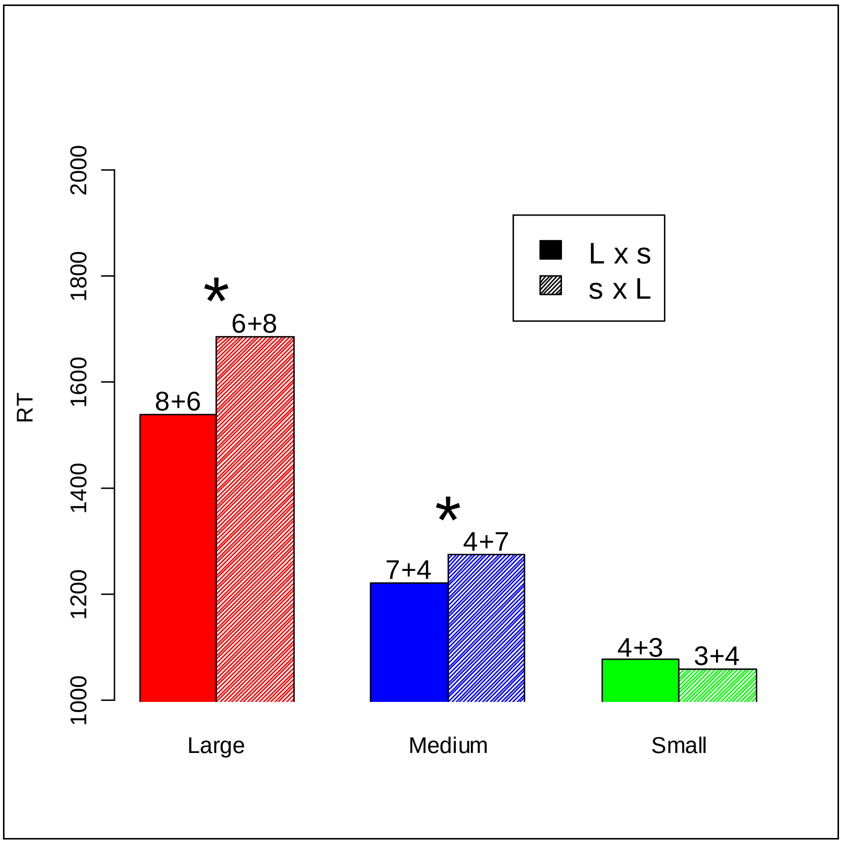


**Table 2.3: RTs as function of size and order of the operands for the addition problems (Italians group), * p<0.05.**

For the addition trials, we analyzed the self-report task by aggregating the percent of each procedure in the same cells as in multiplication (figure 2.4 and table 2.2). For each participants, we used only the problem that were correctly solved in the chronometric task (97%) to the end to have a measure of the procedure in the same set of problems. We also excluded the problems in which the participants made an error in the self-report (16 errors over 1344 addition trials) and we did not considered the "other" strategy (the participants reported "other" 47 time over 1344 trials).

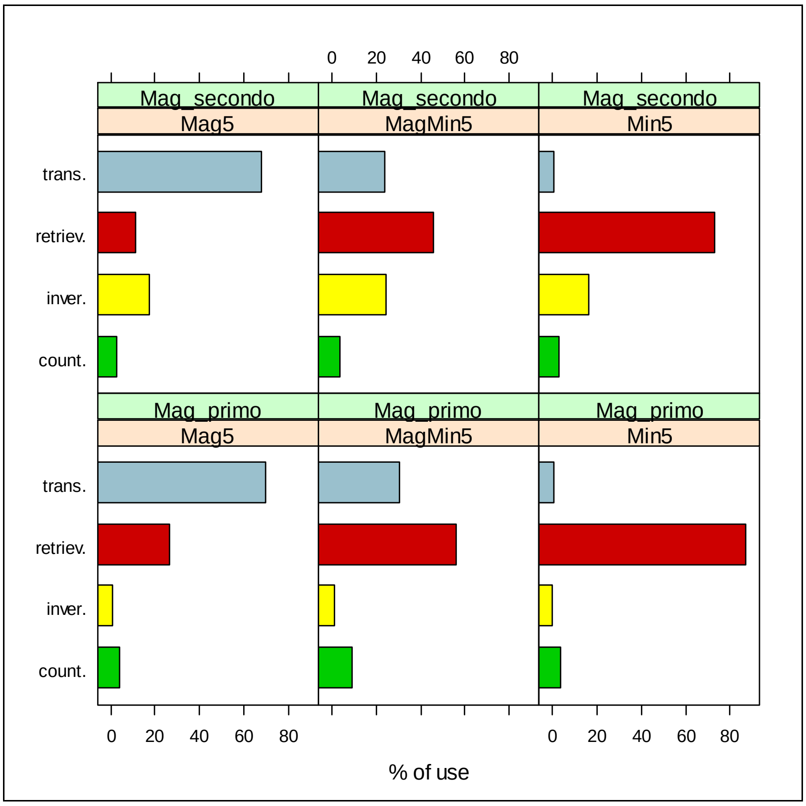


**Figure 2.4: percent of use of the procedures for addition in each experimental cell; trans: transformation, retriev: retrieval, inver: inversion, count: counting (Italians group).**

The percent of use of retrieval and inversion procedures was consistent with the RTs of the chronometric task. As showed in table 2.2, the participants reported to use more often retrieval in the L+s order (solved faster in the chronometric) than in the s+L order, whereas they reported more often inversion in the s+L order (solved slower) than in the L+s order. For transformation and counting were more often reported for the L+s order than for the s+L order, though for these procedures the differences between the two orders were smaller than for retrieval and inversion. More important, differently than for multiplications, the participants reported to use inversion almost only to solve the s×L order regardless the size of the problem. Namely, inversion was used to solve only the problems in the order that takes more time to be solved in the chronometric task.

| | % of retrieval | | | | % of transformation | | |
|---|---|---|---|---|---|---|---|
| | large | medium | small | | large | medium | small |
| L+s | 25.4 | 56.9 | 87.2 | L+s | 70.0 | 29.6 | 0.7 |
| s+L | 11.2 | 45.6 | 73.1 | s+L | 67.2 | 23.2 | 0.7 |
| | % of counting | | | | % of inversion | | |
| | large | medium | small | | large | medium | small |
| L+s | 3.8 | 9.8 | 3.8 | L+s | 0.8 | 0.9 | 0.0 |
| s+L | 1.6 | 3.1 | 3.0 | s+L | 17.6 | 25.2 | 16.4 |

**Table 2.1: the percent of use of the procedures in each experimental cell for addition (Italians group).**

### 2.2.3 Discussion

We found an operands-order effect for additions and an interaction between operands-order and size for multiplications. In the chronometric task for the multiplication trials the L×s order was solved faster than the s×L order in the small and medium size conditions; whereas the s×L order was solved faster in the large problems. The self-report of the participants was consistent with the RTs results. In other words, the retrieval (supposedly the fastest procedure) was reported more often in the order solved faster (s×L in the large size condition, L×s in the medium and small size conditions). On the contrary, inversion was reported more often for the order solved slower (L×s in the large condition, s×L in the medium and small conditions). In the chronometric task for the additions the L+s order was solved faster than the s+L order. Also in this case, the self-report was consistent with the RTs analysis. The retrieval (supposedly the fastest procedure) was reported more often in the order solved faster (L+s). On the contrary, inversion was reported more often for the order solved slower (s+L).

The *interacting neighbors model* (Verguts and Fias, 2005) assumes that only one of the two operands-orders is stored as multiplication fact. However, we found that the order of the operands interacted with the size of the problems. We interpret this interaction as an evidence that the order stored could vary as a function of the size, that is, assuming the architecture of the Vergut and Fias (2005) model, one has to assume that the s×L order is stored for large problems and the L×s order is stored for the medium and small problems. A possible explanation of this order by size interaction can be attempted within the frame that assumes that the multiplication facts memory is shaped during the childhood. Two factors must be considered: the order of acquisition of the problems, and a possible reorganization of the memory due to the repeated use of non-retrieval procedure used to solve the multiplication (see Butterworth et al., 2003). In the Italian education system the problems in the s×L order are taught before the ones in the L×s order. Therefore, the order s×L should be acquired before and more practiced than the inverse order. If the s×L order is acquired before the L×s order, why is the s×L order privileged only for the large problems whereas for the medium and the small problems the privileged order is L×s? This interaction can be explained by the intervention of a reorganization process that reshape the multiplication facts memory according to the non-retrieval procedures adopted by the children during the acquisition of the multiplication table. In the study of Butterworth et al. (2003), Italian children (8, 9, and 10 years old) showed to solve faster the multiplication in the L×s order (in the study were tested only the problems form 2×2 and 6×5). The Authors assumed that their results were due to a reorganization of the multiplication facts memory produced by the use of particular non-retrieval procedures. Namely, the children use repeated addition (e.g., 7×3=7+7+7) and table sequences (e.g., 7×3=7, 14, 21) to solve multiplication problems that are not yet stored in memory or which representation is not yet strongly acquired. The L×s order is easier/faster to solve with these kind of procedures, whereas the s×L order is likely to be reordered (e.g., 3×7 has to be reordered in 7×3) to be efficiently solved with these procedures. Therefore, the L×s order could became the privileged one, and the s×L order could often be solved relying on the L×s order. This asymmetry between the two orders could arrange the memory in way that the L×s order becomes the one stored in the associative network that contains arithmetic facts. Again, the reorganization principle cannot alone explain our results because of it predicts that we would have to found a main effect of order, that is the L×s order would be solved faster than the inverse order for all problems. We propose that these two factors (order of acquiring and reorganization) work together to shape the multiplication facts memory. In fact, both repeated addition and table sequence are really efficient with small and medium problems, but they are not for the large problems. It is highly unlikely that children (or adults) use repeated addition or table sequence procedures to solve large multiplication. For example, given the problem 7×8, the procedure 7+7+7+7+7+7+7+7 (repeated addition) or 7, 14, 21, 28, 35, 42, 49, 56 (table sequence) are very inefficient and difficult, and the use of the other order (8×7=8+8+8+8+8+8+8) is as well as inefficient and difficult. Therefore, the use of procedures can reorganize the small and medium problems (giving an advantage to the L×s order) but has no effect on the large problems that are shaped by the only order of acquisition, maintaining the original advantage for the early learned s×L order.

The above explanation accounts for RTs difference only in term of retrieval within arithmetic facts memory since it is likely that this procedure should be the preferred for rather competent adults when speeded solutions are required, we cannot however exclude that, at least part of these differences, are also due to the use of the non-retrieval procedure the participant reported they used in the unspeeded task where maximal accuracy was required. There are some evidences in fact that adults solve multiplication by using both retrieval and non-retrieval procedures (see chapter 1). The asymmetries we found in the self-report were consistent with the results of the RTs, therefore it is not unlikely that also the non-retrieval procedures could have a role in the differences found between the two orders in the chronometric task. Since we asked to the participants to report the procedures they were using during the self-report and not to try to remember the procedures used in the chronometric task, we cannot assume that the participants used the same procedures to solve the problems in the two tasks. Moreover, in the chronometric task the speed was stressed whereas in the self-report it was not and only the accuracy was stressed. Smith-Chant & LeFevre (2003) showed that instruction of speed can bias the procedures reported by the participants, that is they reported more often to use retrieval. Therefore, it was likely that our participants relied more on retrieval in the chronometric task than in the self-report. So, what can the self-report tell us about the effect of the order of the operands we found in RTs? The self-report could give information about what the participants do when they have to solve problems in a context in which it is required only to be "as accurate as possible". In the chronometric task the order of the operands affected the RTs also for small problems (e.g. 4x2), that are very likely to be solved by using retrieval procedures only. Therefore, we are confident to assume that the RTs effects we found are at least partially due to the retrieval process in terms of speed of access to arithmetic facts in long term memory. Nevertheless, the self-report suggests that also the non-retrieval procedures play a role in the effect we found. The self-report results suggest that the order of the operands can affect the RTs in two ways: 1) one of the two orders is easier/faster to retrieve because of it is the stored one; 2) the two order can be solved with different procedures that could be require different time to be performed.

With respect to additions the *COMP model* (Butterworth et al., 2001) assumes that only the L+s order is stored in the addition facts memory. Butterworth and Colleagues asserted than the preference for the L+s order could be due to the fact that the children use non-retrieval procedures that are simple to solve with the L+s order. Our results are consistent with the this hypothesis. We found that the L+s order is solved faster than the s+L order in the large and small problems. Like in multiplication, our result can be explained by the reorganization of the addition facts memory due to the use of non-retrieval procedures when the children learn the addition. The L+s order is easier to solve with non-retrieval procedures, therefore it could be privileged and then stored in memory. However, unlike multiplication, in addition the non-retrieval procedures could reshape also the large problems because of the use of procedures like “counting for the larger” could be efficient with large problems as well as with smalls. Therefore, the L+s order could be privileged during the acquisition and then stored in memory.

The self-report showed that also for addition problems the use of non-retrieval procedures could play a role in the differences we found between the two orders. Like for the effect found with multiplication, we hypothesized that both retrieval and non-retrieval procedures could work together to generate the operands-order effect we found. On the one hand, the L+s order could be easier/faster to retrieve than the s+L order, because of the former is the stored one. On the other hand, when the participants adopted non-retrieval procedures the L+s order could be easier to solve than the inverse order.

The explanation described above is based on the architectures of the *interacting neighbors model* and the *COMP model*. However, we propose a second possible explanation based on the *network retrieval model* and *network interference model* (see chapter 1), which assume that both orders of the operands are stored as arithmetic facts. We suppose than the use of non-retrieval procedures plays a important role in either producing or at least determining (by shaping the arithmetic facts memory) the order by size interaction we found. Nevertheless, the role of the non-retrieval procedures is assumed to be different for additions and multiplications. Some evidence suggests that the use of non-retrieval procedures is more common in addition whereas multiplication mainly solved by means of retrieval (Campbell & Xue, 2001). If multiplication is mainly solved by retrieval the order by size interaction we found could depend on the spreading of activation inside the architecture of the multiplication facts memory rather than the learning experience or the use of non-retrieval procedures. There are many evidence that suggest that the presentation of a problem automatically activates the closer problems (Galfano, Rusconi, &Umiltà, 2003; Galfano Penolazzi, Vervaeck, Angrilli, & Umiltà, 2009; Niedeggen & Rösler, 1999; Rusconi, Galfano, Speriani, Umiltà, 2004; Rusconi, Galfano, Rebonato, & Umiltà, 2006). For example, when the problem 7×8 is presented the problems 7×6, 7×7, 7×8, 6×8, 9×8, and so on are activated. Above we discussed the effect of the order of the operands by assuming that only one order is stored in memory as multiplication fact, coherently with the architecture proposed by Verguts and Fias (2005). However, the order by size interaction could be explained also within frames that assume that both order are stored in memory. The use of non-retrieval procedures in the childhood could affect the association strengths between the results of the problems. In fact, the use of procedures like repeated addition or table sequence could produce an asymmetry in the association between the results of the problems (i.e., the multiples of the multiplication table). Both repeated addition or table sequence procedures could reinforce the association between a multiple and the following one more than the association between a multiple and the previous one. In the repeated addition procedure the multiplication is transformed in a series of addition, where the intermediate results are the sequence of multiples of the operand used as base for the additions. For example, when the problem 7×4 is solved with repeated addition, it is transformed in the sequence (((7+7)+7)+7)=((14+7)+7)=(21+7)=28, where the results 14, 21, and 28 are the multiple of the number 7 in the small to large direction. Therefore, when a child use this procedure, a multiple could be used to identify the following one in the table, that is 14 is used as base to identify 21 (after 7 is added). The association between a multiple and the following one could be reinforced also by the use of a table sequence procedure. For example, in the procedure 3×4= 7, 14, 21, 28 each multiple could be used as a cue to find the following one. This two procedures could reinforce exclusively the association between a multiple and the following one. Therefore, once activated a result (e.g., 21) could spread more activation in the forward direction (e.g., 28, 35) than in the backward direction (e.g., 14). Moreover, these two procedures could reinforce the association between an operand and the begin of its table because of the first problems and result of the table are often used as starting point to perform the repeated addition and the table sequence procedures. It is assumed that when a problem is presented, the activation inside the arithmetic facts memory spreads also to the closer problems (see chapter 1). In our experiment the presentation of the operands were sequential, that is the first operand were presented 600 ms before the second operand. When it is presented a small or medium problem the first operand (e.g., 7) could strongly activate the problems associated with the begin of its table (e.g., 7×2, 7×3, 7×4). Therefore, the difference between 7×3 and 3×7 is that when the former is presented first operand (7) could activate the problem 7×2, 7×3, 7×4 before the second operand (3) is presented; whereas 3×7 is presented the first operand activate the problem 3×2, 3×3, 3×4 before the second operand is presented (7). Therefore, when the second operand is presented for the problem 7×3 the problem associated with the result is already activated (7 has been presented 600 ms before and it has activated 7×2, 7×3, 7×4). On the contrary, when the second operand of the problem 3×7 is presented the problem associated with the result is not activated (3 has been presented 600 ms before and it has activated 3×2, 3×3, 3×4, but only weakly 3×7). This hypothesis could explain why the L×s order is solved faster than the s×L order. However, the result with the large problems cannot directly be explained by this hypothesis. In fact, the large problems are equally distant from the beginning of the table. Nevertheless, the tie effect (the fact that tie problems are solved faster than the other problems with similar size) could produce the inversion of the order effect with the large problems. Since the tie problem are solved faster, the association between the operands and the tie result could be stronger than for the other results (see for example *network retrieval model* in chapter 1). For example, 7 could be more strongly associated with 49 than with 42, 56, 63. Therefore, when a large problem is presented the first operand could strongly activate the tie problem before the presentation of the second operand and the activation of the tie could be stronger than the activation of all the other problems. Once the tie problem is activated it could spread more activation in the forward direction than in the backward direction for the reasons above described, due to the use of repeated addition and table sequence procedures. Both the stronger activation of the tie and the stronger forward activation spreading can explain the advantage for the s×L with respect to the L×s order in large problems. For example, when the first operand (7) of the problem 7×8 is presented, the tie problem (7×7) is activated and spreads activation mainly to the forward problems (7×8). When the first operand (8) of the problem 8×7 is presented, the tie problem (8×8) is activated and spreads activation mainly to the forward direction (8×9). Therefore, when the second operand is presented for the problem 7×8 the problem associated with the result is already activated (7 has been presented 600 ms before and it has activated 7×7, and then 7×8 and 7×9). On the contrary, when the second operand of the problem 8×7 is presented the problem associated with the result is only weakly activated (8 has been presented 600 ms before and it has activated 8×8, and then 8×9, but only weakly 8×7). Therefore, with the problem 7×8 the result could be activated before the second operand is presented, whereas with the problems 8×7 the result would not be activated. This could explain why the s×L order is solved faster than the L×s order in the large problem.

Summarizing, we found that operands-order affects speed of solutions for both additions and multiplications. For additions the advantage for the L+s order can be explained within the COMP model in terms of reorganization of the addition facts memory. For multiplications the pattern is rather surprising since at our knowledge such an inversion of order preferences across problem sizes has never been reported in the literature and thus it is hard to be explained by any current model, independently it assumes that only one order or both are stored in the arithmetic facts memory. Despite part of the effect could be due to non-retrieval procedure we were able to offer two distinct explanations of the interaction in terms of speed of retrieval from memory (that is somewhat simpler than assuming the effect is driven only by non-retrieval procedures). One is framed within models that assume only one order is stored in arithmetic facts memory (Verguts & Fias, 2005) and depends on both order of acquiring and reorganization of the memory, the other is framed within models that assume both orders are stored in memory and the strength of the association with the correct result depends on the usage of non-retrieval procedures that we used to hypothesise that activation of multiples spreads stronger from left to right starting from both the beginning of the table and from the tie. Crucially, the two explanation differ with respect to the role of the order of acquisition during childhood learning. The former strongly depends on this, capitalizing on the Buttterworth et al. (2003) original idea of reorganization, the latter is completely independent from the order of acquisition. In the next experiment we try to disentangle between these two hypotheses by testing a population where the order in which the problems are acquired is inverse with respect to the Italians.

## 2.3 EXPERIMENT 2: ENGLISH PARTICIPANTS

We decided to test a population that learn the multiplication table in the inverse order with respect to the Italians since one of the explanation given for the results of the previous experiment predicts a different pattern as a function of learning order. In England the name of the table is in second position, whereas in Italian is in first position. For example, the 2-table is 2×1, 2×2, 2×3, ..., 2×9 in Italian education system, and 1×2, 2×2, 3×2, ..., 9×2 in the English system. Therefore, in Italy children learn s×L before L×s (e.g. “2 per 9” before “9 per 2”); whereas in England children learn L×s before s×L (“9 times 2” before “2 times 9”). Given the explanation of the interaction between size and order of experiment 1 (this chapter), the preferred order for large multiplication problems should be due to order in which the multiplication are learned, and therefore we expect opposite preferences in the two groups, Italians and English.

In fact, if the multiplication facts memory is shaped by reorganization and order of acquiring we expect an advantage for the L×s order in the large problems, because of in England the L×s order is learned before the s×L order. Therefore, we should not found any interaction order by size but a main effect of the order, that is the L×s order should be solved faster than the s×L order regardless the size of the problems. If the effect we found with multiplication in the Italian group is due to the asymmetric forward activation spreading in the multiplication facts memory we expect to find the in the English group same result found in the Italian group, that is an order by size interaction.

For addition there should not be differences between Italians and English, because of the two hypotheses have the same prediction, that is an advantage for the L+s order.

### 2.3.1 Method

**Participants.** Twenty-eight native English-speaking students (13 females; mean age: 26, sd: 6.16) from the University College of London (UCL) participated in the experiment as volunteers. All participants had normal or corrected-to-normal vision. This experiment was approved by the Ethic committee of UCL. Six participants were excluded because of low performance in the experiment.

**Material, Procedure, and Data analysis.** The material, the procedure, and the data analysis were was exactly the same as in experiment 1.

### 2.3.2 Results and Discussion

Each participant had to solve 56 multiplications (from 2×2 to 9×9, tie problems excluded). Participants made errors on 8% of the trials on average, 96 errors on 1232 trials (56 problem x 22 participants). A two-way repeated measures ANOVA with size (small, medium, and large) and order (L×s and s×L) as within factors was conducted on the accuracy for the multiplication problems. The ANOVA revealed only a main effect of the size, $F(2,42)= 26.68$, $\varepsilon_{GG}=0.71$, $p<0.001$. Post-hoc analysis revealed that the participants made more errors in the large condition (81% of correct answer) than in both medium (94%) and small (97%)

condition, t(21)=-5.65, p<0.001 and t(21)=-5.74, p<0.001 respectively. The order factor did not reach significance level but the interaction between order and size was significant, F(2,42)=3.92, $\varepsilon_{GG}$=0.63, p<0.05. Post-hoc analysis revealed that the participants tended to make more errors in the L×s order (77%) than in the s×L order (86%), t(21)=2.08, p<0.1. For the large multiplication problems the English participants seem to have better performance for the s×L order, that is the same order privileged by the Italian group. in England the L×s order is taught before the s×L order, therefore, this result is consistent with the asymmetric forward spreading activation hypothesis. However, the accuracy differences are based on very few errors (the participants made errors on only the 8% of the trials) and the post-hoc analysis reveal only a tendency toward significant for the large problems. Hence, we think this result is not strong enough to discriminate between the two hypotheses.

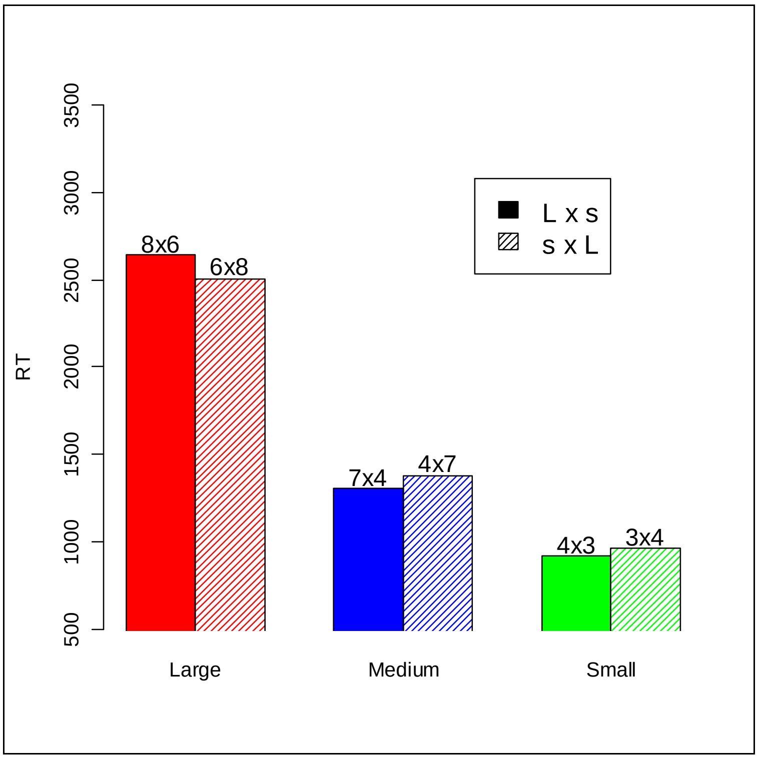


**Figure 2.5: RTs as function of size and order of the operands for the multiplication problems (English group).**

The analysis of the RTs of the two keys showed a correlation of 0.96. Therefore, we analysed only the RTs associated with the first key pressed. A two-way repeated measures ANOVA with size (small, medium, and large) and order (L×s and s×L) as within factors was conducted on the RTs of the first key for the multiplication problems (see figure 2.5). The ANOVA revealed the a significant main effect of the size, $F(2,42)= 30.02$, $\varepsilon_{GG}=0.53$, $p<0.001$. Post-hoc comparison revealed that the participants responded faster in the small condition (937 ms) than in both medium condition (1346 ms) ($t(21)=-7.2$; $p<0.001$) and in large condition (2198 ms) ($t(21)=5.89$; $p<0.001$); and that they responded faster in the medium condition than in the large condition ($t(21)=4.87$; $p<0.001$). Neither the order factor nor its interaction with size were significant.

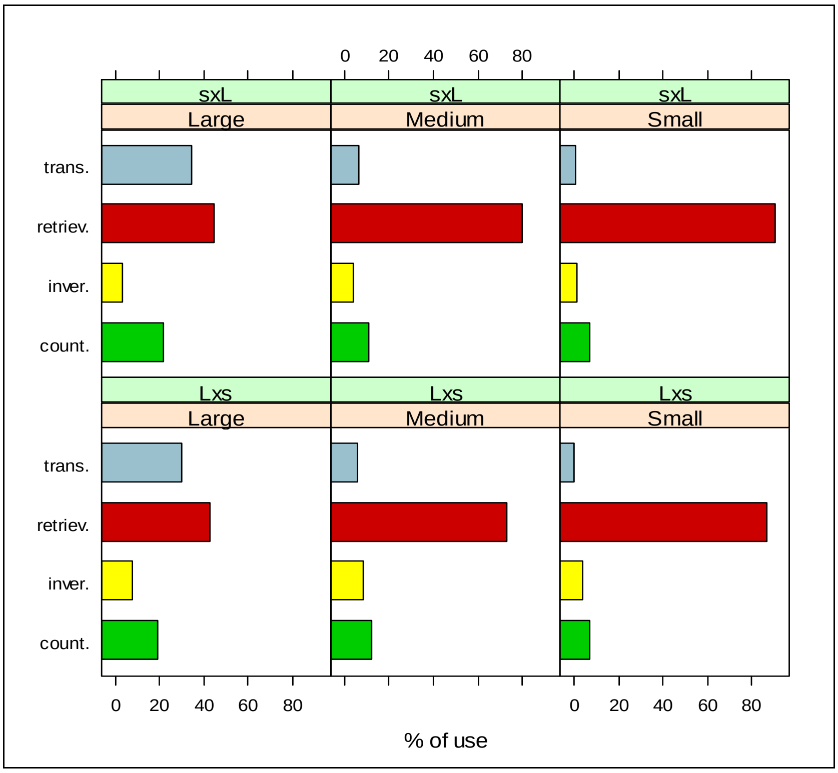


**Figure 2.6: percent of use of the procedures for multiplication in each experimental cell; trans: transformation, retriev: retrieval, inver: inversion, count: counting (English group).**

For the multiplication trials, we qualitatively analyzed the self-report task by aggregating in the same cells as in the ANOVAs the percent of use of each procedure (figure 2.6 and table 2.3). For each participants, we used only the problems that were solved correctly in the chronometric task (92%) to the end to have a measure of the procedure in the same set of problems. We also excluded the problems in which the participants made an error in the self-report (only 29 errors on 1232 multiplication) and we did not considered the “other” strategy (the participants reported “other” 28 times over 1232 multiplication trials of the self-report task). Unlike for the Italians, the self-report results showed very small differences between the two orders of the operands. However, contrary what we expected, the participants reported little more often retrieval for the s×L order than for the L×s order; whereas inversion was reported little more often for the L×s order than for the s×L order. Therefore, unlike in the Italian group the self-report did not show any strong asymmetry.

| | % of retrieval | | | | % of transformation | | |
|---|---|---|---|---|---|---|---|
| | large | medium | small | | large | medium | small |
| L×s | 42.3 | 73.1 | 87.1 | L×s | 30.1 | 5.8 | 0.0 |
| s×L | 44.2 | 79.7 | 90.8 | s×L | 34.2 | 6.4 | 0.8 |
| | % of counting | | | | % of inversion | | |
| | large | medium | small | | large | medium | small |
| L×s | 18.7 | 12.3 | 6.8 | L×s | 7.3 | 8.2 | 3.8 |
| s×L | 21.7 | 11.0 | 6.9 | s×L | 3.3 | 3.8 | 1.5 |

**Table 2.3: the percent of use of the procedures in each experimental cell for multiplication (English group).**

Unfortunately, for both chronometric task and self-report task the results with multiplication problems do not show any relevant effect of the order of the operands. Therefore, they cannot be used to disentangle between our two hypotheses.

Each of the 22 participants had to solve 56 additions (from 2+2 to 9+9, tie problems excluded). Participants made errors on 3% of the trials, 36 errors on 1232 trials (56 problem × 24 participants). A two-way repeated measures ANOVA with size (small, medium, and large) and order (L×s and s×L) as within factors was conducted on the accuracy for the multiplication problems. The ANOVA revealed only a main effect of the size, $F(2,42)=4.05$, $\varepsilon_{GG}=0.78$, $p<0.05$. Post-hoc analysis revealed that the participants tended to make more errors in the large condition (95% of correct answer) than in small condition (98%), $t(21)=-2.32$, $p<0.1$. Neither the order factor nor the interaction reached the significance level.

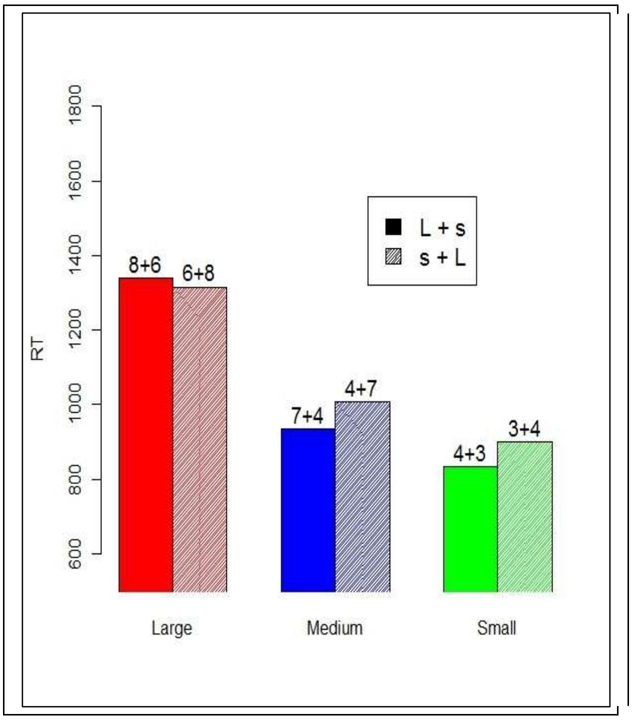


**Figure 2.7: RTs as function of size and order of the operands for the addition problems (English group).**

The analysis of the RTs of the two keys showed a correlation of 0.97. Therefore, we analysed only the RTs associated with the first key pressed. A two-way repeated measures ANOVA with size (small, medium, and large) and order (L+s and s+L) as within factors was conducted on the RTs of the first key for the addition problems (see figure 2.7). The ANOVA revealed the a significant main effect of the size, $F(2,42)= 16.03$, $\varepsilon_{GG}=0.57$, $p<0.001$. Post-hoc comparison revealed that the participants responded faster in the small condition (866 ms) than in both medium condition (965 ms) ($t(21)=2.93$; $p<0.01$) and large condition (1308 ms) ($t(21)=4.17$; $p<0.001$); and that they responded faster in the medium condition than in the large condition, $t(21)=3.88$; $p<0.001$. The ANOVA revealed also a trend toward the significance for order, F(1,21)=3.74, p<0.1. The participants tended to respond faster to the L+s order (993 ms) than the s+L order (1037 ms).

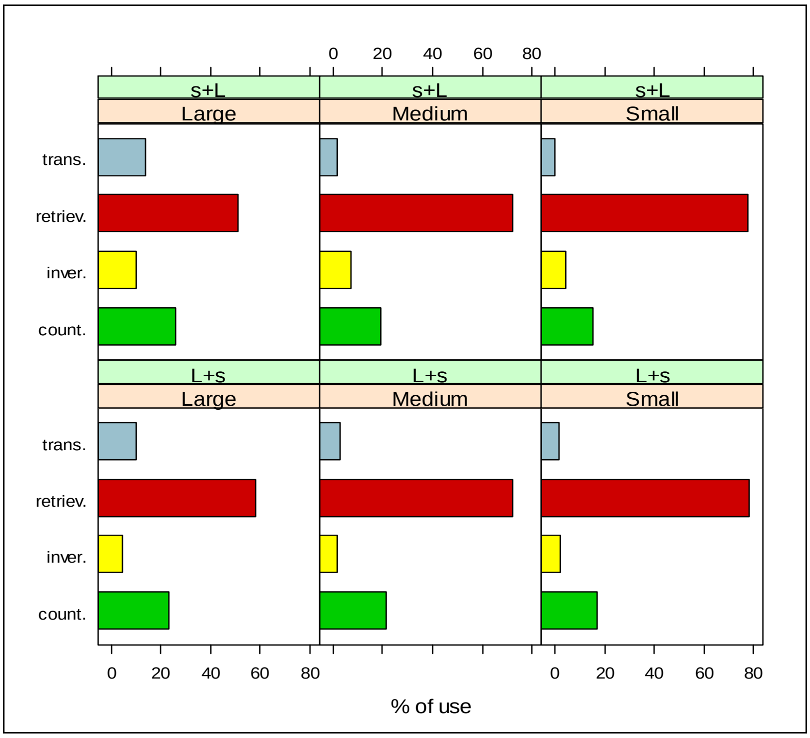


**Figure 2.8: percent of use of the procedures for addition in each experimental cell; trans: transformation, retriev: retrieval, inver: inversion, count: counting (English group).**

For the addition trials we also analyzed the self-report task by aggregating the frequency of report of each procedure in the same cells as done for multiplications (figure 2.8 and table 2.4). For each participants, we considered only the problems that were solved correctly in the chronometric task (97%) to the end to have a measure of the procedure in the same set of problems. We also excluded the problems in which the participants made an error in the self-report (10 errors over 1344 addition trials) and we did not considered the “other” strategy (the participants reported “other” 33 time over 1344 trials in the self-report task). As for the multiplication the differences between the two orders are smaller in the English group than in the Italian group. However, the participants reported little more often retrieval for the L+s order than for the s+L order; whereas they reported inversion little more often for the s+L order than for the L+s order. Like for multiplication, the self-report did not show any strong asymmetry between the two orders.

To resume addition results, the participants tended to solve faster the L+s order than the s+L order. Moreover, they reported to use retrieval a little more often in the L+s order than in the s+L order, and inversion a little more often in the L+s order than in the s+L order. The results for addition problems show a tendency consistent with the results of the Italian group. However, as for multiplication, these results are weaker than in the Italian group.

| | % of retrieval | | | | % of transformation | | |
|---|---|---|---|---|---|---|---|
| | large | medium | small | | large | medium | small |
| L+s | 58.0 | 72.2 | 78.3 | L+s | 9.9 | 2.9 | 1.5 |
| s+L | 50.7 | 72.0 | 77.9 | s+L | 13.6 | 1.7 | 0.0 |
| | % of counting | | | | % of inversion | | |
| | large | medium | small | | large | medium | small |
| L+s | 22.9 | 21.5 | 17.0 | L+s | 4.6 | 1.7 | 2.3 |
| s+L | 25.8 | 18.8 | 15.3 | s+L | 9.8 | 6.8 | 4.6 |

**Table 2.4: the percent of use of the procedures in each experimental cell for addition (English group).**

The results of the experiment 2 for both addition and multiplication with English participants do not show any relevant effect of the order in both RTs and self-report. This null results could be due to the different competences of the participants within English group[6].

## 2.4 GENERAL DISCUSSION

The aim of the present study was to evaluate if the order of the operands could affect the RTs and the selection of procedures (retrieval and non-retrieval) in both multiplication and

[6] In analyses not reported here we did a median split analysis dividing the participants in two different group according to the median of the RTs (low skill participants: with RTs below the median RTs of the whole group; high skill participants: with RTs below the median RTs of the whole group). These analyses showed that the privileged order varied in the two groups. However, the number of participants in each groups was too small to make inferences.

addition. To this end, we adopted a chronometric production task with speed and accuracy stressed and a self-report task with only accuracy stressed. In the chronometric task we found in the Italians group an operands-order by size interaction in both multiplication and addition. The participants solved faster the multiplication in L×s order than in s×L order in small and medium problems; whereas they solved faster the multiplication in s×L order than in L×s order in large problems. Moreover, the addition in L+s order were solved faster than in s+L order in medium and large problems. Finally, in both multiplication and addition, we found that retrieval was more often reported for the problems in the order solved faster, whereas inversion was reported more often for the problems in the order solved slower.

In this study we provide clear evidence that the order of the operands can affect the RTs in the production of the result of multiplication and addition problems. We propose two hypothesis to explain the results. The first hypothesis (hereafter, *reorganization hypothesis*) states that only one arithmetic facts is stored as arithmetic facts, and that which is the stored order is determined by the order is which the problems are learned and by the reorganization due to the use of non-retrieval procedures during when the problems are learned. The second hypothesis (hereafter, *asymmetry hypothesis*) states that both order are stored as arithmetic facts and the effect of order is due to the activation spreading inside the multiplication facts memory.

The results from the self-report suggest that the procedures selection could play a relevant role in the operands-order effect we found in the RTs. The self-report used in this study could have biased the procedures reported by the participants (Kirk & Ashcraft, 2001), since we asked the participants to select between a range of procedures proposed by us. Nevertheless, the relevant result is that the participants reported different procedures for the commuted pairs. This result suggest that the individuals could use different procedures to solve the two orders of the operands. However, since we are not sure that the participants used the same procedures in both chronometric and self-report task, we cannot exclude that the order by size interaction in the chronometric task was mainly due to retrieval and to the asymmetric activation spreading (*asymmetry hypothesis*).

The *reorganization hypothesis* is consistent with the *interacting neighbors model* (Verguts and Fias, 2005) and the *COMP model* (Butterworth et al., 2001), which state that only one order of the operands is stored as arithmetic fact for multiplication and addition, respectively. A fundamental process in this two models is the reordering of the operands when the presented problem is not in the stored order. According to these two models, the *reorganization hypothesis* explains the RTs differences between the two order of the operands by means of the operands reordering process. In fact, the reordering process should occur only if the presented problem is not in the stored order, and then this supplementary process could explain the difference found between the two orders. The *asymmetry hypothesis* explains the RTs differences in terms a different amount of activation of the result when the second operand is presented.

Crucially the two hypothesis differ in that the former (*reorganization hypothesis*) assumes a central role of learning order during acquisition of arithmetical knowledge and the latter (*asymmetry hypothesis*) does not. For these reason we tested a group of English speaker with the same paradigm. Two different predictions were developed: one on the basis of the *reorganization hypothesis* predicted an overall preference for the L×s order in English speakers, independent form size; the other on the basis of the *asymmetry hypothesis* predicted the same interaction of order and size found for Italians. The second experiment however gave a null result on RTs and an interaction for accuracy without a strong support from debriefing. In both cases this can be interpreted as a failure to replicate the effects of order in the solution of arithmetical problems and thus suggest further studies in order to confirm the interaction found in experiment 1, possible with different paradigms in order to gain at the same time a better understanding of the phenomena that we hope can be replicated.

A crucial point of the *asymmetry hypothesis* is that the tie problem spreads more activation in the forward direction than in the backward direction. Therefore, a second way to evaluate the *asymmetry hypothesis* within the same group of participants (Italian speakers) is to test for asymmetries in the amount of activation spread by the tie. The experiments reported in the next two chapters will test the assumption that the tie problems spread the activation asymmetrically. Moreover, in the experiment of the chapter 4, the two assumptions of the *asymmetry hypothesis* have been tested separately. Namely, we tested both if the activation generally spreads more in the forward direction than in the backward direction and if the activation spreading around the tie problems is asymmetric.

# Chapter 3

## Asymmetric activation spreading around the tie problems: matching and multiples tasks

## 3.1 INTRODUCTION

In this chapter we will test a critical assumption of the *asymmetry hypothesis* (see chapter 2). Namely, we aim to verify if the tie problems spread more activation to the result of the table in the forward direction than to the result of the table in the backward direction.

An important question about the architecture of the multiplication fact memory is which kind of association exists between the operands and the results. The *network retrieval model* (see chapter 1) states that the arithmetic facts memory includes three sets of nodes: one set for the first operand; one set for the second operand; one set for the corresponding result. For example, when the problem 7×8 is presented, the node 7 and the node 8 are activated in the sets of the first and the second operands respectively. Furthermore, in the set of the results the node 56 is activated. This model assumes that each operands pair is associated with a specific results. For example, the problem 6×4 and 3×8 activate two different result nodes (both corresponding to the number 24). According to this model the presentation of a number (an operand) could be able to activate its multiples (the result of the multiplication table associated with that number). Moreover, this model assumes that the result nodes are associated each other and spread activation to their neighbourhood. The *network interference model* (see chapter 1) assumes a similar association between operands and result, with the exception that the result nodes are unique regardless the problem. The operand nodes that share the result (e.g., 24 is the result of both 6×4 and 3×8) are associated with the same result node (e.g., the same result node 24 is activated by both 6×4 and 3×8). Moreover, the result nodes are associated with magnitude information that identify the approximate size of the problem (e.g., the problem 2×7 is associated with "small" and 8×9 with "large"). Like the *network retrieval model*, also the *network interference model* assumes that the result nodes are interconnected each other.

An relevant distinction between the two models regards the tie effect (i.e., the tie problem, e.g. 6×6, are solved faster than other problems with similar size). The *network retrieval* *model* explain the tie effect by assuming that tie problems are more frequent than the other problems with similar size and then the strength of association between operands and tie result is higher than between operands and non-tie problems. On the contrary, the *network interference model* assumes that the tie problems are easier to retrieve because of they are stored separately from the other problems. When a tie problem is presented the non-tie problem are weakly activated (since they are stored separately). Therefore, the presentation of a non-tie problem actives more competitors than the presentation of a tie problem. The smaller number of competitors produces an advantage in the identification of the result for the tie problems.

Both models agree that operand nodes are associated with the result nodes and that the result nodes are associated one to each other. There is a clear empirical evidence (Galfano et al., 2003; Galfano et al., 2009; Niedeggen & Rösler, 1999; Rusconi et al., 2004; Rusconi et al., 2006) that the presentation of a number automatically activates the multiples associated to that number (that is the results of the multiplication table of that number). However, the *asymmetry hypothesis* require that the tie and the non-tie problems are stored together and that they can spread activation to the other results of a table. Therefore, according to the *network retrieval model*, we assume that the tie effect is due to the higher association between operands and tie result. We hypothesize that the stronger association between operands and tie results is ascribable to the structural characteristics of the tie problems rather than to the frequency factor. In the non-tie problems the activation of the result is given by the contribute of both operands, whereas in the tie problems a single operand repeated twice has to activate alone the result. For example, the multiple 36 is the result of the tie problem 6×6, then the result 36 is associated only with the operand 6. Hence the activation of 36 is exclusively given by its association with the operand 6. However, the operand 6 is associated also with the other multiples that constitute its multiplication table. Therefore, to allow the memory to discriminate between the other multiples and the tie result, the association between an operand (e.g., 6) and its tie multiple (e.g., 36) has to be particularly strong, that is the operand activate the tie multiple more than the other non-tie multiples.

The fact that the presentation of a problem can activate both the result of the problem and as well the other multiples of the operands around the result is provided by various studies. For example, Niedeggen & Rösler (1999), in an ERPs study, adopted a verification task to investigate the spread of activation in the memory network that encodes the multiplication facts. The task of the participants was to verify if the proposed solution was correct or incorrect. The incorrect solution could be table related to one of the operands (e.g., 5×8=32, 24, or 16) or not related (e.g., 5×8=34, 26, or 18). Furthermore, the distance of the proposed solution from the correct result (e.g., 5×8=40) could be small (e.g., 32 or 34; for the not-related and related condition respectively), medium (e.g., 24 or 26), or large (e.g., 16 or 18). The Authors found a larger N400 for the incorrect trials with respect to the correct ones, and that this effect was modulated by the distance between the actual and the presented result, but only in the trials where the wrong response was in the table of one of the operands. The amplitude of the N400 effect was attenuated for the small and medium problems with respect to large problems. From studies on language processing, the amplitude of the N400 effect is supposed to be associated with the (semantic) relation between a preceding context (e.g. the operands) and the target (the proposed result). Namely the stronger is the association between the context and the target, the smaller is the amplitude of the N400 (Kutas, Van Petten, & Kluender, 2006). The results of this study has been interpreted as an evidence that the activation spreads from the actual result of the presented problem mainly to the multiples of the operands that are close to the actual result.

Two different hypotheses about the architecture of the multiplication facts network are consistent with the result of Niedeggen & Rösler (1999). The first hypothesis (indirect activation) claims that the product is activated by the operands and the multiples are indirectly activated via product, namely the activation inside the network spreads from the product (Galfano et al., 2003). In other words, two operands (e.g., 6×4) activate their product (e.g., 24) and then the activation spreads from the product to the closer members of the table of the two operands (e.g., considering the operand 4 the activation spreads from 24 to 20, 28, 16, 32, and so on). The second hypothesis (direct activation) claims that the operands can directly activate both the product and the multiples at the same time (Galfano et al., 2009). To disentangle between these two hypothesis, Galfano and Colleagues (2009) carried out a ERPs study adopting the number-matching task (LeFevre et al., 1988; see also Galfano et al., 2003; Rusconi et al., 2004; Rusconi et al., 2006). In the number-matching task the participants are presented with two numbers (cue) displayed together followed by a third number (probe). The task of the participants was to decide whether the probe number matched or not with one of the two cue numbers. For example, given the cue numbers 3 and 7, the probe could be 7 (matching trials) or 16 (no-matching trials). Since arithmetic knowledge is not required to accomplish the task, this paradigm allows to implicitly study the strength of the associations between a number and its multiples. In fact, in the no-matching trials the probe can be arithmetically related to the cue numbers (e.g., given the cue numbers 3 and 7, the probe can be the product (21) or a multiple (28)) or not-related (e.g., the probe 23 is neither the product nor a multiple). Adopting this paradigm, Galfano et al. (2009) analysed the brain activity evoked by the presentation of product probes, multiple probes, and no-related probes. The results showed that the brain activity are consistent with the direct activation architecture. Immediately after the presentation of the probe stimulus, the activity evoked by the product probes and by the multiple probes is similar, then the activity evoked by the multiples decay and became similar to the activity evoked by the not-related probes. These results suggest that the presentation of two numbers automatically activates the nodes associate with both the product and other multiples of that numbers.

The direct activation architecture states that a single number (an operand) can directly activate its multiples. However, it does not specify whether an operand activates all its multiples with the same strength or whether some multiples receive more activation than others. What we are interested to test in the present experiment is if tie multiples receive a stronger activation and is results following ties receive a larger activation than multiples preceding ties. In fact, this is a critical assumption of the *asymmetry hypothesis* (see chapter 2). Namely, we aim to test if the tie multiples (the results of the problems 2×2, 3×3, 4×4, and so on, that is the multiples 4, 9, 16, 25, and so on) receive a particular strong activation. This assumption is supported by the evidence that the advantage of the tie problems relative to the non-tie problems is likely due to the easier access of the formers to the result nodes in the multiplication facts network (Campbell & Gunter, 2002).

The *network retrieval model* and the *network interference model* state that the result nodes are interconnected. Therefore, once activated a result nodes the activation should spread at least to the closer results node (see Niedeggen & Rösler, 1999). According to this kind of architecture, we hypothesize that the activation spreading between the result nodes could follow the forward direction (that is the small to large numbers direction). We claimed that the forward direction is privileged because of when the people use the number system they typically adopt the forward direction: when the people have to enumerate a set of object they enumerate in the forward direction (1, 2, 3, 4, ...); when the people try to remember a multiplication table they “count” in the forward direction (e.g., the 7 table is remembered by using the series 7, 14, 21, 28,...); etc. Moreover, the adults report sometime to use the repeated addition procedure (e.g., 6×3=6+6=12+6=18) and the table sequence procedure (e.g., 6×3=6, 12, 18) to solve multiplication problems (LeFevre et al., 1996a; Smith-Chant and LeFevre, 2003; Hecht, 1999). These two strategies should consolidate the association of the multiples to each others, and in particular they should strengthen the association between a multiple and its subsequent one inside the multiplication table. In other words, since when the numbers are used they are typically ordered in the forward direction, we assume that this practice could have effects on the network that encodes the arithmetic facts, and in particular it should produce an asymmetry during the spread of activation between the result nodes.

To summarize, we assume that each number is associated with its multiples (the results of its multiplication table) and that the tie multiples have a particularly strong activation. Moreover, according to the *network retrieval model* and the *network interference model*, we assume that the result nodes are associated each other and then that the activation spreading inside the result nodes. The hypothesis that we would like test in this study is the asymmetric spread of activation across multiples. Namely, we think that inside the result nodes set the activation spreads mainly to the forward direction. To test this hypothesis we decided to evaluate the activation spreading around the tie results because of it is a critical assumption of the *asymmetry hypothesis*. To this end, we had conceived two paradigms: a modified matching task and a multiples task. The matching task we used is similar to the task used in the experiment described above used by Galfano and Collaborators (2009; see also LeFevre et al., 1988; Galfano et al., 2003; Rusconi et al., 2004; Rusconi et al., 2006). In this task the participants were presented with two numbers sequentially presented (a cue followed by a probe) and they had to decide if the cue and the probe matched. For example, after presenting the cue 6, in the matching condition the probe was 6, whereas in the non-matching condition the probe could be 42 (multiple trial) or 34 (neutral trial). The multiple trials were divided into tie-1 and tie+1 conditions. In the tie-1 condition the probe was the multiple before the tie (e.g., the cue was 6 and the probe was 30=36–6), whereas in the tie+1 condition the probe was the multiple after the tie (e.g., the cue was 6 and the probe was 42=36+6). Our prediction is that the multiple after the tie (tie+1) should generate more interference than the multiple before the tie (tie-1). This larger interference should affect the RTs, that is the participants should respond no slower in the tie+1 condition than in the tie-1 condition. The larger interference of the tie+1 multiple would be a direct test of the asymmetric activation, spreading from the tie to the closer multiples. In the matching task the cues in the target trials are typically two one-digit numbers (see Galfano et al., 2009; Galfano et al., 2003; LeFevre et al., 1988; Rusconi et al., 2004; Rusconi et al., 2006). However, due to our specific purpose we presented only one one-digit number as cue. This choice was made to avoid possible confound due to the presentation of two numbers. In fact, the presentation of two one-digit numbers would have activated the product of that numbers (e.g., the presentation of 6 and 7 as cues actives the representation of 42). The activation of other problems could interfere with the activation spreading around the tie that we aim to study. Therefore, we decided to present only one one-digit number as cue.

In the multiples task the participants were simply asked to report if in a sequence of two numbers, sequentially presented, the second one was a multiple of the first one. For example, the participants had to respond “yes” if the sequence presented was “6 42” (multiple trial), “no” if the sequence was “6 45” (non-multiple trial). The multiple trials included tie-1 condition (e.g., “6 30”), tie condition (e.g., “6 36”), and tie+1 condition (e.g., “6 42”). According to the asymmetric activation spreading we expected that the participants responded faster to the tie+1 condition (easier to access) than to the tie-1 condition (more difficult to access). In this task the tie multiples were part of the set of stimuli, then their representations were activated more time during the experiment.

## 3.2 METHOD

**Participants.** Seventeen students of the University of Trento participated in the present experiment as volunteers (6 females; mean age: 29.8; *sd*: 3.11). All participants were native Italian speakers and had normal or corrected-to-normal vision. This experiment was approved by the Ethic committee of the University of Trento. Each participant performed the matching task followed by the multiples task. The data of the matching task for the first three participants will not be analysed due to technical problems. Therefore, we analysed the data of 14 participants for the matching task and 17 participants for the multiple task.

***Matching task***

**Material.** The stimuli used in the matching task are reported in Appendix 1. Each trials consisted in the presentation of a sequences of two numbers, that is a cue followed by a probe. Since we wanted to test the hypothesis of an asymmetry in the activation spreading, we decided to present as cues the number 4, 5, 6, 7, 8, and 9 and as probes the multiples around the tie multiples of the cue numbers (e.g., if the cue was 6 the probe were 30 (tie–1) and 42 (tie+1)). For the cues 4 and 9 we presented only the tie+1 and tie–1 multiples respectively. In the matching trials the cues and the probes were the same number (e.g., cue=6, probe=6); in the non-matching trials the cue and the probe was two different numbers (e.g., cue=6, probe=42). In the non-matching trials there were 6 conditions: tie+1, tie–1, neutral+1, neutral-1, and fillers. In the tie-1 condition the probe was the multiple before the tie in the multiplication table of the cue (e.g., cue=6, probe=6×(6–1)=30); whereas in the tie+1 condition the probe was the multiple after the tie (e.g., cue=6, probe=6×(6+1)=42). In the neutral+1 and neutral-1 conditions the cues were the same numbers used in the tie+1 and tie–1 conditions, but the probes were numbers that are not member of any multiplication tables (e.g., cue=6, probe=34). Since, each probe was presented in both tie–1 and tie+1 conditions with two different cues (e.g., 20 was presented as probe of 4 in the tie+1 condition and as probe of 5 in the tie–1 condition), the probes of the neutral+1 and neutral–1 conditions were presented twice (once in each neutral condition). Moreover, the cues that shared the same probe in the tie condition shared the same probe also in the neutral conditions. For example, the number 20 was presented in both the tie+1 and the tie–1 conditions with the cues 4 and 5, the probe 38 as well as were presented in both the neutral+1 and neutral–1 conditions with the cues 4 and 5. The mean of the probe numbers in the tie+1 and tie–1 conditions was 44.0 and the mean of the probe numbers in the neutral+1 and neutral–1 conditions was 42.4. In the fillers condition both cue and probe were numbers that were not members of any multiplication table.

In the matching trials there were 6 conditions: cue-balancing+1, cue-balancing–1, probe-matching multiple, probe-matching neutral, and fillers. The cue-balancing+1 and the cue-balancing–1 conditions had as cues and probes the same cues used in the tie+1 and tie–1 conditions of the non-matching trials. The probe-matching multiple and the probe-matching neutral conditions presented as cues and probes the same probes used in the tie and neutral conditions of the non-matching trials. The fillers condition presented as cues and probes some of the numbers used in the non-matching fillers condition.

The total number of cue-probe pairs in the matching task was 50: 5 (stimuli per condition) × 5 (conditions) × 2 (matching or non-matching sequences). Each trial stimulus was repeated 12 times with a total of 600 trials in the whole experiment. The participants were presented with the same number of matching and non-matching trials.

**Procedure.** The stimuli were presented in white on a black background. The procedure we used was similar to the procedure of Galfano et al. (2009), with the exception that we presented one cue instead of two. The 600 trials (50 stimuli repeated 12 times) were divided in 10 blocks of 60 trials each. Between the blocks the participants could take a short break. The order in which the stimuli were presented was randomized for each participants. Each trials started with a fixation point ("#") shown for 400 ms at the centre of the screen. After this time the cue replaced the fixation point. The cue was presented for 60 ms and was immediately followed by a mask frame consisting of the "####" string presented for 40 ms. After the mask frame a black screen was presented for 20 ms and after the black screen the probe was presented until the participants responded. The interval between the onset of the cue and the onset of the probe (stimulus onset asynchrony) was 120 ms. Between the trials there were intervals of 1100 ms. The participants had to respond by pressing the Z keyboard key  with the left hand for yes answers and the M key for no answers. Contrarily to what we planned response hand was not counterbalances across participants, due to an error in the program used for the presentation of the stimuli.

The matching task was preceded by 12 practice trials (6 matching and 6 non-matching), in which both cues and probe were numbers that were not member of any multiplication table. The matching task required about 30 minutes. After the matching task the participants could take a short break. When the participants were ready they could start the multiple task.

### *Multiples task*

**Material.** The stimuli used in the matching task are reported in Appendix 2. Each trials consisted in the presentation of a sequences of two numbers, that is a one-digit number (cue) followed by a two-digit number (probe). Cues were the numbers 4, 5, 6, 7, and 8 and the probes were the multiples around the ties and tie multiples of the cues (e.g., if the target cue was 6 the target probe were 30 (tie–1), 36(tie), and 42 (tie+1). In the table trials the probe was a multiple of the cue (e.g., cue=6, probe=36); in the non-table trials the probe was a number that was not in the table of the cue (e.g., cue=6, probe=39). In the table trials there were 3 conditions: tie–1, tie, and tie+1. In the tie–1 condition the probe was the multiple before the tie, in the tie condition the probe was the tie, and in the tie+1 condition the probe was the multiple after the tie (e.g., given the cue 7, the tie–1 probe was 42, the tie probe was 49, and the tie+1 probe was 56).

In the non-table trials there were 4 conditions: below tie–1, below tie, above tie, and above tie+1. The probes in the below tie condition were the number in the middle between the tie and the tie–1. The probes in the above tie condition were the number in the middle between the tie and the tie+1. In both below tie and above tie conditions the probes were rounded off towards the tie. The distance between the tie probe and the below tie probe was subtracted to the tie–1 multiple or added to the tie+1 multiple to generate the below tie–1 and above tie+1conditions respectively.

Each cue-probe pair in the table set was repeated 12 times, whereas the stimuli in the non-table was repeated 9 times each. The total number of table trials was 180: 5 operands (4, 5, 6, 7, and 8) × 3 condition (tie–1, tie, and tie+1) × 12 repetitions. The total number of non-table trials was 180: 5 operands (4, 5, 6, 7, and 8) × 4 condition (below tie–1, below tie, above tie, and above tie+1) × 9 repetitions. The participants were presented with the same number of table and non-table trials (360 trials in the experiment).

**Procedure.** The stimuli were presented in white on a black background. The 360 trials were divided in 8 blocks of 45 trials each. Between the blocks the participants could take a short break. The order in which the stimuli were presented was randomized for each participants. Each trials started with the presentation of the cue for 600 ms followed by a black screen presented for 200 ms. After the black screen the probe was presented until the participants responded. The interval between the onset of the cue and the onset of the probe (stimulus onset asynchrony) was 800 ms. Between the trials there were intervals of 1000 ms. The participants had to respond by pressing the keys “M” and “Z” on the keyboard. All the participants were required to press “Z” with the index finger of the left hand if the cue and the probe matched, and “M” with the right index of the right hand if the cue and probe did not match. Like in the matching task and for the same reasons, we did not balance the key to press across the participants.

The matching task was preceded by 10 practice trials (5 table trials and 5 non-table trials), in which each cue was repeated twice (1 table and 1 non-table trials) and the probes were either multiple numbers or non-multiple not used in the experiment. The matching task required about 15 minutes.

## 3.3 RESULTS

### *Matching task*

The accuracies for each experimental cell are reported in table 3.1. The table shows that the participants made more errors in the matching conditions than in the non-matching conditions. Contrarily to what we expected, these results suggest that the participants found easier to respond when the cues and the probes were not the same number (non-matching condition).

| **matching** | | **no matching** | |
|---|---|---|---|
| **condition** | **acc.** | **acc.** | **condition** |
| cue-balancing+1 | 0.82 | 0.97 | tie+1 |
| cue-balancing-1 | 0.81 | 0.96 | tie-1 |
| probe-balancing M | 0.81 | 0.94 | neutral+1 |
| probe-balancing N | 0.79 | 0.95 | neutral-1 |
| filler | 0.81 | 0.95 | filler |

**Table 3.1: the accuracies in each experimental cell in matching and non-matching conditions.**

The figure 3.1 shows the RTs aggregated across participants and congruence (that is matching trials vs non-matching trials). The participants responded faster in the non-matching condition (578 ms) than in the matching condition (623 ms), $t(13)=2.04$, $p=0.06$. Both the smaller number of errors and the faster RT in the non-matching condition suggest that the participants found easier to respond in the non-matching conditions than in the matching condition.

The following analysis have been performed on the RTs for the correct responses. For each participant outliers were removed using the Van Selst and Jolicoeur (1994) procedure. We recursively removed the RTs that beyond 3.5 standard deviation from the mean of each participant. Then, we calculated the mean RTs for each participant for both tie-1 and tie+1 conditions. The analysis showed that there was no differences in the RTs between the two conditions (541 ms for both conditions).

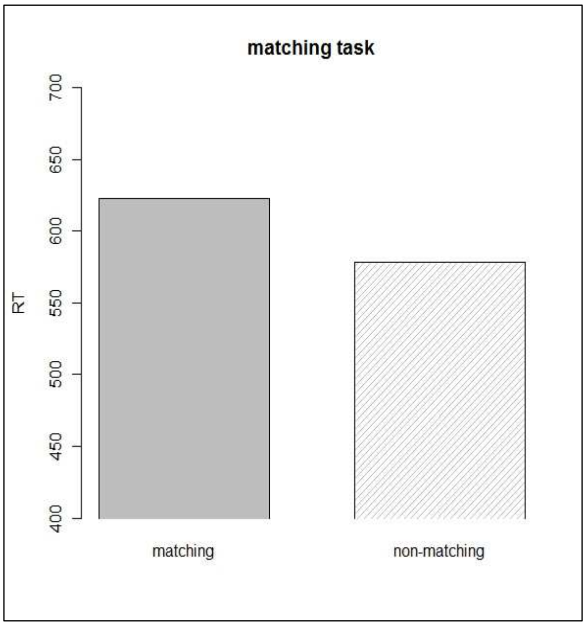


**Figure 3.1: mean RTs in milliseconds as function of the congruence (matching vs non-matching trials).**

A further analysis was conducted on the cues that were presented in both conditions (the cues 5, 6, 7, and 8 were used in both tie-1 and tie+1 conditions; see Appendix 1). The cues 5, 6, 7, and 8 were presented followed by probe numbers that were either tie-1 multiple or tie+1 multiple of the cues. For example, the cue 6 was presented once in the tie-1 condition followed by the probe 30 (6×6=36–6=30) and once in the tie+1 condition followed by 42 (6×6=36+6=42). A repeated measure ANOVA was conducted on the RTs with condition (tie-1 and tie+1) and cue (5, 6, 7, and 8) as within factors. Again, no significant effects emerged. In an exploratory view, it is interesting to have a look to the means reported in table 3.2 and figure 3.2.

| | cues | | | |
|---|---|---|---|---|
| conditions | 5 | 6 | 7 | 8 |
| tie-1 | 545 | 541 | 524 | 550 |
| tie+1 | 530 | 540 | 549 | 523 |

**Table 3.2: mean RTs in milliseconds for the cues 5, 6, 7, and 8 in the tie-1 and tie+1 conditions.**

The mean RTs in each experimental cell show that, even though not significant, there are some asymmetries comparing the multiples. In the trials in which the cues were 5 and 8 the probes in the tie+1 (30 an 72, respectively) category showed an advantage with respect to the probe in the tie-1 category (20 an 56, respectively); whereas the cue 7 showed the inverse pattern, the tie-1 probe (42) showed an advantage with respect to the tie+1 probe (56).

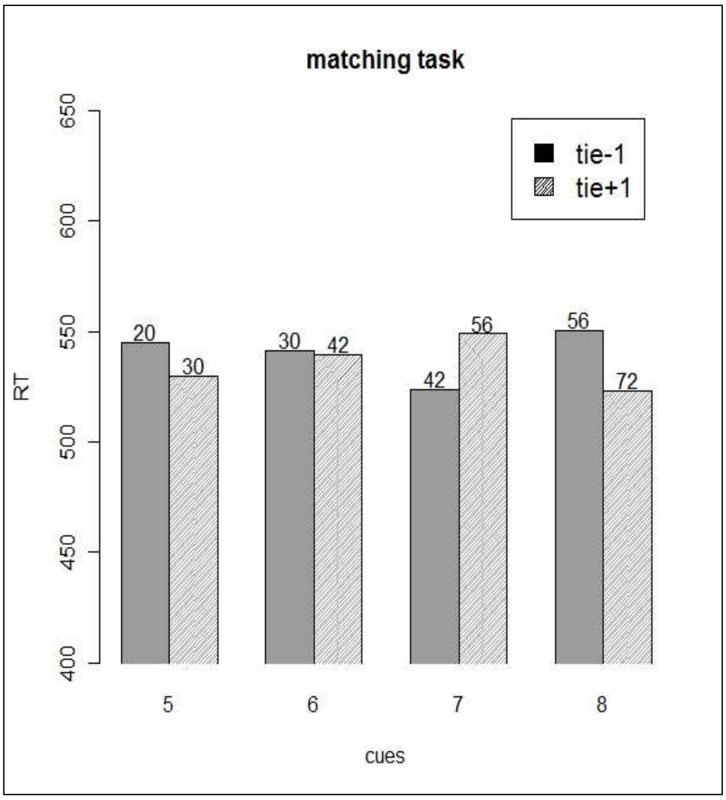


**Figure 3.2: mean RTs in milliseconds as function of the cues (5, 6, 7, and 8) and conditions (ties-1 and ties+1)**.

### *Multiples task*

The accuracies for each experimental condition are reported in table 3.3. Unlike than for the matching task, the number of errors were similar in the conditions. Furthermore, the participants made less errors in the table conditions (that is the conditions in which the probe was a multiple of the cue), and tended to respond faster to the table trials (the probe is a multiple of the cue, 763 ms) than to the non-table trials (the probe is not a multiple of the cue, 817 ms), t(16)=–2, p<0.1 (figure 3.3).

| condition | accuracy |
|---|---|
| below tie-1 | 0.87 |
| below tie | 0.9 |
| above tie | 0.85 |
| above tie+1 | 0.86 |
| tie | 0.94 |
| tie-1 | 0.94 |
| tie+1 | 0.92 |

**Table 3.3: the accuracies in each experimental cell in the multiple task.**

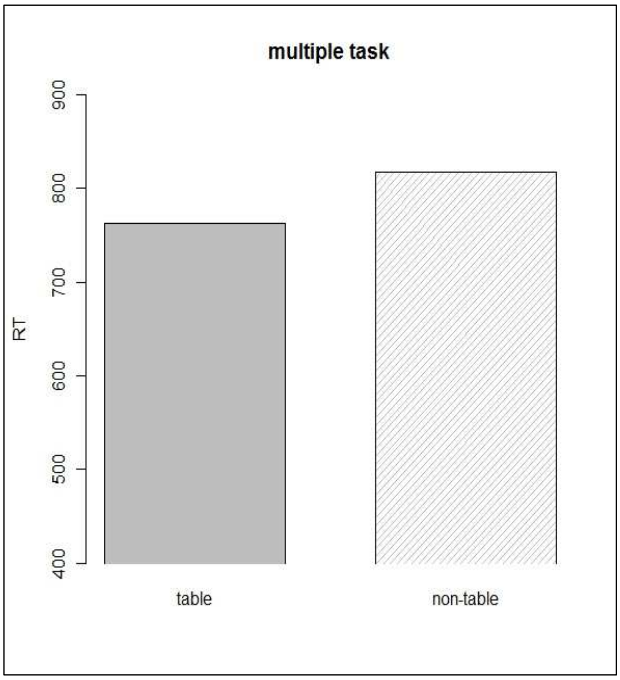


**Figure 3.3: mean RTs in milliseconds as function of the congruence; in the table trials the probe was a multiple of the cue, in the non-table trials the probe was not a multiple of the cue.**

The following analysis have been performed on the RTs for the correct responses. The outliers have been removed with the same method used for the matching task ( Van Selst and Jolicoeur, 1994 with a threshold 3.5 standard deviations). Since we were interested to the differences between multiples, we carried out the analysis only on the table trials (that is, the trials where the probe was multiple of the cue). We performed an repeated measure ANOVA on the RTs for the correct trials with condition (tie-1, tie, and tie+1) and cue (4, 5, 6, 7, and 8) as within factors (see figure 3.4).

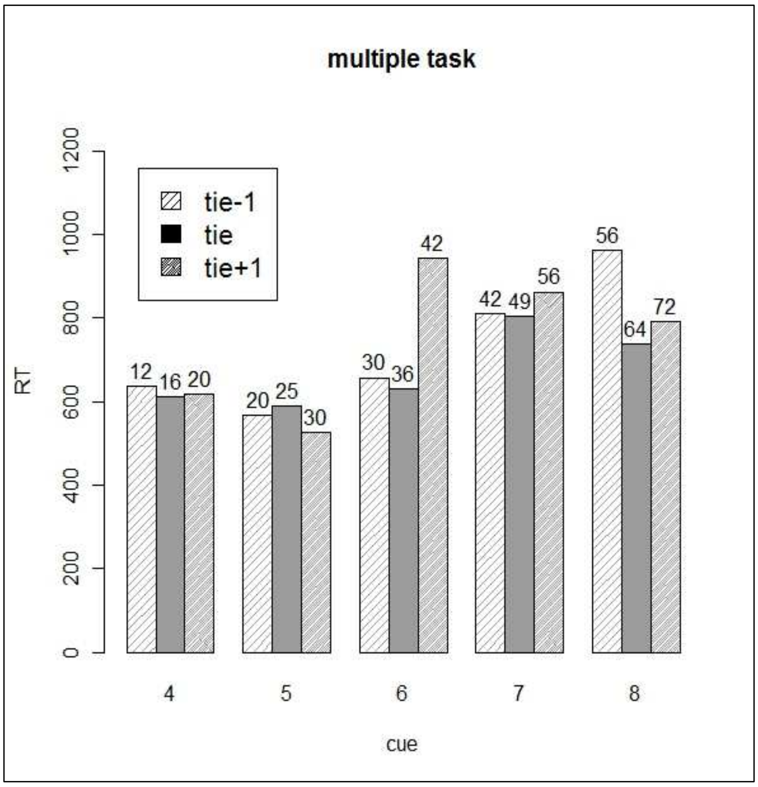


**Figure 3.4: mean RTs in milliseconds as function of cue and condition.**

The main effect of condition was significant, F(2,32)=3.65, p<0.05. Post-hoc analysis showed that the participants tended to respond faster in the tie condition (672 ms) than in tie+1 condition (752 ms), p<0.1. Moreover, contrarily what we expected, the participants tended to be faster in tie-1 condition (721 ms) than in tie+1 condition (752 ms), p<0.1. The difference between tie and tie-1 conditions was however not significant. The main effect of cue was significant, F(4,64)=8.07, p<0.001. The effect of the cue likely reflected the problem size effect (Zbrodoff & Logan, 2005). In fact, the RTs were modulated by the size of the cues (621, 562, 741, 854, and 839 for the cue 4, 5, 6, 7, and 8 respectively), that is the recognition of the multiples was more difficult for the larger cues. The condition by cue interaction was also significant, $F(8,128)=3.79$, $p<0.001$. The post-hoc analysis showed a tendency to respond faster in the tie+1 (529 ms) condition than in the tie-1 condition (568 ms) when the cue was 5, $p<0.1$; whereas, contrarily what we expected, the tie-1 condition (655 ms) tended to be response faster than the tie+1 condition (979 ms) when the cue was 6, $p<0.1$. No others comparisons were significant.

### 3.4 DISCUSSION

This study was based on the assumption that there is a strong association between a number (e.g., 6) and its tie multiple (e.g., 36). The aim of this study was to investigate the possibility that the activation spreading around the tie result nodes in the arithmetic facts memory would have an asymmetry. Namely, we hypothesized that the activation spreads more in the forward direction (i.e., from tie to larger multiples) than in the backward direction (i.e., from tie to smaller Multiples).

The data collection was interrupted after 17 participants when, after a first analysis of the data, we discovered that in the matching task the “no” responses for non-matching trials were faster and more accurate than the “yes” responses for matching trials. Going back to check the paradigm we discovered a possible cause of it, the lack of counterbalancing of the response hand across participants. The response criteria in the matching task, as well as in similar paradigms (e.g. lexical decision), is typically biased trough a “yes” response and a time limit criteria is typically used by participants to balance accuracy and speed (e.g. if after a given time there is no positive evidence for a match do respond “no”). In fact previous studies that implemented a paradigm similar to the one here used (Galfano et al., 2009; Galfano et al., 2003; LeFevre et al., 1988; Rusconi et al., 2004; Rusconi et al., 2006) typically found more accurate and fast answer in the match condition. One possible explanation of our results is that the “yes” response was associated with the left, non dominant hand, however accuracy in the second task (where the same behavioural confound was present) was better for “yes” answers and thus this can be not the whole story. We cannot disregard another possible explanation of the faster and more accurate responses to the non matching trials in the first task. One other difference from the original paradigm (Galfano et al., 2009; Galfano et al., 2003; LeFevre et al., 1988; Rusconi et al., 2004; Rusconi et al., 2006) was the use of a single cue instead of two. The procedure to replace the probe in the same position of the cue could have lead participants to adopt a specific strategy: given that cues and targets were presented in fast sequence at the same point of the screen, the use of simple perceptive features (or even just the number of digits) to check whether the two numbers were the same or not could have been used. This can have affected our data and explain a null result, since the activation of abstract numeric representations needed to accomplish the task could have been very weak.

Despite this, the qualitative analysis of the first task (see figure 3.2) showed numerical asymmetries that do not show a global tendency as expected by our hypothesis (a right skewed activation around ties) but different tendencies to an asymmetry that depend on specific cues.

In the second task, where the confound of the use of mere perceptual strategies in order to solve the task does not apply, the condition by cue interaction provided little evidence for an asymmetry, but the direction of this asymmetry again depended on the specific of cue. For example, the direction of the asymmetry is inverted between the cue 6 and 8.

Despite the null results obtained with this two paradigms, the qualitative results suggest that the presentation of a single number could not lead to the activation spreading we was looking for. In the literature in the matching task are usually used two cue numbers (see for example Galfano et al., 2003). Given our interest in the spreading of activation from the ties the presentation of the same number repeated twice as cues could have been used. However, the multiples above or below the tie presented as probe (e.g., 6 6 as cue and 42 as probe; 6 6 as cue and 30 as probe) could have made the task too easy.

The first task tested whether the implicitly activation of multiples within a task that does not require multiplication knowledge give rise to the asymmetry of spreading of activation around ties that could explain our results of chapter 1. The second task requires the activation of nodes in the associative memory for arithmetic facts. In both cases numerical asymmetries were found but without any overall direction. Therefore we think that it could be possible that in order to obtain an asymmetry inside the arithmetic facts memory may be necessary to present both operands of a problem in an arithmetic context, namely a context for which the problems nodes are also activated within the arithmetic facts memory. Therefore, we decided to test the *asymmetry hypothesis* in a result verification task that will be described in the next chapter.

# Chapter 4

## The operands-order effect and the asymmetric spreading activation inside the multiplication facts memory: a study with verification task

## 4.1 INTRODUCTION

In this Chapter we report a verification task experiment, in which the whole components of the arithmetic facts memory, including problem nodes and not only multiples of single digits (operands) are directly involved. In the verification task the participants are presented with an equation (two operands and a result) that can be correct (e.g., 2×9=18) or incorrect (e.g., 2×9=23). The participants have to judge if the presented result is correct or not.

The experiment here presented has two aims: the first is to evaluate if the operands-order by size interaction we found for Italians participants in the production task (see chapter 2) affects also the performance when the participants have to judge if a presented equation is correct (e.g., 6×4=24) or incorrect (e.g., 6×4=18). First of all, given that the results of the first experiment reported in this thesis were not replicated with English speakers it is interesting to replicate the results with a different pool of participants. Instead of running a second production experiment we however decided to test if the same interaction emerges also in a verification task. The two tasks differ for a number of reasons, first of all a difference between production and verification is that while in the former an explicit selection of the correct result within a pool of active result nodes is necessary to achieve the task (Zbrodoff & Logan, 1990; Zbrodoff & Logan, 2000). In a verification task this selection may come from an interplay with the representation of the presented result that largely facilitates to achieve a rather good accuracy with a lesser effort. For this reason in a verification task is more likely than in a speeded production task that the participants rely mainly on retrieval procedures. Thus, finding the same results could help us in attributing the interaction between order and size as a property of the internal organization of arithmetic facts memory, despite being shaped by learning and use of non-retrieval procedures.

The second aim of using a verification task is to evaluate, by carefully manipulating the trials where the result is incorrect (e.g., 6×4=18), whether the activation spreads asymmetrically between the multiples that constitute the multiplication table. The second aim is interesting for three reasons. First, it can be used to test the *asymmetry hypothesis* (see chapter 2), that is to evaluate if in the multiplication facts memory the activation of a tie result spreads more to the larger multiple (forward direction) than to the smaller multiple (backward direction). Second, it can be evaluated if the asymmetric activation spreading is a general feature of the associations between result nodes in the associative memory that is assumed to store multiplication facts. Third, the asymmetric activation spreading of the tie results, if exist, can be compared with the asymmetric activation spreading of non-tie results.

In the production task experiment (chapter 2), we found a size by order interaction. Namely, for small and medium problems the L×s order (e.g., 4×3, 7×3) was solved faster than the s×L order (e.g., 3×4, 3×7); whereas, for large problems the s×L order (e.g., 7×8) was solved faster than the L×s order (e.g., 8×7). We proposed two possible explanations: *reorganization hypothesis* and *asymmetry hypothesis*. The *reorganization hypothesis* states that the interaction is due to both the order in which the multiplication problems are learned and the reorganization of the memory due to the use of non-retrieval procedures in the childhood. The *asymmetry hypothesis* explains the interaction with a different amount of activation of the result when the second operand is presented. The order by size interaction has been found in a production task, in which the participants had to identify the result of the problem by accessing to the arithmetic facts. In other words, the production task requires the full identification of the result of the arithmetic problem in the multiplication facts memory (Zbrodoff & Logan, 1990; 2000). On the contrary, in the verification task the participants use the all the element of the equation (operands and result) to judge whether the equation is correct (Zbrodoff & Logan, 1990; 2000). According to Zbrodoff and Logan (1990; 2000), verification and production are two different processes that rely on the same "arithmetic knowledge base". If we consider the multiplication facts memory as the "arithmetic knowledge base" on which the verification and the production tasks operate, then the organization of the multiplication facts memory should affect in a similar way the performance in both tasks with respect to the operands-order by size interaction. In fact, both the *reorganization hypothesis* and the *asymmetry hypothesis* state that the order by size interaction is produced by specific features of the “arithmetic knowledge base”. According to the *reorganization hypothesis* only one of the two orders of the operands is stored in memory, and then both product the result and verify the result (or the whole equation) should be easier/faster for the stored order. According to the *asymmetry hypothesis* the activation spreading advantage one of the two orders, and then, as well as for the reorganization hypothesis, the advantage should be the same across production and verification. Therefore, we expect to find in the verification a similar order by size interaction we found in the production task discussed above.

The second aim of this experiment is to evaluate whether the activation of the result (e.g., 21) generated by the operands of a problem (e.g., 7×3) spreads in an asymmetric way around the result. Various studies (Galfano et al., 2003; Galfano et al., 2009; Rusconi et al., 2004; Rusconi et al., 2006) showed that the simple presentation of two one-digit numbers in a matching task (in which the arithmetic knowledge are not required) is able to activate the at least the two multiples around the result of the multiplication problem with that numbers as operands (e.g., 4 and 6 activate 20 (below the product 24) and 28 (above 24)). For example, the presentation of the two numbers 4 and 6 is able to activate the multiple 20, 28, 18, and 30, which are around the product 24. One explanation of this result could be that the multiple are associated one other (see the *network retrieval model* and the *network interference model* in chapter 1). Namely, once presented the two numbers, their product is automatically activated and then this activation spread at least to the closer multiples (but see Galfano et al., 2009, for a different architecture with a direct activation not mediated by the true result). In the verification task, when the presented result is incorrect, the two operands could activate the correct result and this result could spread the activation to the multiples that are close to the results. However, we hypothesized that this activation spreading is not symmetric. More precisely, we hypothesized that the activation spreads more in the forward direction (from the result to the larger multiples) than in the backward direction (from the result to the smaller multiples). This assumption is based on the observation that when people try to remember a table they typically “count” from the smaller multiple to the larger (e.g., remembering the 7-table “counting” 7, 14, 21, 28, and so on). Therefore, the multiples could be more associated with the subsequent multiple than to the previous one because for producing the multiple list correctly each multiple has to be associated with the next one and not with the previous one. Moreover, the non-retrieval procedures like repeated addition (e.g., 6×3=6+6=12+6=18) and table sequence (e.g., 6×3=6, 12, 18) reinforce only the association between a multiple and the next one, and not vice versa. Therefore, we expected that, when presented a incorrect equation, a incorrect result above the correct result (e.g., 7×3=21) would produce more interference than a incorrect result below (e.g., 7×3=14). In fact, the above multiple (that is in the forward direction) should be more activated than the below multiple (backward direction) due to the asymmetric activation spreading. To test this hypothesis, in the incorrect equations we used as incorrect results only the multiple above and below the correct results. This should also make the task more difficult (all the incorrect results are the correct result for another multiplication) and then it is more unlikely that the participants would base their responses on a generic familiarity judgement.

## 4.2 METHOD

**Participants.** Twenty-two students of the University of Trento participated in the present experiment as volunteers (11 females; mean age: 21.8; *sd*: 6.3). All participants were native Italian speaker and had normal or corrected-to-normal vision. This experiment was approved by the Ethic committee at the University of Trento.

**Material.** We use as stimuli the multiplication problems from 3×3 to 8×8 (overall 36 problems). For each problem we presented 4 correct equation (e.g., 7×3=21) and 4 incorrect equation (e.g., 7×3=28) in order to balance the number of yes and no responses. The result of each incorrect problem was one of the multiples around the correct solution of the problem and this was done for each of the two operands. Namely, given a problem (e.g., 7×3) the 4 incorrect solution were: 1) the multiple of the first operand above the result (e.g., 7×3=28); 2) the multiple of the first operand below the result (7×3=14); 3) the multiple of the second operand above the result (e.g., 7×3=24); 4) the multiple of the second operand below the result (7×3=18). Therefore, each problem was presented 8 times (4 correct and 4 incorrect). The participants performed 4 blocks with 72 problems each (36 correct and 36 incorrect) and in each block there were a correct result and a incorrect one for each problems. In each block the participants were presented with all the 4 incorrect multiple conditions and the order in which the 4 incorrect solution of a problem were presented in the 4 blocks varied randomly for each participant. Therefore, given the problem (e.g., 7×3), the order in which the incorrect solutions of the problem (14, 18, 24, 28) were presented varied randomly across the participants. Totally, the participants performed 288 problems, 144 was presented with the correct solution and 144 with the incorrect one. In each block was presented 72 problems, 36 correct and 36 incorrect. All the incorrect results were a multiple of one of the two operands and were close (above or below) to the correct result.

**Procedure.** During the experiment the participants sat alone in a partially sound-proof room. In order to familiarize with the experimental procedure the participants performed a block of practice in which they were presented 10 trials which problems had 2 or 9 as operand. The 2 and 9 tables were not used as experimental trials given that multiples below and above the correct results were trivial (1 and 10 table). After the practice the experimenter made sure that the procedure was clear to the participants. The problems were sequentially presented at the centre of a monitor of a PC. Each trial started with a fixation point ("#") presented for 1 second. The first operand, the sign ("×"), and the second operand were presented for 300 ms each. After the second operand the proposed result was presented until the participants responded. Once one stimulus disappear the next one was immediately presented on the screen without any interstimulus interval. The operands, the sign, and the results had a dimension of about 1 cm and the participants were at about 60 cm from the monitor. Participants were required to judge whether the presented result was correct or not and to respond with the keys “Z” and “M” of the keyboard with the left and right hand respectively, as quick and accurate as possible. One half of the participants had to respond “Z” to the correct result and “M” to the incorrect ones, whereas the other half had to respond with the inverse keys response code. The participants had to perform 4 experimental blocks and between one block and the other they could take a little break.

## 4.3 RESULTS

One participant was excluded from the analysis due to extremely slow reaction times (RTs). The mean RTs of that participant (2162 ms) was beyond 2 standard deviation from the mean of the group (1138 ms). Moreover, in the following analysis, for the reaction times we removed the outlier values for each participant by using the outlier elimination procedure of Van Selst and Jolicoeur (1994). We recursively removed the RTs that beyond 3.5 standard deviation from the mean of each participant. Furthermore, ANOVAs were Greenhouse-Geisser corrected when the degrees of freedom of the numerator exceeded one (uncorrected degrees of freedom and epsilon values are reported).

First, we analysed the performance of the participants between the correct solution and the incorrect solution trials (see figure 4.1). The accuracy (the proportion of correct response) and the mean reaction times were calculated for each participant in each of the two correctness condition (correct trials vs incorrect trials). The participants made more errors in the incorrect condition (8% of error; 248 error on 3024 total trials) than in the correct condition (5%; 146), $t(20)=3.91$, $p<0.01$. Moreover, the participants were faster to judge the correct trials (1021 ms) than the incorrect ones (1194 ms), $t(20)=-5.64$, $p<0.001$. Therefore, the correct trials were easier to judge than the incorrect ones. In the following analysis the correct and incorrect trials have been analysed separately to test the different hypotheses outlined in the introduction.

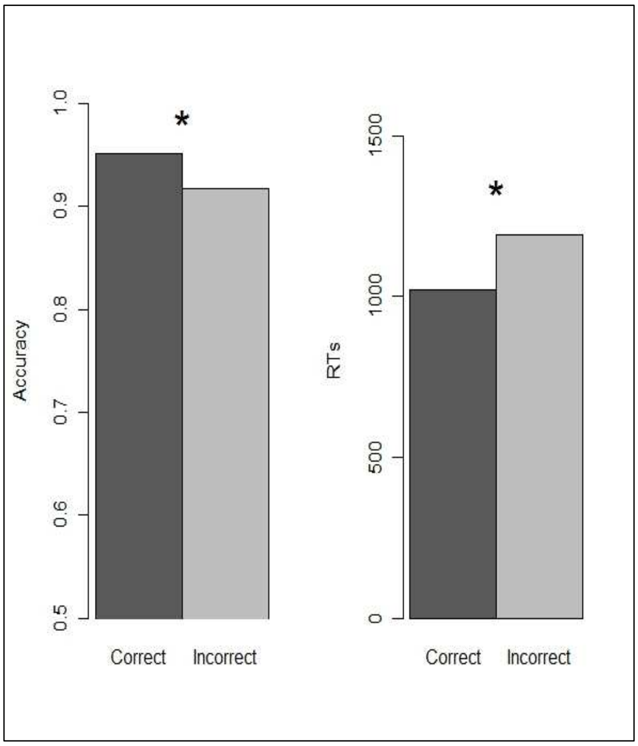


**Figure 4.1: the accuracy (on the left) and the mean RTs (in millisecond, on the right) for the correct and incorrect presented result conditions.** *p<0.05**.**

The trials in the correct condition were analysed to evaluate if the order by size interaction we found in the production task (see Chapter 2 of this thesis) affects a verification task as well. For the correct condition, the accuracy and the mean RTs were calculated for each participants in each of the six experimental cells given by the two factors size and order. The problems were classified in three size category: small, medium, and large. The problems with both operands larger than five were classified as “large” (e.g., 7×8); the problems with one operand larger and one smaller than 5 were classified as “medium” (e.g., 7×3); the problems with both operands smaller than 5 were classified as “small” (e.g., 3×4). The problems 6×5, 7×5, 8×5, and their commuted were classified as "medium"; whereas the problems 3×5, 4×5, and their commuted were classified as "small". The tie problems (e.g., 3×3) were excluded from the analysis. The order factor involved two levels: L×s problems (larger operand in first position, e.g. 7×3) and s×L problems (smaller operand is first position, e.g. 3×7).

On the trials in the correct condition, a 2x3 repeated measure ANOVA was performed on the accuracies with size and order as within factors. Only the main effect of size was significant, $F(2,40)=10.11$, $\varepsilon_{GG}=0.69$ , $p<0.01$. Post-hoc t-test[7] analysis revealed that the participants made more errors in the large condition (89% of correct answers) than in both medium (95%) and small conditions (97%), $t(20)=-3.00$, $p<0.05$ and $t(20)=-3.58$, $p<0.01$, respectively.

On the trials in the correct condition, a 2x3 repeated measure ANOVA was performed on the RTs with size and order as within factors. The main effect of size was significant, $F(2,40)=33.13$, $\varepsilon_{GG}=0.61$, $p<0.001$. Post-hoc t-test analysis revealed that the participants were slower in the large condition (1274 ms) than in both medium (1029 ms) and small conditions (746 ms), $t(20)=4.51$, $p<0.001$ and $t(20)= 6.08$, $p<0.001$, respectively. Moreover, the participants were slower in medium condition than in small condition, $t(20)=6.11$, $p<0.001$. The ANOVA revealed also a significant order by size interaction (figure 4.2), $F(2,40)=6$, $\varepsilon_{GG}=0.99$, $p<0.01$. Post-hoc t-test analysis showed that in the small condition the participants responded faster in the L×s order (703 ms) than in the s×L order (788 ms), $t(20)=-2.95$, $p<0.01$; likewise, in the medium condition the participants responded faster in the L×s order (991 ms) than in the s×L order (1066 ms), $t(20)=-3.14$, $p<0.01$; whereas, in the large condition the participants showed a tendency to respond faster in the s×L order (1242 ms) than in the L×s order (1305 ms), $t(20)=1.56$, $p<0.1$. These results confirm the results of the production task discussed above (chapter 2) and provide further evidences that the order of the operands can affect the performance in multiplication.

[7] All the p-value of the post-hoc t-test reported in this experiment have been corrected with the FDR method.

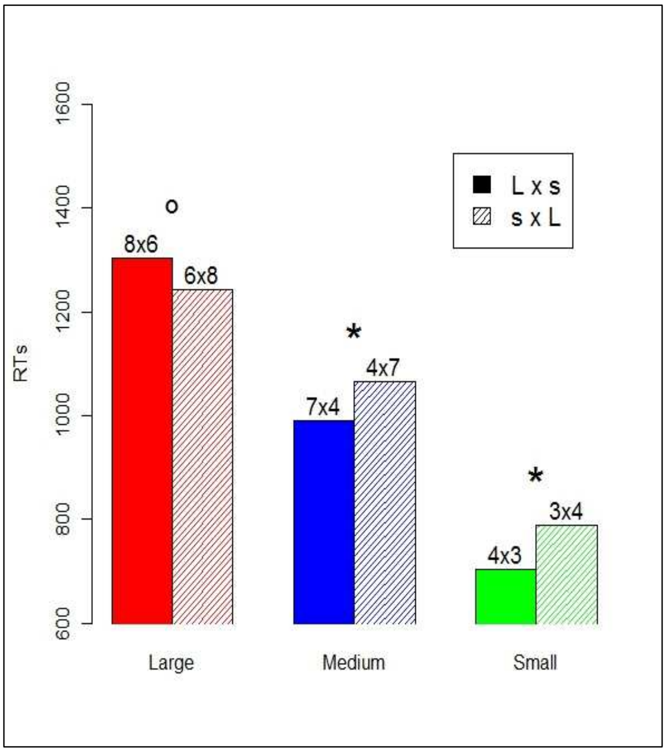


**Figure 4.2: the mean RTs for the size by order interaction. Above each bar is presented an example of a multiplication of that size and order. °** p<0.1**;** *p<0.05**.**

The trials in the incorrect condition were analysed to evaluate if the activation inside the multiplication facts memory spread asymmetrically, that is the activation spread more in the forward direction (i.e., to the activated result to the larger multiples) than in the backward direction (i.e., to the activated result to the smaller multiples). The asymmetry has been compared between tie and non-tie problems. For the incorrect condition, the accuracy and the mean RTs were calculated in the eight experimental cells given by the three factors direction (below: the presented result was below the correct result, e.g. 7×3=18; above: the presented result was above the correct result, e.g. 7×3=24), position (first: the presented result was a multiple of the first operand, e.g. 7×3=28, second: the presented result was a multiple of the second operand, e.g. 7×3=24), and type (tie, e.g. 4×4; non-tie, e.g. 4×7).

On the trials in the incorrect condition, a 2x2x2 repeated measure ANOVA was performed on the accuracies with direction, position, and type as within factors. The ANOVA revealed an direction by type interaction, F(1,20)=8.75, p<0.01. In the non-tie condition, the participants made more errors when the incorrect result where above (89%) than when it was below (94%), t(20)=–3.38, p<0.01. No differences emerged in the tie condition. Contrary to what we expected, the direction factor affected more the performance in the non-tie condition than in the tie condition.

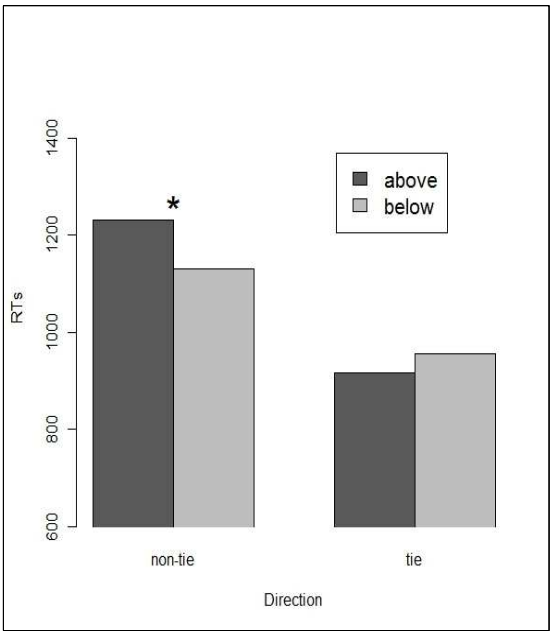


**Figure 4.3: mean RTs (in millisecond) for the direction by type interaction.** *$p$<0.05**.**

On the trials in the incorrect condition, a 2x2x2 repeated measure ANOVA was performed on the RTs with direction, position, and type as within factors. The ANOVA revealed a main effect of type, F(1,20)=25.36, p<0.001, given that wrong result of tie problems were rejected faster (935 ms) than that of the non-tie problems (1180 ms),. More interesting given the aims of this experiment, the ANOVA showed a direction by type interaction (figure 4.3), F(1,20)=13.85, p<0.01. Post-hoc analysis reveal that in the non-tie condition the participants responded slower when the incorrect result was above (1230 ms) then when it was below (1130 ms), t(20)=2.77, p<0.01. On the other side no differences emerged between above incorrect results (915 ms) and below incorrect result (955 ms) for the tie problems. This result is not consistent with the critical assumption of the *asymmetry hypothesis*, that is the activation around the ties spread more in the forward than in the backward direction. However, this result shows that the direction factor affects the performance on non-tie problems. Therefore, we made a further analysis on the non-tie problems to better describe this effect. Since the direction factor does not affect the performance on the tie problems, in the following analysis the tie problems are excluded.

For the incorrect condition (only non-tie problems), the accuracy and the mean RTs were calculated in each of the eight experimental cells given by the three factors direction (below: the presented result was below the correct result, e.g. 7×3=18; above: the presented result was above the correct result, e.g. 7×3=24), distance (small: the presented result was a multiple of the smaller operand, e.g. 7×3=18; large: the presented result was a multiple of the larger operand, e.g. 7×3=14), and position (first: the presented result was a multiple of the first operand, e.g. 7×3=28, second: the presented result was a multiple of the second operand, e.g. 7×3=24).

On the trials in the incorrect condition, a 2x2x2 repeated measure ANOVA was performed on the accuracies with direction, distance, and position as within factors. The main effect of direction was significant, $F(1,20)=11.42$, $p<0.01$ (figure 4.4 on the left). The participants made more errors when the incorrect result was above (89% of correct answers) the correct result than when it was below (94%). This result suggest that the participants found more difficult to judge the equations when the incorrect result was a multiple above the correct result of the problem. Moreover, the main effect of position was also significant, $F(1,20)=4.89$, $p<0.05$ (figure 4.4 on the right). The participants made more errors when the incorrect result was a multiple of the second operand (93% of correct answers) than when it was a multiple of the first (90%).

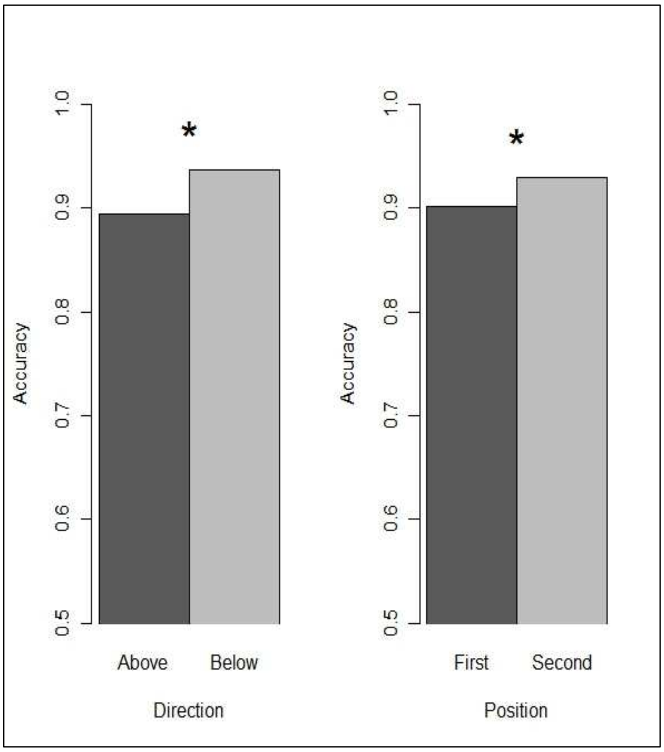


**Figure 4.4: the accuracy in the direction factor (on the left) and position factor (on the right).** *$p$<0.05.

On the trials in the incorrect condition, a 2x2x2 repeated measure ANOVA was performed on the RTs with direction, distance, and position as within factors. Like for the accuracy analysis, the main effect of direction was significant, $F(1,20)=8.04$, $p<0.05$ (figure 4.5). The participants responded slower when the presented result was a multiple above (1230 ms) the correct result than when it was below (1130 ms). Consistently with the accuracy analysis, this result show that the participants found more difficult to judge the equation when the presented incurred result was above the correct result. This suggests that, in the non-tie problems, the multiples above the correct results made more interference than the multiples below, and that the former could be more activated than the latter providing a evidence that the activation spread from the correct result of the problem mainly to the forward direction.

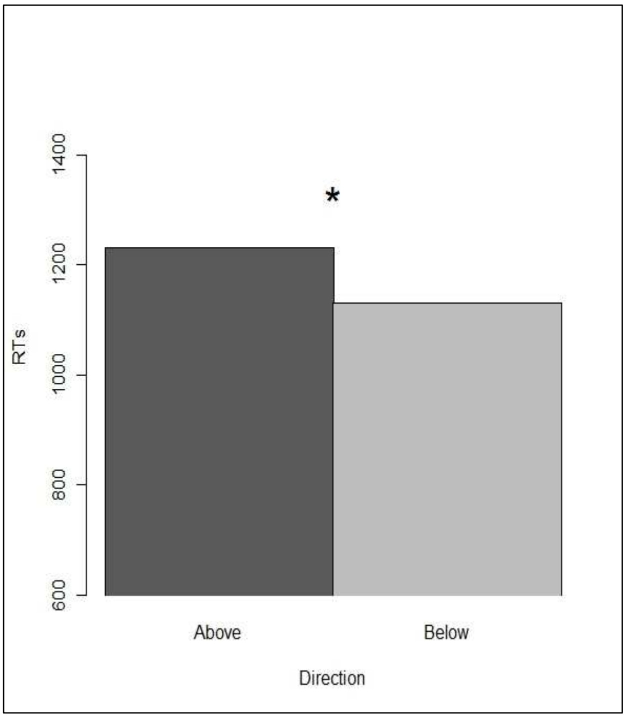


**Figure 4.5: the mean RTs for the two levels of the direction factor.** *$p<0.05$**.**

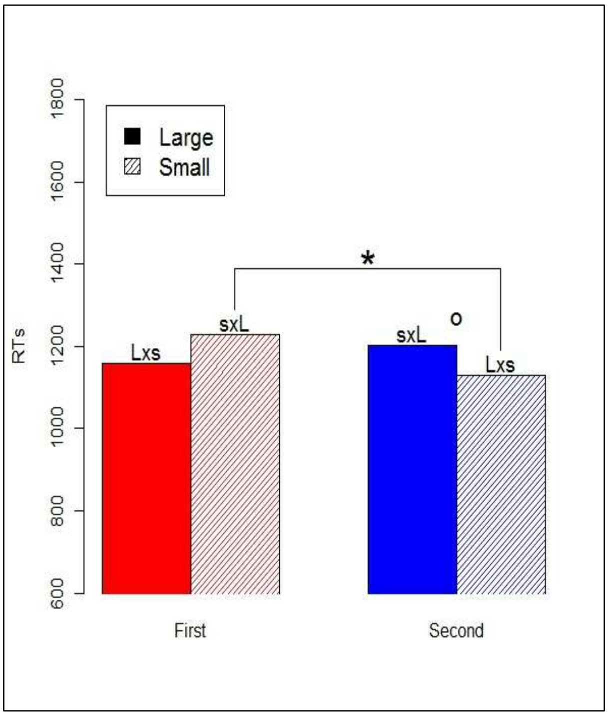


**Figure 4.6: the mean RTs (in millisecond) for the position by distance interaction. Above each bar is presented the order of the operands (L×s: the first operand is larger than the second one; s×L: the first operand is smaller than the second one). °** $p<0.1$**;** *$p<0.05$**.**

Moreover, the position by distance interaction was significant, $F(1,20)=5.21$, $p<0.05$ (figure 4.6). When the incorrect result was a multiple of the second operand the participants tended to respond slower in the large distance condition (e.g., 3×7=14 or 28; 1229 ms) than in the small distance condition (e.g., 7×3=18 or 24; 1130 ms), $t(20)=2.03$, $p=0.06$; when the incorrect result was a multiple of the smaller operand the participants responded slower in the first position condition (e.g., 3×7=18 or 24; 1203 ms) than in the second position condition (e.g., 7×3=18 or 24; 1130 ms), $t(20)=3.13$ ,$p<0.05$.

**A**

| | position | |
|---|---|---|
| distance | first | second |
| small | **3** × 7 = 24 (s×L) | 7 × **3** = 24 (L×s) |
| large | **7** × 3 = 28 (L×s) | 3 × **7** = 28 (s×L) |

**B**

| | position | |
|---|---|---|
| distance | first | second |
| small | 1203 | 1130 |
| large | 1158 | 1229 |

**Table 4.1: A: an example of the position by order interaction. The incorrect result is a multiple of the underlined operand. Below the equation is reported the order of the operands in which the problem is presented. In this example are reported only the incorrect result above the correct result, but the effect is the same with the incorrect result below. B: the mean RTs (in millisecond) in each experimental cell.**

This interaction can be interpreted as the fact that the order of the operands affects the time required to judge the correctness of the equation. In other words, the two factor position and direction can be to trace back to the order of the operands. When in the presented equation the incorrect result was a multiple of the first operand and the first operand was smaller than the second one the equation was in the s×L order; whereas when the incorrect result was a multiple of the first operand and the first operand was larger than the second one the equation was in the L×s order (see first column of the table 4.1). When in the presented equation the incorrect result was a multiple of the second operand and the second operand was smaller than the first one the equation was in the L×s order; whereas when the incorrect result was a multiple of the second operand and the second operand was larger than the first one the equation was in the s×L order (see second column of the table 4.1). In table 4.1an example the problems in which the operands were 7 and 3 is reported. As showed in the table, the s×L order was judged slower than the L×s order regardless the position or the distance of the operand which multiple was presented as incorrect result. This result is another strong evidence than the order of the operand affects the performance with multiplication problems and that the interference product by an incorrect result is stronger for the s×L order than for the L×s order. Unfortunately, although it would be interesting, it has not been possible to statistically evaluate whether there was an interaction with the size of the problems, that is to investigate if the distance by position interaction changes across small, medium, and large problems like the operands-order effect we found in the production task discussed in the chapter 2 and in verification task (for the correct results) discussed above in this chapter, or if this interaction is stable across the size. In fact, adding to the ANOVA the size factor the number of trials in the experimental cells would be too small.

The ANOVA revealed also a trend toward the significance for the three-way interaction (i.e., direction, distance, position), $F(1,20)=3.46$, $p<0.1$ (figure 4.7). The mean RTs in the experimental cells and the difference in RTs between the two levels of the direction factor are reported in the table 4.2. As showed in the box B of the table, the effect of the direction factor tended to be stronger for the s×L order than for the L×s order regardless the levels of the position and distance factors. This three-way interaction suggests that the s×L order was much more sensitive to the interference produced by the manipulation of the incorrect results than the L×s order.

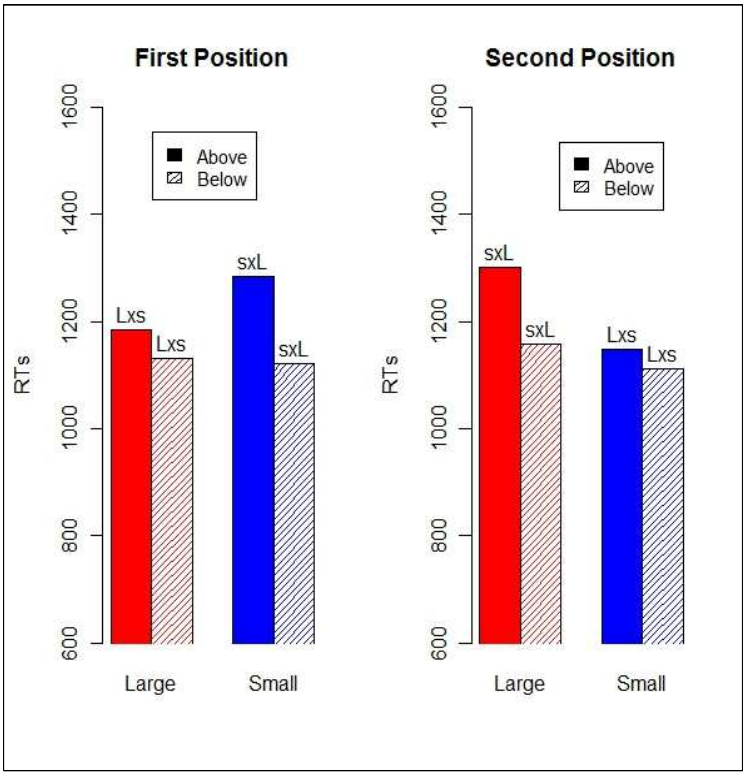


**Figure 4.7: mean RTs for the three-way interaction. On the left there are the mean RTs for the first position condition (when the incorrect result was a multiple of the first operand). On the right there are the mean RTs for the second position condition (when the incorrect result was a multiple of the second operand; on the right). Both graphic show also the mean RTs across direction and distance factors. Above each bar is presented the order of the operands (L×s: the first operand is larger than the second one; s×L: the first operand is smaller than the second one)**

Finally, in table 4.3.A the differences between the mean RTs of the two orders for the correct result condition are reported. The mean difference between the privileged and the not-privileged order is 74.3 ms. Table 4.3.B shows the difference between the mean RTs of the two orders for the incorrect result condition. The difference between the two orders is 72 ms. The differences between the two orders are very similar in the two conditions (74,3 vs 72). Adopting a strong subtractive logic and according to the assumption that only one operands order is stored in memory for each commuted pair, this last result suggests that a reordering process that could allow to access to arithmetic fact acts in both correct and incorrect result conditions and that this process could need (at least in our sample) about 70 ms to reorder the operands.

**A**

| | | position | | | |
|---|---|---|---|---|---|
| | | first | | second | |
| | | direction | | direction | |
| | | above | below | above | below |
| distance | small | 1285 (**s**×L) | 1121 (**s**×L) | 1149 (L×**s**) | 1112 (L×**s**) |
| | large | 1186 (**L**×s) | 1131 (**L**×s) | 1301 (s×**L**) | 1157 (s×**L**) |

**B**

| | | position | |
|---|---|---|---|
| | | first | second |
| | | direction | direction |
| | | above - below | above - below |
| distance | small | 164 (**s**×L) | 37 (**L**×s) |
| | large | 55 (**L**×s) | 144 (**s**×L) |

**Table 1: A: mean RTs (in milliseconds) for the position by direction by distance interaction. B: the differences of the mean RTs (in milliseconds) between the two conditions of the direction factor (above – below). In both tables the incorrect result is a multiple of the underlined operand, and below the RTs is reported the order of the operands in which the problem is presented.**

**A**

| | order | | | |
|---|---|---|---|---|
| size | not-privileged | privileged | difference | mean difference |
| small | 788 | 703 | 85 | 74.3 |
| medium | 1066 | 991 | 75 | |
| large | 1305 | 1242 | 63 | |

**B**

| mean of s×L order | mean of L×s order | difference |
|---|---|---|
| 1216 | 1144 | 72 |

**Table 4.3: A: for each size are presented the mean RTs (in milliseconds) for the privileged and the not-privileged orders, the differences between the RTs of the two order, and the mean of the differences (correct result condition). B: the mean of the mean RTs for the orders reported in the table 4.1.B and the difference between them (incorrect result condition).**

## 4.4 DISCUSSION

The experiment here presented had to aims: 1) to evaluate if the order by size interaction affects the verification task in the same way it affects the production task; 2) to evaluate if inside the arithmetic facts memory the activation spread asymmetrically privileging the forward direction. The second aim concerns also the test on the critical assumption of the *asymmetry hypothesis* about the asymmetric activation spreading around the tie problems. The results show that when the presented result was correct the L×s order was solved faster than the s×L order with small (e.g., $RT(4\times3=12)<RT(3\times4=12)$) and medium problems (e.g., $RT(8\times3=24)<RT(3\times8=24)$); whereas the s×L order was solved faster than the L×s order with large problems (e.g., $RT(7\times8=56)<RT(8\times7=56)$). Thus replicating the results of the production task of Chapter 2 with a different paradigm.

Moreover, when the result was incorrect we found that the multiples above the correct result were rejected with a greater difficulty than the multiples below (e.g., $RT(7\times3=28)>RT(7\times3=14)$). However, the lack of an effect of direction on the tie problems ($RT(7\times7=42)=RT(7\times7=56)$) is not consistent with the *asymmetry hypothesis*. Finally, in the incorrect result condition the position by distance interaction (and the distance by position by direction tendency) revealed that the s×L order was more difficult to reject than the L×s order (i.e., $RT(s\times L)>RT(L\times s)$).

There are at least five interesting inferences that the results of this experiment allow. First, the order by size interaction we found in this experiment is consistent with the result of the experiment of the chapter 2. Second, the lack of an effect of direction in the tie problems shows that the *asymmetry hypothesis* we proposed in the discussion of the results of the production experiment (Chapter 2) cannot explain the order by size interaction, because of the results here reported invalidate the critic assumption than the activation is asymmetric around the ties. On the contrary, according to the *network interference model* (see chapter 1), this results suggest that the tie problems are stored in memory separately by the non-tie problems. The separate representation could explain why the direction effect we found with the multiple around the non-tie problems does not shows up for tie problems. In fact, it could be possible that the activation does not spread from the tie results to the closer multiples (which are non-ties) because of they are stored separately. Therefore, we reject the *asymmetry hypothesis* in favour of the *reorganization hypothesis* (see chapter 2).

Third, the order of the operands and the size of the problem interact in the same way in the production and in the verification tasks. This suggests that both tasks rely on the same "arithmetic knowledge base" (Zbrodoff and Logan, 1990; 2000). Moreover, similar results in both tasks suggest that the order by size interaction in due to the multiplication facts memory, because of this interaction occur independently by the task used (at least in the production and verification tasks).

The fourth, the multiples are asymmetrically associated. There are various studies suggesting that a number is associated with its multiples (Galfano et al., 2003; Galfano et al., 2009; Niedeggen & Rösler, 1999; Rusconi et al., 2004; Rusconi et al., 2006). We can introduce here a new hypothesis about the association between the multiples, that is each multiple is more associated with its above multiple than with its below multiple. The results provide evidence for the asymmetry of activation spreading. Once presented two operands they automatically activate their product, which is associated with the other multiple. However, the asymmetry of this association causes a stronger activation in the multiple above the product than in the one below. Therefore, when presented two operands the multiple above the product is more difficult to reject than the multiple below because of the former is more activated than the later. For example, given the operands 7 and 5, the product 35 is automatically activated; then, since the activation spreads more in the forward direction, the multiple 42 (above) is more activated than the multiple 28 (below). Therefore, the equation 7×5=42 is more error prone and slower to reject than the equation 7×5=28.

Finally, the order of the operands is relevant also when the presented equation is incorrect. The s×L order is more difficult to reject than the L×s order regardless the type of multiple presented as incorrect result. This can be explained by assuming a different strength of association between the two orders and the correct result. The result of both production task and verification task (in the correct result condition) suggest that the L×s order is privileged for the most of the multiplication problems (medium and small problem); whereas the s×L order is privileged only for the large problems. Therefore, the privileged order could activate the correct result more and therefore make it easier to recognize that the presented incorrect result is not the product of the operands. If only one order is stored as arithmetic facts, when the incorrect equation is presented in the stored operands-order (e.g., 7×3=28) the incorrect result (28) can be detected either comparing the whole equation with the multiplication fact or comparing the stored result with the presented result. When the incorrect equation is presented in the non-stored order (e.g., 3×7=28) it could be more difficult to reject the incorrect result because of the operands of the equation do not match with the arithmetic fact and then it could be needed to reorder them in the stored order. Comparing the order effect found in the correct and in the incorrect result conditions we found that the differences between the two order were very similar (74.3 ms and 72 ms, respectively). This could suggest that this reordering mechanism is automatically activated before judging the equation and that it requires about 70 ms to reorder the operands in the stored order before to access to the multiplication facts, differently to what assumed by Verguts and Fias (2005) that state reordering could be done with no behavioural cost.

# Chapter 5

## ERPs correlate of size-dependent reordering preferences

## 5.1 INTRODUCTION

In this Chapter we will present a production task experiment by using the ERPs (event-related potentials) methodology. ERPs can in fact be used to better interpret the effects found in the behavioural experiments of the previous chapters, and especially the interaction between size and order of operands we found for multiplications in Chapters 2 and 4. First, the ERPs methodology has a high temporal resolution which allows us to study the time course of the reordering process we proposed to explain the RTs effects found in the previous chapters. Second, the comparison of the qualitative nature of the effects of size and order, in terms of topography, latency and polarity can aid in disentangle between a same or a different locus of size and order effects. Different ERPs correlates for size and order effects would in fact suggest the existence of a reordering process explicitly different from a mere order preference in terms of weighs between nodes within the associative memory for arithmetic facts.

Only two studies in the ERPs literature on numerical cognition studied order of operands effects in multiplications. Both of them adopted a delayed verification task, one (Zhou et al. 2007) testing a group of Chinese participants (that learn only one half of the multiplication table) and one (Kiefer and Dehaene, 1997) a group of American participants (that learn all the multiplication table). Zhou et al. (2007) compared Mainland Chinese with Hong Kong and Macao Chinese, since only the former study one half of the multiplication table at school whereas the latter study as the Western population all the table. Both the operands and the result were auditorily presented. The participants heard the first operand, then after 50 ms of silence they heard the second operand, and finally after 2300 ms the result was presented. Only the Mainland Chinese group showed an effect of the order of the operands on the ERPs waveform. In this group the L×s order (not taught at school) evoked a large negativity compared to the s×L order (taught). This negativity was evident between 120 and 750 ms after the presentation of the second operand. The topography of the effect varied during this interval: broad distributed on the whole scalp in the early stage, 120-500 ms, and centro-frontal in a second stage, 500-750 ms interval. The Authors interpret this negativity for the not taught order as the ERPs correlated of the "reversal of the operands" mechanism. Moreover, they suggest that the broad distribution of this negativity was due to multiple neural sources involved in the "reversal of the operands" mechanism. Nevertheless, we think that the broad distribution and the fact that the effect lasted from 120 ms to 750 ms suggests not only that multiple neural sources are involved in the effect they found, but that also multiple processes could be involved in the effect described by Zhou et al. (2007) such as, for example, the processing of the size of operands, the comparison of the size of the operands, and the operand reversal mechanism.

In the second and only ERPs study (Kiefer and Dehaene, 1997) that faced the operands order effect in a Western population both a visual and an auditory presentation where adopted. The operands of multiplication problems were sequentially presented with the same timing procedure: the first operand was presented for 150 ms followed by an interstimulus interval of 350 ms, then the second operand was presented for 150 ms followed by an interstimulus interval of 1250 ms, and finally the proposed result was presented. Kiefer and Dehaene (1997) used as stimuli only the multiplications that we classify as small (both operands equal to or smaller than 5) and large (both operands larger than 5). An order effect was found only when the problem was presented in auditory modality. In the 270-397 ms interval after the second operand the s×L problems were more positive than the L×s problems on the temporal sites (bilaterally), whereas on the frontal sites the s×L problems were more negative than the L×s problems. Then, in the 630-1399 ms time interval the s×L problems were more negative than the L×s problems on central sites. The Authors concluded that the operands order effect was likely due to "strategy of reordering the operands" (Kiefer and Dehaene, 1997, p. 25) and distinct from effect of size that is mainly due to a posterior long lasting positivity for larger problems with respect to small ones preceded by a negativity at the same sites. The Authors explain the results within the context of the *triple code model*

(Dehaene, 1992; Dehaene & Cohen, 1995) by assuming that multiplications are stored in the verbal memory only in a given order, without making hypothesis about which order could be the stored one even if they implicitly assume the stored order is L×s since they discuss the effects with this condition as a baseline. According to the Authors the *triple code model* can explain why the effect was present only in the auditory format and not in the visual format. In fact, according to the *triple code model* the reordering process is different between auditory and visual presentation. In the auditory presentation the process is: 1) the operands are compared; 2) if not presented in the stored order they are reordered; 3) the reordered sequence can access to the rote verbal memory for the arithmetic facts. In the visual presentation (and assuming for example than the multiplication facts are stored in the s×L) the process is: 1) the operands are compared; 2) convert the smaller operand into a verbal representation; 3) convert the larger operand into a verbal representation; 4) access to the rote verbal memory. Since with a visual presentation the operands have to be converted into a verbal code regardless the order in which they are presented, the order has not effect on the processes that allow to retrieve the result.

Zhou et al. (2007) found that the non-privileged order evoked negativity compared to the privileged order in the Chinese group they tested. Despite there is no clear evidence of a preferred order for English speakers, following the implicit assumption that the L×s order is the privileged, Kiefer and Dehaene (1997) also found a negativity at central sites (in the 630-1399 time window) and at frontal sites (in the 270-390 time window) that was interpreted to be the ERPs signature of a reordering process. According to these studies we expect to find that in an Italian group the non-privileged order (solved slower in the experiment reported in Chapter 2 and 4) to exhibit a central-frontal negative effect.

Differently from Zhou et al. (2007) and Kiefer and Dehaene (1997) we have however used a delayed production task instead of a verification task. Despite the fact that we were able to replicate the interaction size by order found in production (see Chapter 1) in a verification task (see Chapter 4), the effect was larger and more clear with the former task. The logic behind delayed tasks is that ERPs deflections cannot be clearly interpreted if response preparation and execution potentials are superimposed with cognitive potentials. For this reason, especially when large differences in RTs can be expected across participants and conditions (size effect), a delayed paradigm is very useful in distinguishing early perceptual and medium latency, cognitive potentials from potentials linked to task execution. However, differently from a verification task where the response cannot be prepared or anticipated until the result is presented, in a production task the motor preparation can affect the ERPs even before the presentation of a cue. We thus decided to ask participants to respond in different modalities, typing versus speaking, depending on the cue appearing after the delay. This allows implementing a delayed production task in which the participants cannot prepare the motor response before the cue is presented. Moreover a variable time interval between problem and cue can also aid in avoid superimposition of problem processing stage and response selection and execution.

A further difference with respect to the literature is the use of self-report that we have seen in Chapter 2, which can give reliable information in line with behavioural data. Moreover, we can use the self-report in order to select the problems for which participants report to use inversion of the operand as an explicit and aware strategy, and compare the waveforms elicited by this condition with all the others. In our knowledge, this kind of analysis has not been used as far in ERPs research on arithmetical cognition but it has been efficiently used by Grabner et al. (200) in a fMRI study.

Before presenting our experiment we will briefly review some other ERP studies that have investigated the electrophysiological correlated of cognitive arithmetic, especially those studies that used a delayed paradigm. Since, beside order we will also manipulate the operation (additions vs multiplications) and clearly the size, given our interest in size by order interaction in multiplications.

Two experiments (Zhou et al., 2006; Zhou et al., 2011) compared additions and multiplications by adopting a verification task. In Zhou et al. (2006) addition, subtraction, and multiplication problems were presented, while in the experiment of Zhou et al. (2011) only addition and multiplication were presented[8]. In both experiment participants had to judge if a visually presented equation was correct (e.g. 7×4=28) or not (e.g., 7×4=35). In Zhou et al. (2006) the proposed result (e.g., 28 or 35) was showed to the participants only 1300 ms after the presentation of the operands (e.g., 7 and 4), while in the experiment of Zhou et al. (2011) the operands and the results were presented at the same time. Despite this differences both study showed similar results, that is multiplication problems elicited frontal left negativity compared to addition problems, in the 275-334 ms interval after the presentation of the operands in Zhou et al. (2006) and in the 400-900 ms interval after the presentation of the whole problems in Zhou et al. (2011). In both studies the Authors interpreted the effect as an evidence that verbal processes are more involved in solving multiplication than in additions.

The problem-size effect (i.e., the large problems are slower to solve and more prone to errors than small problems) is one of the more (if not the most) debated effect in literature about the arithmetical cognition. Therefore, we have compared also problems with different size in both additions and multiplications. The size effect, as well in the other studies reported above, has been mainly investigated adopting a verification task. With respect to addition problems, various studies have found a than the large problems elicit positivity compared to small problems at the centro-parietal sites (El Yagoubi, Lemaire, & Besson, 2003; Kong et al., 1999; Núñez-Peña, Cortinas & Escera, 2006; Núñez-Peña & Escera, 2007; Núñez-Peña, Gracia-Bafalluy, & Tubau, 2011). This positivity for large problems has been interpreted as associated to more demanding process and to the procedures selection. Namely, according to this interpretation large addition problems are more often solved with non-retrieval procedures which are associated with a greater mental effort.

[8] Given the aim of our experiment we report only the result of Zhou et al. (2006) about multiplication and addition. In the study of Zhou et al. (2001) participants were both adults and children. Here we will report only the result of the adult group.

With respect to multiplications, a study of Jost, Hennighausen & Rosler (2004b), in which a delayed verification task with sequential presentation of the operands was used, large problems evoked a right lateralized negativity with respect to small problems between 400 ms and 510 ms. The Authors interpreted this negativity as associated to the use of different procedures between large and small problems, given that it is likely that large problems rely more on non-retrieval procedures than small problems. In another experiment Jost and Colleagues (Jost, Beinhoff, Hennighausen, & Rosler, 2004a) adopted a implicit production task in which the participants were sequentially presented with two one-digit numbers. Three-thousand ms after the second number a third one-digit number was presented, the participants had to add that number to the product of the first two numbers and to judge if the following proposed result was correct or not. Between 900 ms and 2700 ms after the presentation of the operands of the multiplication, the large multiplications evoked a long lasting negativity with respect to small multiplications localized mainly bilaterally at frontal sites and over right temporal sites. The Authors interpret the larger negativity as associated with the larger mental load and with the more frequent use of non retrieval procedures in large multiplication.

Another study that is worthwhile to cite is Pauli et al. (1994) that used a production task in which participants had to perform multiplications in different sessions, thus manipulating practice. A central-parietal positivity was associated with practice. The participants had to perform four sessions in different days and the EEG was recorded in each session. The Authors found a frontal-central positivity diminishing with the practice, that is the positivity became smaller from session to session. Moreover, the focus of the positivity move on the centro-parietal regions in the last sessions. This pattern of brain activity was interpreted as the fact that two different processes are involved in the earlier and the later sessions. In the earlier sessions, when the performance in less automatized, the role of the frontal lobe (supposed to have a relevant role in generating the fronto-central positivity) is stronger, whereas in the later session the centro-parietal regions (supposed to be the area associated with the arithmetic facts memory) show stronger positivity because the performance rely more on automatic procedures that involve the “retrieval of arithmetic facts from the cortical network” (Pauli et al., 1994, p. 28). Capitalizing on these results and interpretation, one can expect different topographies associated with the two macro-areas (frontal vs centro-parietal) distinguishing effects that are due to less automatized process (frontal areas) or to more automatized process (centro-parietal areas).

To summarize, a difference between multiplications and additions has found in terms of a frontal left negativity for the former with respect to the latter, and this effect is interpreted as an indication that multiplications rely more on verbal processes than addition problems. Size effects are different for multiplications and additions but typically start after 300 ms and show a long lasting time development. Large multiplications typically show a larger negativity mainly on anterior and right temporal area, while large additions  show posterior positivity with respect to small ones. In both cases the effects have been interpreted to the use of non-retrieval procedures in large problems compared to small problems. Specifically the posterior positivity for large additions as a correlate of the selection of non-retrieval procedures (Núñez-Peña et al., 2011), while anterior negativities for large multiplications to memory load or more generically mental effort (Jost et al., 2004a).

Despite the relevance of the finding of these studies in the arithmetical cognition, it has completely been underestimated the role of the retrieval procedure both in producing the size effect (see the models reported in the Chapter 1) and in generating the ERPs differences found. ERPs correlates of spreading of activation within the retrieval process have been instead addressed by a number of studies, not reviewed here, that looked at ERPs elicited by the presentation of the result in verification tasks (see for example Niedeggen, Rosler; & Jost, 1999; Niedeggen & Rosler, 1999; Prieto-Corona Prieto-Corona, Rodríguez-Camacho, Silva-Pereyra, Marosi, Fernández, & Guerrero; 2010) where an N400 effect proportional to the association between the problems and the results was found. Even for the ERPs elicited by problems it is possible that some of the effects could be also due to differences within retrieval procedures in terms of the difficulty to activate and select the result within the associative network that stores arithmetic facts memory. Despite the non-retrieval procedures play a relevant role in the process of solution of the arithmetic problems (see Chapter 1), classical models assume in fact that the direct retrieval is likely to be the most used procedure adopted by adults to solve simple one-digit arithmetic problems.

Our predictions, based on the literature are thus that reordering processes will elicit a central (Kiefer and Dehaene, 1997) or frontal (Zhou et al, 2007) negativity. As well, possibly at a later latency, effects of this type should also distinguish between additions and multiplication, and size effects for multiplications (possibly with a more broad, right lateralized focus). Size effects for additions should elicit a larger posterior positivity for larger problems.

## 5.2 METHOD

**Participants.** Twenty-three (14 females; mean age: 26.5, *sd*: 4.71) from the University of Trento participated in the experiment. All participants were native Italian speakers and had normal or corrected-to-normal vision. All the participants were right-handed and were refund 20 euro on their participation in the experiment.

**Material.** The participants were asked to perform two tasks: a delayed chronometric task and a self-report task. During the delayed chronometric task the EEG of the participants was recorded. In both tasks the stimuli were the same. Single-digit multiplication and addition problems were used during the experiment (tie problems included). The problems with 0 or 1 (e.g., 0×3, 0+5, 1×3, etc.) were excluded because they are likely solved by means of rules (LeFevre et al., 1996a). Therefore, there were 64 problems for each operation. In order to have a sufficient number of trials for the EEG analysis, in the delayed chronometric task each problem was presented three times. Therefore, there was a total of 192 problems for each operation (384 problems in the whole chronometric task). Problems of different operations were presented in different blocks (see below), half of the participants performed additions before multiplications, half the converse. Block order in the self-report was the same as in the chronometric task. In both the delayed chronometric task and in the self-report tasks the order in which the problems were presented within each block was randomized for each participant.

**Procedure.** In the delay chronometric task each operation (multiplication and addition) was presented in separated blocks (6 blocks of 64 problems each). In order to familiarize with the experimental procedure, before each block the participants performed some practice trials with problems with 0 and 1 as operands (e.g., 0×3, 0+5, 1×3, etc.). There were six blocks, three for each operation. In each block only one operation was presented. The order in which the operations (addition and multiplication) were presented was alternated and balanced across participants. For example, the first participant performed the operation in the following order: multiplication (block 1), addition (block 2), multiplication (block 3), addition (block 4), multiplication (block 5), addition (block 6). The second participant performed the operation in the following order: addition (block 1), multiplication (block 2), addition (block 3), multiplication (block 4), addition (block 5), multiplication (block 6). The problems were sequentially presented at the centre of a monitor of a PC. Each trial started with a fixation point (“#”). The fixation point lasted on the monitor until the participant pressed a key on the keyboard. Once pressed the key the first operand was presented for 300 ms, followed by the sign of the operation (“+” or “×”) for 300 ms, and by the second operand for 300 ms. After the second operand, the equal symbol (“=”) was presented. The equal symbol was used to help the participant to do not move the eyes and to maintain the fixation on the centre on the screen during the delay. The equal symbol could have lasted on the monitor for 1700, 2000, 2250, or 2500 ms. This four delay intervals were used to reduce the expectation of the cue at a fixed time after the equal symbol. Each interval was used the same number times in each block. After the equal symbol, a cue was presented on the monitor. We used two cues: one representing a finger and one representing a mouth. The participants had to respond by using the keyboard when the “finger” cue was presented and by spelling the result verbally when the “mouth” cue was presented. When the answer had to be performed with the keyboard, the participants were required to respond with the right hand, by using the numeric keypad on the right of the PC keyboard just after the presentation of the “finger” cue. The participants had to press the keys corresponding to the digit of result of the problem (one key if the result was with one digit; two keys if the result was with two digits). When the answer had to be performed verbally it was recorded with a microphone, the participants were required to verbally respond just after the presentation of the “mouth” cue. The use of two different response type (verbal or manual) was implemented to avoid any motor preparation. In fact, with this delay double response procedure the participant did not know how to respond until the cue was presented. This was done to avoid that participants could prepare their response in advance since this can generate ERPs deflections due to motor preparation brain activity. The participants were instructed to be “as quick and accurate as possible”, and the response cue lasted on the screen until the participants responded. However, if the participants did not responded within 9 seconds the cue disappeared and the fixation point of the next trials was showed on the screen. The operands and the operation signs had a dimension of about 1 cm and the participants were at about 80 cm from the monitor. RTs and accuracy of the keys pressed  (one or two according to the number of digit of the result) and the voice response were recorded. Between the blocks the participants could have taken a little break.

After the delayed chronometric task, the participants had to perform the self-report task on a notebook computer in which they had to solve the same arithmetic problems. The self-report was similar to that performed after the behavioural experiment done after the production experiment described in chapter 2. The order in which the operations blocks were presented was the same as in the previous tasks. In this task each problem of both operation was presented only once. The participants had to report for each problem the result and the procedure used to solve the problem, and they were required to be "as accurate as possible" without time pressure (they might take all time they need to solve the problem and report the strategy). Before starting the task a sheet with the description of the procedures was given to the participants, who could take the sheet during the task to remember the procedures description. There were 5 procedures among which the participants could choose: retrieval, transformation, counting, inversion, and other. On the sheet given to the participants the procedures were described as following:

- Retrieval: "you remember the solution of the problem, that is you retrieve the result directly by the memory".
- Transformation: "you solve the problem by using other problems that can be members of the same arithmetical operation or of another operation (e.g., you solve the problem 9×9=? by using 9×10=90–9 = 81)".
- Counting: "you solve the problem counting (maybe in a quiet voice) a certain number of times until you obtain the result of the problem (e.g., you solve the problem 4×4 by counting 4..8..12..16; or you solve the problem 13+4 by counting 13..14..15..16..17)".
- Inversion: "you reverse the two operands to be able to find the result of the problem (e.g., you solve the problem N1×N2 by using the problem N2×N1)".
- Other: "you solve the problem by using another procedure or you are not sure about the procedure used".

The problem and the procedures to select were presented together in the same screen. Therefore, unlike in the chronometric task, the operands and the sign of the problem was simultaneously presented, and they remained on the screen until the participant reported the result and the strategy. The participants were required to solve first the problem and then to select the used procedure. The participants had to use the numeric keys on the notebook keyboard to report the result and the mouse to select the strategies. The problems were presented on the screen with on the right a white space in which the participants had to report the result of the operation (the white space had been selected with the mouse before write the result). Below the problem there was the strategies (5 options). Once the participant filled in all the required information they could go to the next trials by pressing the “Enter key”. If the participant forgot to fill in one or more information a message dialog appeared on the screen asking to complete all the sections. The participants were asked to report the procedure associated to the problem solved during the self-report and not trying to remember how they solved the task during the delay chronometric experiment.

### 5.2.1 Behavioural data analysis

The statistical analysis we used were the same for both multiplication and addition. The two operations have been analysed separately. For the chronometric task, on RTs a three-way ANOVA was separately performed with size, order, and response type as within factors. The size factor included three levels: the problems with both operands larger than 5 were coded as “large” (e.g. 7×8); the problems with one operand larger and one smaller than 5 were coded as “medium” (e.g. 7×3); the problems with both operands smaller than 5 were coded as “small”. Both orders of the problems 6×5, 7×5, 8×5, and 9×5 were coded as “medium”, whereas both orders of the problems 2×5, 3×5 and 4×5 were coded as “small”. The order factor had two levels: L×s (or L+s for addition) and s×L (or s+L). The response type factor had two levels: manual (the participant had to respond with the keyboard) and vocal (the participant had to respond with the microphone). For each participant we calculated the mean RTs in the twelve experimental cells (order X size X response type). The analysis of the RTs was performed on the correct trials. In case the main effects or the interactions were significant, the post-hoc analysis were performed by using t-test corrected with the FDR method. ANOVAs were Greenhouse-Geisser corrected when the degrees of freedom of a factor exceeded one (uncorrected degrees of freedom and epsilon values are reported). The accuracy (proportion of correct answers) has been only qualitatively analysed in order to evaluate the participants performed the task with an adequate accuracy.

The self-report results have not be statistically analysed, but they have been used only to evaluate the use of the inversion procedure in the EEG analysis (see below).

### 5.2.2 EEG recording and data analysis

*EEG recording*

The EEG was recorded using a 64 channels BrainAmp amplifier (Brainproducts, gmb). Sixty-two electrodes were placed on scalp sites (Fpz, Fp1, Fp2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT8, T7, C5, C3, C1, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, A2, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2, F9, F10, Cz) according to the 10% system with the aid of an elastic cap (Easycap, Gmb). Additional electrodes were also placed on left and right mastoids (A1, A2), below the left eye (Ve1). The ground site was placed frontal to Fz (AFz site). Data, referenced to the left mastoid (A1), were amplified and filtered with a band-pass filter with a high pass time constant of 10s and a 100 Hz low-pass cut-off, and digitalized at 250 Hz (amplitude resolution 0.1 μV). Impedance was kept below 10kΩ and for most of the channels it was below 5kΩ. Trigger were sent from the stimulation program to the EEG recoding system trough a parallel port for the onset of each operand, the operation symbol, the equal symbol, and the response cue.

*ERPs waveform extraction*

Data were analysed using the EEGLAB (Delorme & Makeig, 2004), an open source Matlab© toolbox. After the recording, data were re-referenced to the linked mastoids and further filtered with band pass filter (0.08Hz – 30Hz). Noisy channels, detected by visual inspection, were interpolated using spherical interpolation (15 channels in the whole pool of 23 participants). Marker information for the second operand was enriched by integrating all the useful information extracted from the output of the stimulation program, allowing selective averages of epochs in different experimental conditions, for 3 participants (1, 16, 19) this was not possible given a misalignment of marker information, probably due to pauses or interruptions of the EEG recording during the experiment.

Epochs, from 500 ms before to 1600 ms after the onset of the second operand, were extracted after an automatic artifacts rejection procedure. This procedure rejected epochs where the amplitude of the EEG, after a pre-stimulus baseline correction, exceeded +/- 70μV for channels on sites around eyes (F9, F10, Fp1, Fp2, Ve1) and +/- 90μV for all the channels.

Data from 4 participants (17, 18, 20, 23) were excluded from subsequent analysis given the high number of epochs affected by artifacts (mainly blinks). Within the resulting pool of data from 16 participants the average number of residual epochs were 192 (mean 84.47%, median 91.02%). Single participant average waveforms in the different experimental cells of interest (see below) were computed for grandaverage calculation and statistical analysis.

*ERPs analysis*

Single participant averages were calculated for each experimental cell of interest (see below) and then the statistical analysis was performed using the software R-project and the library Ez. Data were clustered by averaging single channel data on 10 groups of sites (see table 5.1) to simplify graphical representation and reduce the number of degrees of freedom in the statistical analysis. The resulting 10 virtual sites (hereafter, called simply 'sites') can in fact be used for two separate analyses. The first analysis considered midline sites with a single topographic factor (*Longitude*, 4 levels: FP, FC, CP, PO), while the 6 lateralized sites can be organized into 2 topographic factors (*Longitude*, 3 levels: F, C, P; *Lateralization*, 2 levels: L, R) as shown in table 5.2.

| FP | FC | CP | PO | AL | AR | CL | CR | PL | PR |
|---|---|---|---|---|---|---|---|---|---|
| Fp1 | FC1 | CP1 | PO3 | F9 | F10 | FT7 | FT8 | TP7 | TP8 |
| Fpz | FCz | CPz | POz | F7 | F8 | FC5 | FC6 | CP5 | CP6 |
| Fp2 | FC2 | CP2 | PO4 | F5 | F6 | FC3 | FC4 | CP3 | CP4 |
| F1 | C1 | P1 | O1 | F3 | F4 | T7 | T8 | P7 | P8 |
| Fz | Cz | Pz | Oz | AF7 | AF8 | C5 | C6 | P5 | P6 |
| F2 | C2 | P2 | O2 | AF3 | AF4 | C3 | C4 | P3 | P4 |

**Table 5.1: table reporting the groups of channels used for graphic display and statistical analysis. Top column names in bold are the names of the new sites (FP: prefrontal, FC: fronto-central, CP: central-parietal, PO, parietal-occipital, AL: anterior-left, AR: anterior right, CL: central left, CR: central right, PL: posterior left, PR: posterior right), the new sites were computed as the average for each time point of the event related potential values of the sites reported in each column.**

| ch | lat | long |
|---|---|---|
| AL | L | A |
| AR | R | A |
| CL | L | C |
| CR | R | C |
| PL | L | P |
| PR | R | P |

**Table 5.2: definition of the two topographic factor used for the analysis of the lateral clusters defined in table 5.1: Lateralization (L:left, R:right) and Longitude (A: anterior, C: central, P: posterior)**

Repeated measure ANOVAs have been performed using as dependent variable the mean voltage amplitudes in specific time intervals selected on the basis of qualitative analysis of the grand-averages plots. The qualitative analysis was performed both on clustered plots and on original single channel plots (not reported here) for the specific comparisons of the experimental manipulations under study (i.e., operation type, size, and order). These experimental manipulations, together with the topographical factors described above, constitute the independent variables for the ANOVAs. Greenhouse-Geisser corrections for deviations from the sphericity of variance were computed when numerator degrees of freedom were larger than one, uncorrected degrees of freedom are reported with the Greenhouse-Geisser epsilon value and corrected p-values. In order to make the results more readable the effects involving only topographic factors have not been reported.

Given the large number of experimental conditions we proceed by analyzing first the overall comparisons between *Additions* and *Multiplications*[9], followed by the comparisons of *Size* and *Ties* effects for each operation independently from the order of the operands. The aim of these first analyses is to understand overall main effects in the ERPs waveforms elicited during the interval between the presentation of the second operand and the response cue before analyzing the specific conditions of interest for the present work, that is the order by size interaction (especially for multiplications for which the experiments reported in chapters 2 and 4 allow specific hypotheses described in the introduction). This analysis scheme with multiples ANOVAs performed on the same sets of data may be considered incorrect from a strictly statistical way. However, we consider the separate ANOVAs for multiplications and additions (with order and size as factors) to verify the hypotheses of the previous experiments (order by size interaction), whereas the other comparisons are performed just to qualitative compare the present results with previous ERPs studies, where *Size* and *Order* variables have never been manipulated together.

[9] The italic is used to distinguish between the arithmetic operation (multiplication and addition) and the factors used in the ANOVAs (*Additions* and *Multiplications*). The same typography rule is used for the other factors of the ANOVAs.

## 5.3 RESULTS

### 5.3.1 Behavioural results

The following analysis has been performed by calculating the accuracy and the RTs only for the participants for which the data was used in the EEG analysis (16 participants, see above). In the chronometric task, the participants are very accurate in both multiplication (mean: 95% of correct responses; median: 97%; range: 84-100%) and addition (mean: 98%; median: 98%; range: 92-100%), and in both manual condition (mean: 96%; median: 96%; range: 83-99%) and vocal condition (mean: 97%; median: 98%; range: 93-100%). Table 5.3 shows the accuracy for all the experimental cells given by the three factors: size, order, and response type.

A

| | | response type | |
|---|---|---|---|
| Size | order | manual | vocal |
| Large | Lxs | 0.91 | 0.89 |
| medium | Lxs | 0.95 | 0.97 |
| small | Lxs | 0.98 | 0.99 |
| Large | sxL | 0.85 | 0.94 |
| medium | sxL | 0.97 | 0.97 |
| small | sxL | 0.99 | 0.99 |
| Tie | tie | 0.92 | 0.97 |

B

| | | response type | |
|---|---|---|---|
| Size | order | manual | vocal |
| Large | Lxs | 0.98 | 0.97 |
| medium | Lxs | 0.97 | 0.98 |
| small | Lxs | 0.97 | 0.98 |
| Large | sxL | 0.95 | 0.95 |
| medium | sxL | 0.98 | 0.99 |
| small | sxL | 0.99 | 0.97 |
| Tie | tie | 0.96 | 0.99 |

**Table 5.3: the accuracy in the experimental cells for both multiplication (A) and addition (B).**

For both multiplications and additions we performed a repeated measure ANOVA on the RTs with size, order, and response type as within factors[10]. For multiplication the ANOVA revealed that the participants responded faster in the vocal condition (729 ms) than in the manual condition (1236 ms), $F(1,15)=54.13$, $p<0.001$. The main effect of size was also significant ($F(2,30)=27.74$, $\varepsilon_{GG}=0.55$, $p<0.001$): large (1166 ms), medium (911 ms), and small problems (871 ms). Finally, the size by response type interaction was significant, $F(2,30)=15.88$, $\varepsilon_{GG}=0.70$, $p<0.001$. This interaction is given by the fact that the size effect was larger in the manual condition (large: 1500 ms, medium: 1122 ms, small: 1086 ms) than in the vocal condition (large: 832 ms, medium: 699 ms, small: 655 ms). The presence of the size effect after the presentation of the cue can be just due to the fact that larger results require more motor planning to be produced (both in the manual and verbal conditions) even if we cannot exclude that the problems were not all completely solved before the cue presentation.

For additions the ANOVA reveals that the participants solved faster the problems in the L+s order (855 ms) than in the s+L order (882 ms), $F(1,15)= 6.38$, $p<0.05$. This result suggests that the order of the operands is also relevant in a delay production task. The difference between the two orders however is very small (only about 30 ms) with respect to the differences that emerged in the non-delayed production task (about 130 ms, see chapter 1). Moreover, the participants responded faster in the verbal condition (682 ms) than in the manual condition (1055 ms), $F(1,15)=38.33$, $p<0.001$. Finally, the size by response type interaction was significant, $F(2,30)=4.95$, $\varepsilon_{GG}=0.84$, $p<0.05$. This interaction is given by the fact that in manual condition the medium problems (995 ms) were solved faster than large (1089 ms) and small problems (1082 ms); whereas in vocal condition small problems (644 ms) were solved faster than large (702 ms) and medium problems (700 ms). This interaction

[10] Each problem was repeated three times in the different blocks. The block factor (three levels: “first block”, “second block”, and “third block”) has been analysed in ANOVAs not reported here. These ANOVAs showed that, even if the block factor was significant (participants solved the problems in the third block faster than in the first block), it did not interact with the other variables.

is given by the bizarre fact that for manual responses medium problems were solved faster than smalls. We have no explanation for this difference that is likely to be due to the delayed procedure we used.

For the self-report task we simply qualitatively analysed the proportions of use of the inversion procedures reported by the participant in the trials used in the EEG analysis (see below). Table 5.4 report the proportion in the experimental cell given by the size and order factors in the two operations. As we expected the participant reported to use inversion much more often in the non-privileged orders than in the privileged orders.

| order | size | N | tot. | prop. | operation |
|---|---|---|---|---|---|
| sxL | large | 62 | 126 | 0.49 | addition |
| Lxs | large | 7 | 126 | 0.05 | addition |
| sxL | medium | 158 | 336 | 0.47 | addition |
| Lxs | medium | 6 | 336 | 0.02 | addition |
| sxL | small | 42 | 126 | 0.33 | addition |
| Lxs | small | 0 | 126 | 0 | addition |
| sxL | large | 48 | 288 | 0.17 | multiplication |
| Lxs | large | 92 | 288 | 0.32 | multiplication |
| sxL | medium | 291 | 768 | 0.38 | multiplication |
| Lxs | medium | 56 | 768 | 0.07 | multiplication |
| sxL | small | 54 | 288 | 0.19 | multiplication |
| Lxs | small | 0 | 288 | 0 | multiplication |

**Table 5.4: for each experimental cell the proportion of use of the inversion procedure is reported. N: number of times the participants report to use inversion; tot.: the total number of trials in the cell used in the EEG analysis; prop.: the proportion of the use of the inversion procedure.**

### 5.3.2 EEG results

We first compared ERPs for additions and multiplications for an overall qualitative analysis. Then, we compared for each operation the effect of size along the three levels of the *Size* factor used in the previous behavioural experiments (*Small, Medium, Large*) and keeping *Ties* separate. Finally, given our hypothesis and the results of the previous behavioural experiments, we compared the combined effect of operand order and size.

The comparison between *Additions* and *Multiplications* (see figure 5.1) shows an early deflection between 250 ms and 350 ms on central-frontal sites where *Multiplications* elicit a larger bilateral negativity. This early effect is followed by a long lasting sustained slow wave for the whole epoch, where *Multiplications* show a larger positivity with respect to *Additions*. The latter effect emerges mainly on posterior parietal sites and appears to be slightly larger over the left hemisphere (especially when the effect on PL and PR sites are visually compared).

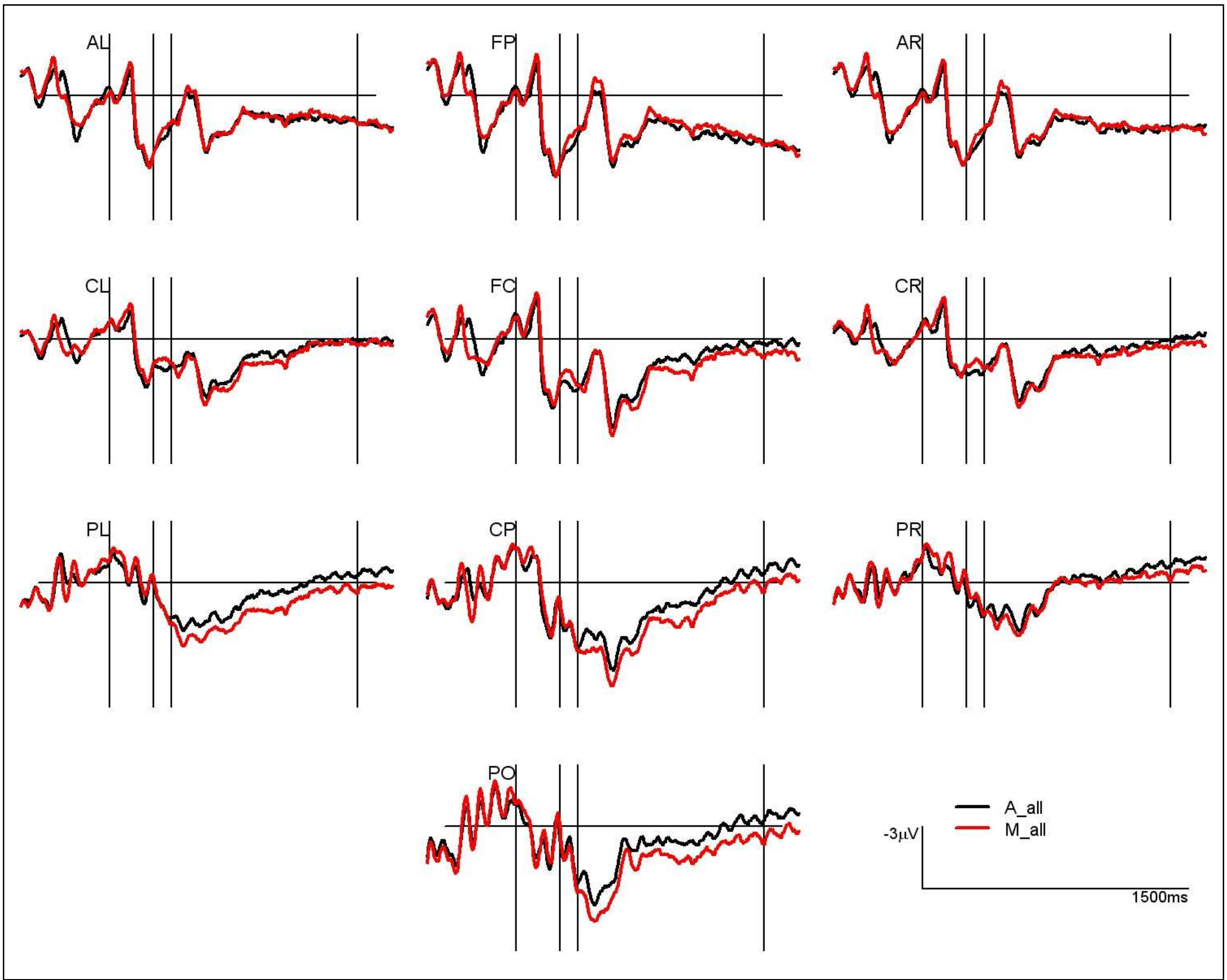


**Figure 5.1: Grandaverage plot for the comparison between additions (A_all), in black, and multiplications (M_all) in red. Vertical lines are plotted at the onset of the second operand, and after 250ms, 350ms and 1400ms.**

For midline sites, the ANOVA in the 250-350 ms time window shows a marginally significant interaction between *Longitude* and *Operation* (2 levels: *Additions, Multiplications*), F(3,45)=2.85 , $\varepsilon_{GG}$=0.5, p<0.1. No significant effects emerge for lateral sites. The ANOVA in the 350-1400 ms interval shows for the midline sites a significant *Longitude* by *Operation* interaction, F(3,45)=7.33, $\varepsilon_{GG}$=0.5, p<0.05. In the same time window, the analysis on lateral sites shows significant *Longitude* by *Operation* interaction (F(2,30)=4.84, $\varepsilon_{GG}$=0.54, p<0.05), and *Longitude* by *Lateralization* by *Operation* interaction (F(2,30)=6.25, $\varepsilon_{GG}$=0.74, p<0.05), confirming the posterior and left lateralized nature of the described effects.

### ***Size effect for Multiplications***

Figure (figure 5.2) shows the grandaverages for the three levels (*Small*, *Medium*, and *Large*) of the *Size* factor for multiplications. The *Ties* are plotted separately, independently from size. A large positivity for *Ties* with respect to all the other three *Size* levels is evident between 300 and 500 ms, indicating a very different processing of ties problems. The comparison between the three levels of *Size* factor shows a long lasting positivity for larger problems starting from about 300 ms. This lasting positivity for larger problems is initially focused on parietal sites, but at later latency it is more widely distributed on the scalp for *Large* problems with respect to *Small* ones. The *Medium* problems elicited a waveform somewhat in the middle between *Large* and *Small*, but more closer to those elicited by the *Small* problems. Starting from about 800 ms post-stimulus, a sustained frontal negativity for *Small* problems relatively to both *Medium* and *Large* problems emerges. Even if both slow waves can be interpreted as parts of the slow-waves family of components, the different topography suggests that the latter effect, despite the polarity differences are always in terms of a larger positivity for *Large* (more difficult) problems, is likely to be distinguished from the positivity previously discussed.

The statistical analysis of the so far described effects were preformed into two stages. First, we analysed the 300-500 ms interval with the aim to statistically assess the difference between *Ties* and the other *Size* levels (small, medium, and large). Second, given the clear difference between *Ties* and the other *Size* levels in the ERPs, we proceeded by analyzing *Size* factor without considering *Ties* anymore in three time windows: 300-500 ms, 500-800 ms, and 800-1400 ms.

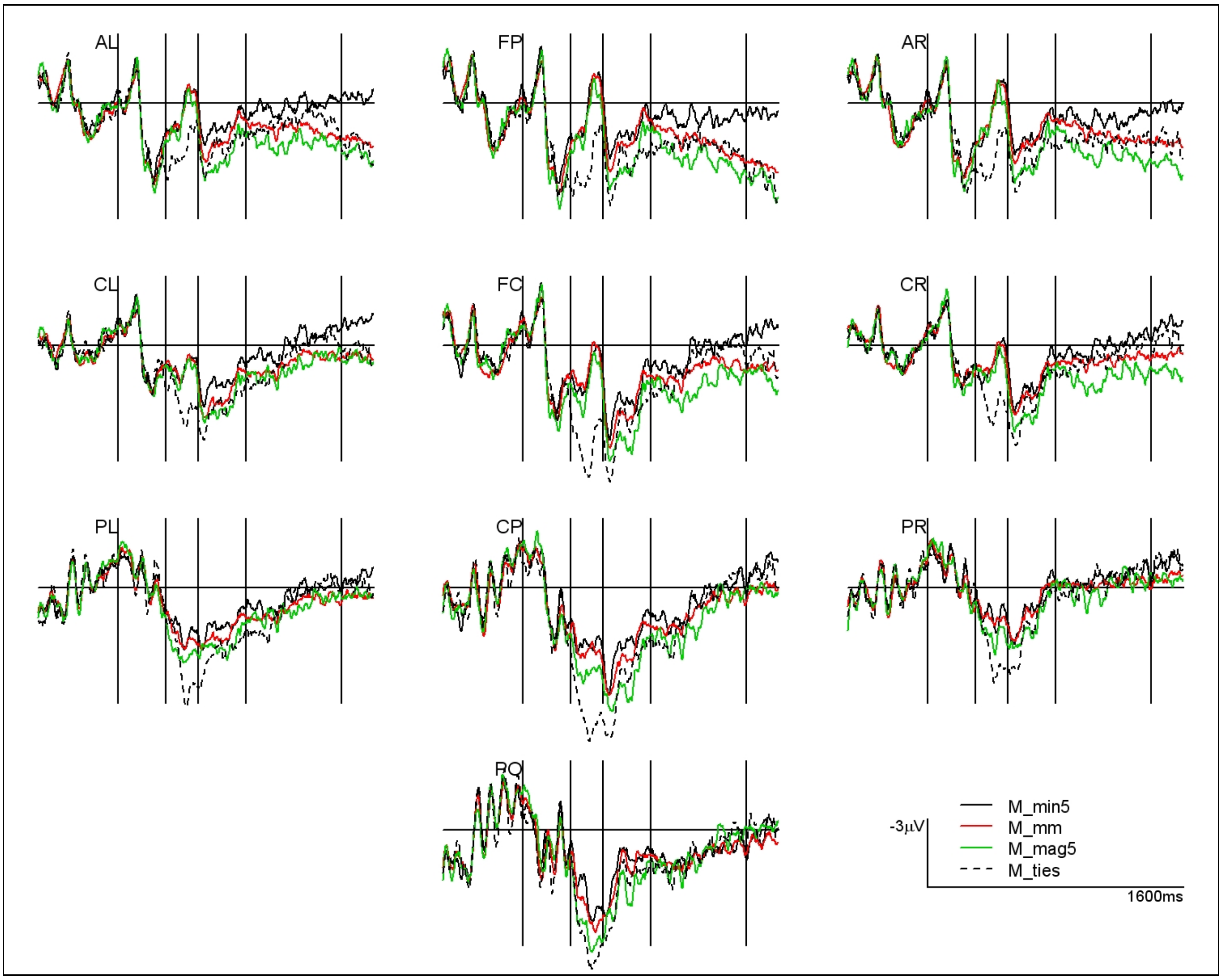


**Figure 5.2: Grandaverage plot for the size and ties comparison for multiplications. *Ties* problems (M_ties) are plotted in black dashed, *Small* problems (M_min5) in black, *Medium* problems (M_mm) in red, and *Large* problems (M_mag5) in green. Vertical lines are plotted at the onset of the second operand, and after 300ms ,500ms ,800ms ,1400ms.**

The ANOVA in the 300-500 ms for midline sites gives a main effect of four levels *Size+Ties* factor (F(3,45)=19.41, $\varepsilon_{GG}$=0.60, p<0.01), and an interaction of the same factor with *Longitude* (F(9,135)=3.98, $\varepsilon_{GG}$=0.39, p<0.01). Similarly, the ANOVA in the same time window on lateral sites shows a main effect of *Size+Ties* (F(3,45)=12.60, $\varepsilon_{GG}$=0.67, p<0.01). However, the *Size+Ties* factor does not interaction with the two topographical factors (*Longitude* and *Lateralization*). To compare the *Ties* with the three levels of the *Size* factor we performed post-hoc t-test analysis with FDR correction for each site of both midline and central line. All differences were significant with the only exception of the difference between *Ties* and *Large* problems in the PO cluster (difference 0.45uV, t(15)=1.61, p>0.1). For all the other comparisons the difference between *Ties* and the other three *Size* levels was between 0.54uV and 3.27uV (median 1.73uV), with all t(15)>2.19 and all p<0.05.

After having confirmed the large widespread positivity for *Ties* with respect to all other three *Size* levels, we considered the effect of *Size* only. In the 300-500 ms time windows, both ANOVAs on midline and on lateral sites restricted to the *Size* factor did not give any significant results, but only a marginal significant effect of *Size* in the midline analysis (F(6,90)=1.52, $\varepsilon_{GG}$=0.72, p<0.1). In the 500-800 ms interval a main effect of *Size* emerges in the midline site analysis (F(2,30)=9.23, $\varepsilon_{GG}$=0.82, p<0.01), and in the lateral sites analysis (F(2.30)=8.21, $\varepsilon_{GG}$=0.85, p<0.01). The post-hoc t-tests FDR corrected for all sites (see table 5.5) show that the effect is mainly due to a difference between *Large* problems and the other two levels (*Medium* and *Small*) widespread in all the scalp. The *Large-Small* comparison reaches significance at AL, FC, CL, CP, CR, PL, PO sites. The *Large-Medium* comparison reaches significance at AL, AR, CP, CR, PO, PR sites. The *Medium-Small* comparison reaches significance only at the CL site. In the last part of the interval between the presentation of the second operand and the response cue (800-1400 ms) an effect of *Size* emerges only in the lateral sites analysis, F(2,30)=7.78, $\varepsilon_{GG}$=0.85, p<0.01. However, despite the effects appears more larger at frontal and right sites, no interaction with *Lateralization* or *Longitude* reaches significance. Moreover, post-hoc t-tests with FDR correction (see table

5.6) were performed as an explorative analysis. This t-tests suggest that, like in the 500-800 ms interval, the effect is mainly driven by the comparison of *Large* problems with the *Small* ones. Even if a difference between S*mall* and *Medium* problems reaches significance at the AL site, the absence in the ANOVA of an interaction of the *Size* factor with topographical factors does not allow to infer strong implications.

| comparison | ch | diff | t | df | p.value | p.adj | sign |
|---|---|---|---|---|---|---|---|
| large-small | AL | 1,34 | 3,06 | 15 | 0,008 | 0,037 | * |
| large-small | AR | 1,07 | 2,24 | 15 | 0,041 | 0,069 | . |
| large-small | CL | 0,99 | 2,58 | 15 | 0,021 | 0,048 | * |
| large-small | CP | 1,51 | 3,03 | 15 | 0,008 | 0,037 | * |
| large-small | CR | 0,95 | 2,92 | 15 | 0,010 | 0,037 | * |
| large-small | FC | 1,32 | 2,55 | 15 | 0,022 | 0,048 | * |
| large-small | FP | 1,25 | 2,23 | 15 | 0,042 | 0,069 | . |
| large-small | PL | 0,81 | 2,90 | 15 | 0,011 | 0,037 | * |
| large-small | PO | 1,11 | 3,88 | 15 | 0,001 | 0,015 | * |
| large-small | PR | 0,67 | 2,22 | 15 | 0,042 | 0,069 | . |
| large-medium | AL | 0,82 | 2,72 | 15 | 0,016 | 0,043 | * |
| large-medium | AR | 0,89 | 2,97 | 15 | 0,010 | 0,037 | * |
| large-medium | CL | 0,40 | 1,46 | 15 | 0,165 | 0,206 | |
| large-medium | CP | 1,07 | 2,76 | 15 | 0,015 | 0,043 | * |
| large-medium | CR | 0,70 | 4,20 | 15 | 0,001 | 0,012 | * |
| large-medium | FC | 0,88 | 2,20 | 15 | 0,044 | 0,069 | . |
| large-medium | FP | 0,92 | 2,47 | 15 | 0,026 | 0,052 | . |
| large-medium | PL | 0,47 | 1,61 | 15 | 0,129 | 0,181 | |
| large-medium | PO | 0,84 | 4,21 | 15 | 0,001 | 0,012 | * |
| large-medium | PR | 0,59 | 3,25 | 15 | 0,005 | 0,037 | * |
| medium-small | AL | 0,53 | 1,73 | 15 | 0,105 | 0,157 | |
| medium-small | AR | 0,18 | 0,39 | 15 | 0,706 | 0,730 | |
| medium-small | CL | 0,60 | 2,54 | 15 | 0,022 | 0,048 | * |
| medium-small | CP | 0,44 | 1,55 | 15 | 0,142 | 0,185 | |
| medium-small | CR | 0,25 | 0,75 | 15 | 0,468 | 0,520 | |
| medium-small | FC | 0,44 | 1,14 | 15 | 0,270 | 0,319 | |
| medium-small | FP | 0,33 | 0,69 | 15 | 0,503 | 0,539 | |
| medium-small | PL | 0,34 | 1,59 | 15 | 0,132 | 0,181 | |
| medium-small | PO | 0,27 | 1,13 | 15 | 0,276 | 0,319 | |
| medium-small | PR | 0,09 | 0,33 | 15 | 0,743 | 0,743 | |

**Table 5.5: post-hoc t-tests comparison FDR corrected for all sites and condition. The column "comparison" reports the conditions between the t-test was performed; the column "ch" reports the channels; "diff" is the difference between the conditions in uV; "t" is the t-value; "df" is the degrees of freedom; "p.value" is the uncorrected p-value; "p.adj" is the corrected p-value; "sign" report if the t-test was significant (*) or marginally significant (.). [multiplication: 500-800 ms interval]**

| comp | ch | diff | t | df | p | p.adj | sign |
|---|---|---|---|---|---|---|---|
| large-small | AL | 1,55 | 3,77 | 15 | 0,002 | 0,019 | * |
| large-small | AR | 1,45 | 2,80 | 15 | 0,013 | 0,051 | . |
| large-small | CL | 0,99 | 2,56 | 15 | 0,022 | 0,073 | . |
| large-small | CP | 0,75 | 1,01 | 15 | 0,331 | 0,499 | |
| large-small | CR | 1,31 | 4,10 | 15 | 0,001 | 0,014 | * |
| large-small | FC | 1,38 | 2,45 | 15 | 0,027 | 0,081 | . |
| large-small | FP | 1,71 | 3,27 | 15 | 0,005 | 0,039 | * |
| large-small | PL | 0,50 | 1,00 | 15 | 0,333 | 0,499 | |
| large-small | PO | -0,05 | -0,07 | 15 | 0,946 | 0,946 | |
| large-small | PR | 0,34 | 0,69 | 15 | 0,504 | 0,676 | |
| large-medium | AL | 0,57 | 1,83 | 15 | 0,088 | 0,219 | |
| large-medium | AR | 0,79 | 2,87 | 15 | 0,012 | 0,051 | . |
| large-medium | CL | 0,17 | 0,58 | 15 | 0,568 | 0,687 | |
| large-medium | CP | 0,36 | 0,66 | 15 | 0,518 | 0,676 | |
| large-medium | CR | 0,71 | 4,17 | 15 | 0,001 | 0,014 | * |
| large-medium | FC | 0,66 | 1,47 | 15 | 0,163 | 0,305 | |
| large-medium | FP | 0,67 | 1,64 | 15 | 0,122 | 0,244 | |
| large-medium | PL | 0,07 | 0,19 | 15 | 0,852 | 0,913 | |
| large-medium | PO | -0,09 | -0,22 | 15 | 0,829 | 0,913 | |
| large-medium | PR | 0,13 | 0,44 | 15 | 0,665 | 0,767 | |
| medium-small | AL | 0,98 | 3,07 | 15 | 0,008 | 0,046 | * |
| medium-small | AR | 0,66 | 1,39 | 15 | 0,186 | 0,312 | |
| medium-small | CL | 0,82 | 2,84 | 15 | 0,013 | 0,051 | . |
| medium-small | CP | 0,39 | 0,90 | 15 | 0,384 | 0,548 | |
| medium-small | CR | 0,60 | 1,66 | 15 | 0,118 | 0,244 | |
| medium-small | FC | 0,73 | 1,78 | 15 | 0,096 | 0,221 | |
| medium-small | FP | 1,03 | 2,21 | 15 | 0,043 | 0,117 | |
| medium-small | PL | 0,43 | 1,38 | 15 | 0,187 | 0,312 | |
| medium-small | PO | 0,04 | 0,11 | 15 | 0,912 | 0,944 | |
| medium-small | PR | 0,21 | 0,58 | 15 | 0,573 | 0,687 | |

**Table 5.6: post-hoc t-tests comparison FDR corrected for all sites and condition. The column "comparison" reports the conditions between the t-test was performed; the column "ch" reports the channels; "diff" is the difference between the conditions in uV; "t" is the t-value; "df" is the degrees of freedom; "p.value" is the uncorrected p-value; "p.adj" is the corrected p-value; "sign" report if the t-test was significant (*) or marginally significant (.). [multiplication: 800-1400 ms interval]**

To summarize, the most clear effect of *Size* for multiplications can be described in terms of a broad distributed positivity in the whole interval. In the early time window (300-500 ms), this broad distributed positivity is only marginally significant. From 500 ms after the presentation of the second operand to the presentation of the response cue, the broad distributed positivity reaches significance for *Large* problems compared to *Medium* and *Small* ones, and it has a peak in the 500-800 ms interval.

### *Size effect for addition*

Figure (figure 5.3) shows the grandaverages for the three levels (*Small*, *Medium*, and *Large*) of the *Size* factor for addition. The Ties are plotted separately, independently from size. The pattern appears more complicated than for the size effects in multiplication described above. *Ties* again show an early positive deflection with respect to all other conditions. Similarly to multiplications the positivity seems to start at around 300 ms post-stimulus. However, differently from multiplications, in additions the positivity starts with a anterior-right focus (see sites FP and AR compared to AL) and only subsequently spreads to the whole scalp. The comparison between the three levels of *Size* factor seems to show a biphasic pattern. There is an early anterior negativity for *Large* problems with respect to *Medium* and *Small*, peaking at around 400ms. However, in the following slow waves a more positive trend at centro-parietal right sites appears for *Large* and *Medium* problems with respect to the *Small* in the 500-700 ms time window, and for *Large* with respect to *Medium* and *Small* in the 900-1200 ms time window. Overall, in line with what found for multiplications, the pattern for addition can be resumed as a larger posterior positivity for larger problems.

Similarly to what we have done for multiplications, we performed the analysis in two stages. First, we analyzed the difference between *Ties* and the other *Size* levels (*Size+Ties*: 4 level factors) in an adequate time window (i.e., 300-500 ms). Second, we analyzed the effect of *Size* only in the time windows where the differences above described are evident in the grandaverage plot: the interval 400-500 ms for the frontal negativity; the intervals 700-900 ms and 900-1200 ms for the following positivity.

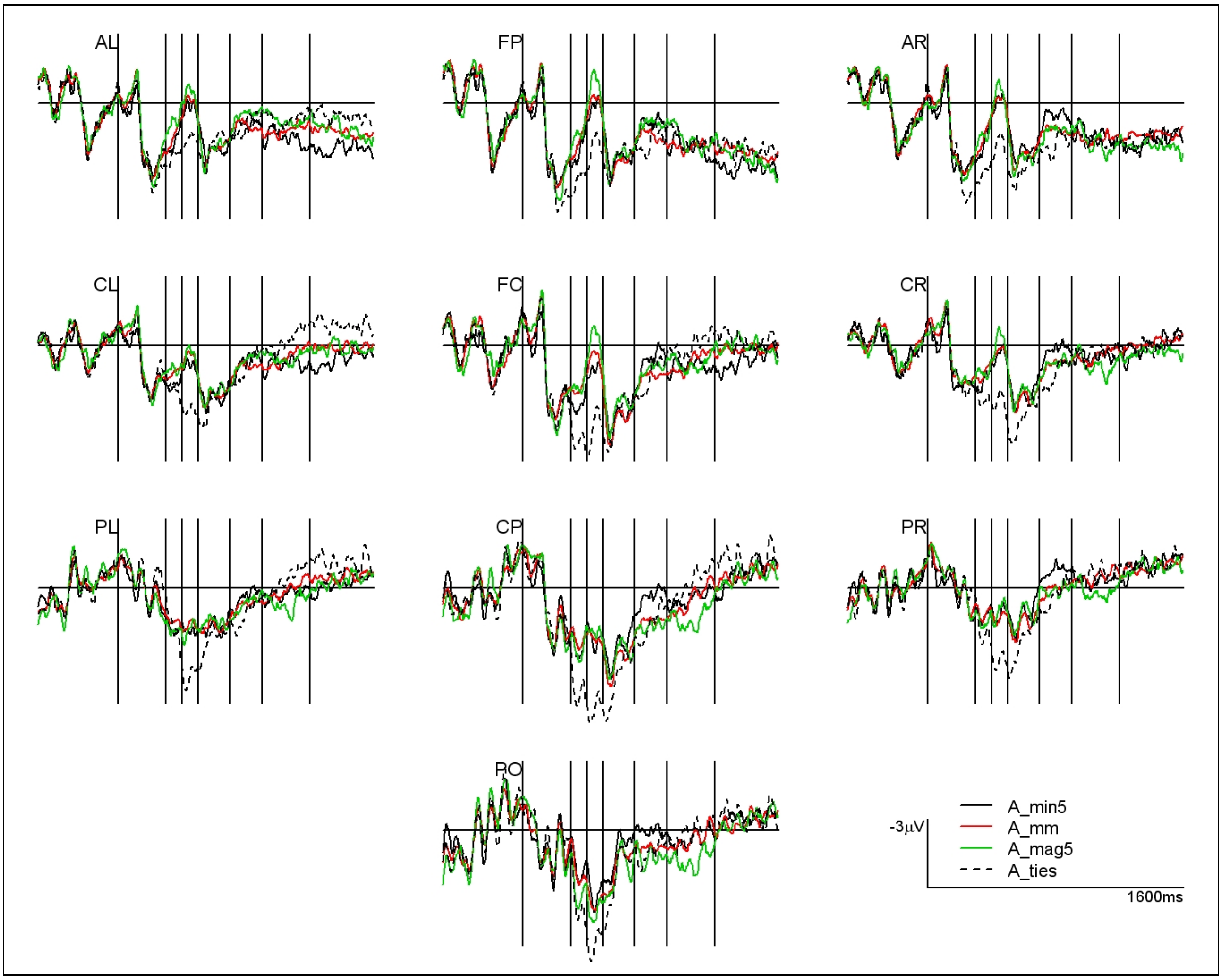


**Figure 5.3: Grandaverage plot for the size and ties comparison for additions: *Ties* problems(A_ties) are plotted in black dashed, *Small* problems (A_min5) in black, *Medium* problems (A_mm) in red and *Large* problems (A_mag5) in green. Vertical lines are plotted at the onset of the second operand, and after 300,400,500,700,900,1200ms.**

The ANOVA in the 300-500 ms for midline sites showed a main effect of four levels *Size+Ties* factor ($F(3,45)=10.12$, $\varepsilon_{GG}=0.55$, $p<0.01$), and an interaction of the same factor with *Longitude* ($F(9,135)=3.42$, $\varepsilon_{GG}=0.30$, $p<0.05$). Similarly, the ANOVA in the same time window on lateral sites shows a main effect of size ($F(9,135)=10.23$, $\varepsilon_{GG}=0.70$, $p<0.01$), but no interaction with the two topographical factors. Post-hoc t-tests with FDR correction were performed for each site of both midline and central line. The t-tests compared *Ties* with the three levels of the *Size* factor. Similarly to multiplications, the differences were significant for most sites and comparisons (see table 5.8).

| comp | ch | diff | t | df | p | p.adj | sign |
|---|---|---|---|---|---|---|---|
| ties-small | AL | 0,99 | 1,59 | 15 | 0,133 | 0,138 | |
| ties-small | AR | 1,59 | 3,10 | 15 | 0,007 | 0,014 | * |
| ties-small | CL | 1,00 | 1,77 | 15 | 0,098 | 0,105 | |
| ties-small | CP | 2,48 | 3,75 | 15 | 0,002 | 0,008 | ** |
| ties-small | CR | 1,42 | 3,61 | 15 | 0,003 | 0,008 | ** |
| ties-small | FC | 2,01 | 2,98 | 15 | 0,009 | 0,016 | * |
| ties-small | FP | 1,49 | 2,14 | 15 | 0,049 | 0,061 | . |
| ties-small | PL | 0,95 | 1,85 | 15 | 0,085 | 0,094 | . |
| ties-small | PO | 1,93 | 2,67 | 15 | 0,017 | 0,027 | * |
| ties-small | PR | 1,37 | 3,57 | 15 | 0,003 | 0,008 | ** |
| ties-medium | AL | 1,22 | 2,47 | 15 | 0,026 | 0,036 | * |
| ties-medium | AR | 1,73 | 4,56 | 15 | 0,000 | 0,003 | ** |
| ties-medium | CL | 1,44 | 3,20 | 15 | 0,006 | 0,012 | * |
| ties-medium | CP | 2,62 | 3,46 | 15 | 0,003 | 0,009 | ** |
| ties-medium | CR | 1,65 | 4,73 | 15 | 0,000 | 0,003 | ** |
| ties-medium | FC | 2,63 | 3,85 | 15 | 0,002 | 0,008 | ** |
| ties-medium | FP | 1,80 | 2,87 | 15 | 0,012 | 0,019 | * |
| ties-medium | PL | 1,19 | 2,27 | 15 | 0,038 | 0,050 | * |
| ties-medium | PO | 1,60 | 2,10 | 15 | 0,053 | 0,064 | . |
| ties-medium | PR | 1,33 | 2,54 | 15 | 0,023 | 0,032 | * |
| ties-large | AL | 1,69 | 3,34 | 15 | 0,004 | 0,010 | * |
| ties-large | AR | 1,89 | 4,83 | 15 | 0,000 | 0,003 | ** |
| ties-large | CL | 1,55 | 3,22 | 15 | 0,006 | 0,012 | * |
| ties-large | CP | 2,57 | 3,49 | 15 | 0,003 | 0,009 | ** |
| ties-large | CR | 1,87 | 4,98 | 15 | 0,000 | 0,003 | ** |
| ties-large | FC | 3,09 | 4,40 | 15 | 0,001 | 0,003 | ** |
| ties-large | FP | 2,32 | 3,67 | 15 | 0,002 | 0,008 | ** |
| ties-large | PL | 0,92 | 1,97 | 15 | 0,068 | 0,078 | . |
| ties-large | PO | 1,01 | 1,42 | 15 | 0,176 | 0,176 | |
| ties-large | PR | 1,29 | 2,66 | 15 | 0,018 | 0,027 | * |

**Table 5.8: post-hoc t-tests comparison FDR corrected for all sites and condition. The column “comparison” reports the conditions between the t-test was performed; the column “ch” reports the channels; “diff” is the difference between the conditions in uV; “t” is the t-value; “df” is the degrees of freedom; “p.value” is the uncorrected p-value; “p.adj” is the corrected p-value; “sign” report if the t-test was significant (*) or marginally significant (.). [addition: 300-500 ms interval]**

The only the exceptions are: the difference between *Ties* and *Large* problems in the PL and PO sites; the difference between *Ties* and *Medium* problems in the PO site; and the difference between *Ties* and *Small* problems in the PO, AL, CL, FP, and PL site. For all other comparisons the difference between *Ties* and the other three conditions was between 1.13uV and 3.10uV.

To test the apparent frontal negativity for *Large* problems in the 400-500 ms time windows, we performed an ANOVA restricted to *Size* on both midline and lateral sites. However, this ANOVA did not show any significant effect. The ANOVA in the 700-900 ms time window showed an interaction between *Laterality* and *Size* for lateral sites ($F(2,30)=5.45$, $\varepsilon_{GG}=0.79$, $p<0.05$), but no effect involving *Size* in the midline sites analysis. Despite no one of the post-hoc t-test on single sites reaches significance after FDR correction, the interaction can be easily interpreted on the basis of grandaverage plot (see figure 5.3). The grandaverage plot shows that *Medium* and *Large* problems are more positive than *Small* at lateral right sites (AR, CR, PR). No effects for ANOVAs in the 900-1200 ms time windows was found.

### ***Size and Order effects in multiplication***

In figure 5.4, 5.5, and 5.6 the grandaverages for the two operands orders (s×L in black and L×s in red) are reported in separate plots for the three *Size* levels (*Small*, *Medium*, *Large*). The effects, as predictable on the basis of the interactions found in the behavioural experiments of Chapter 2 and 4, are rather different. First of all, we must notice that the effects are rather large in amplitude with respect to the effects of problem size previously described. Moreover, this effects start very early (at around 300ms), similarly to the effect of *Ties* that we attributed to the perception of the problem (repetition priming on the second operand).

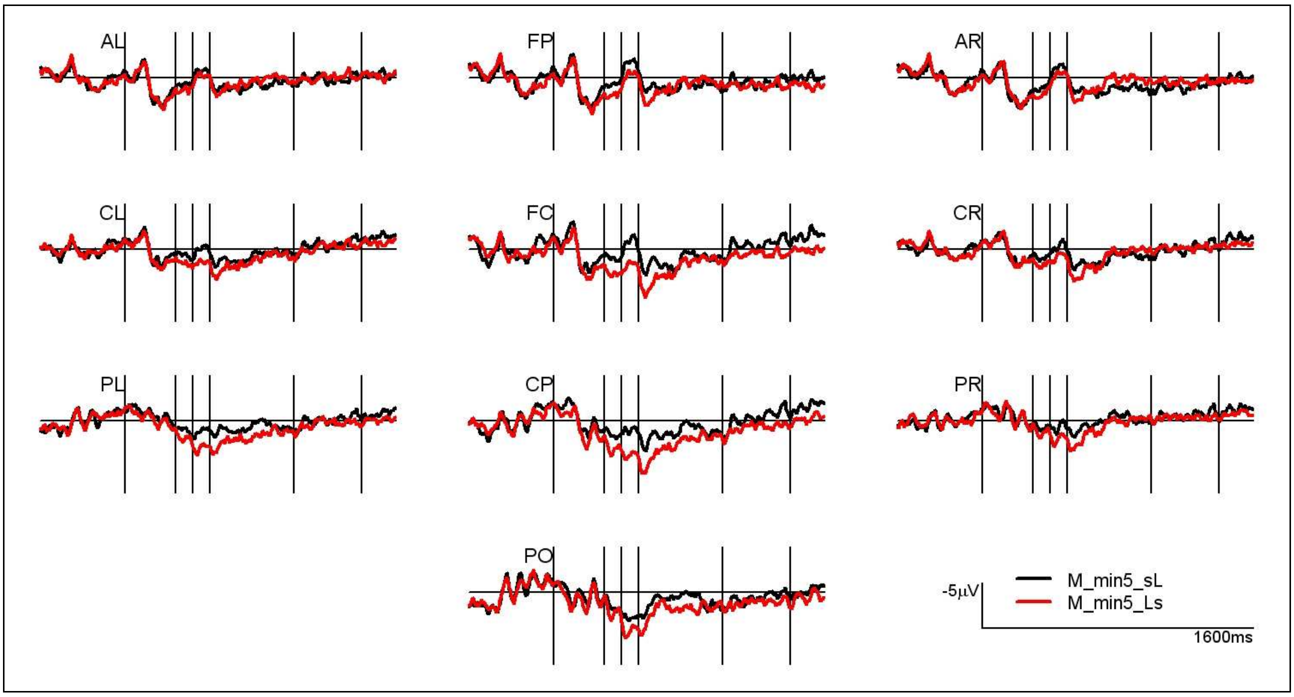


**Table 5.4: Grandaverage plot for the two orders for *Small* problems: sxL (M_min5_sL) waveforms are plotted in black and Lxs (M_min5_Ls) in red. Vertical lines are plotted at the onset of the second operand, and after 300,400,500,1000,1400ms**

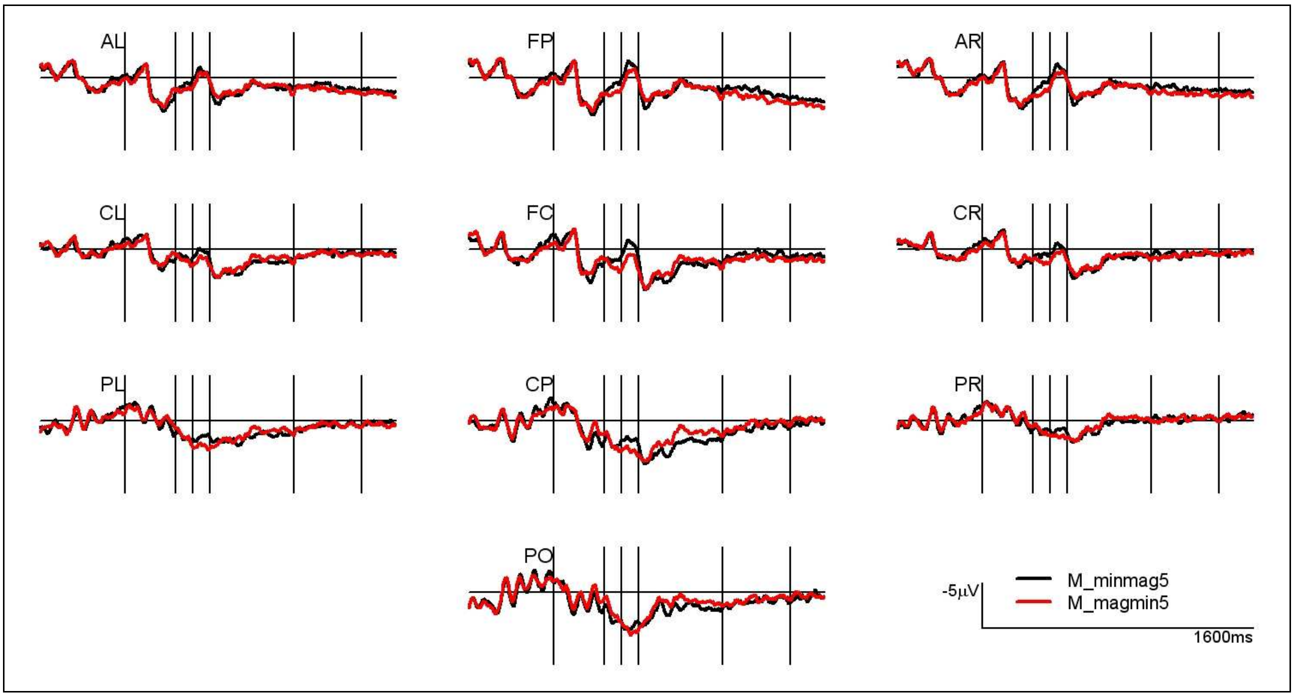


**Figure 5.5: Grandaverage plot for the two orders for *Medium* problems: sxL (M_minmag5) waveforms are plotted in black and Lxs (M_magmin5) in red. Vertical lines are plotted at the onset of the second operand, and after 300,400,500,1000,1400ms.**

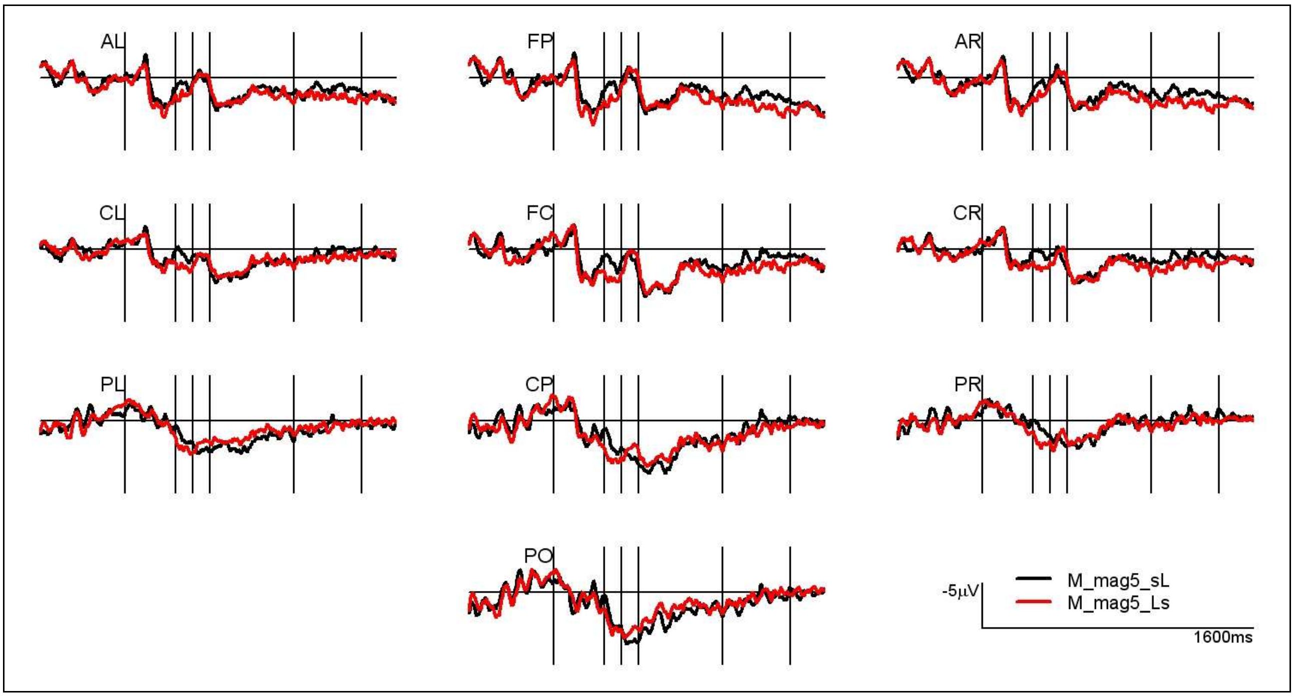


**Table 5.6: Grandaverage plot for the two orders for *Large* problems: sxL (M_mag5_sL) waveforms are plotted in black and Lxs (M_mag5_Ls) in red. Vertical lines are plotted at the onset of the second operand, and after 300,400,500,1000,1400ms.**

We first describe the effects and then report the statistical analysis performed on appropriate time windows. For the *Small* problems the effect appears as a broad distributed negativity, more evident at central sites, for the s×L order (non-privileged order) with respect to the L×s[11]. The effect starts at about 300 ms post-stimulus and last until about 800 ms. Both waveforms in this interval show a negative peak more frontally distributed at around 450 ms, followed by a positive peak at parietal sites at around 550 ms. For *Medium* problems the negativity for s×L order (non-privileged order) with respect to L×s order is restricted to the peak at 450ms in the interval 300-500 ms, with no clear effect in the earlier time window. However, the effect is inverted in the later interval (from about 600 to 1000 ms), with L×s

[11] Here and in the following analysis we prefer to interpret order effects in terms of negativity. In fact, Zhou et al. (2007) reported a widely distributed negativity more evident on frontal sites as the putative ERPs correlate of a operands reordering operation for the unprivileged order. In the ERPs study of Zhou and Colleagues, Chinese participants had to solve multiplication problems. The Chinese population learn only one half of the multiplication table (i.e., only the s×L order is learned).

order more negative than s×L order but rather focal in topography since it is evident only on the CP site. For *Large* problems a negativity for s×L order (privileged order) with respect to L×s , similar to *Small* problems, is widespread distributed between 300 ms and 400 ms. However, the opposite pattern abruptly develops between 400 ms and 500 ms, with the L×s order (non-privileged order) more negative than the s×L order (mainly at parietal sites CP, PO). In the 600-1000 ms time window, a negative deflection for L×s order seems to persist until around 800 ms only at posterior left sites, whereas in the last period (between 1000 ms and 1400 ms) an anterior positivity for L×s order is evident.

Given the complexity of the design and of the effects (both topographically and in terms of the experimental design), and the fact that the data comes from a limited number of trials and are thus more noisy than the previous comparisons, our statistical analysis needs to be firmly grounded on a priori hypotheses. One hypothesis comes from Zhou and Colleagues (2007) who reported a widely distributed negativity for non-privileged order with Chinese participants. The other comes from the behavioural experiments (see chapters 2 and 4), where we found that in the Italian adult population the L×s order is privileged with *Small* and *Medium* problems, whereas the s×L order is privileged with *Large*. From the qualitative analysis reported above at least four time windows of interest may be selected. The first time window is between 300 ms and 400 ms. In this time window, since the s×L order shows a negativity with respect to L×s order in both *Small* and *Large* problems but not in the *Medium* ones, this effect cannot be interpreted in terms of a reordering process toward the privileged order. The second time window is between 400 ms and 500 ms. In this time window the *Small* and *Medium* problems show a larger negativity for s×L order (non-privileged) than for L×s order, and *Large* problems shows a larger negativity for L×s order (non-privileged) than for s×L order. This effect is in line with a possible interpretation in terms of a reordering process because of the non-privileged orders show a negativity for all three levels of the *Size* factor. The effects following 500 ms are rather confusing and it is not clear which time windows to choose in order to compare different deflections in the different conditions in an overall ANOVAs with both *Size* and *Order* as factors. It is likely that after 500 ms for some problems and/or participants the result of the problem is obtained soon, whereas for others a memory search or an alternative procedures are implemented by the cognitive system. Moreover, different effects such as the early stage of the CNV, slow posterior positive waves, and negativities due to inversion or other explicit symbolic transformations of the problem may superimpose. This is not unusual in ERPs research, especially for slow wave and for data with a rather low SNR (signal to noise ratio). In fact, in the "*Psychophysiology Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria*" under the subsection "*Mean Amplitude Measurements Over a Period of Time Should Not Span Clearly Different ERP Components*", Picton and Colleagues underline that "*when measuring slow or sustained potentials the latency range can span several hundred milliseconds. However, if the scalp distribution of the ERP changes significantly during the measurement period, the resultant measurements may become impossible to interpret*" (Picton et al., 2000, p.143). Despite the grandaverages with *Size* only as factor were interpretable (see above), the patterns after 500 ms with both *Size* and *Order* above described are very difficult to be attributed to a single latent component or effect. This patterns do not match with any hypotheses developed in the introduction and thus have not been analyzed. In fact, the ERP differences due to order in the first stages after 500 ms at posterior sites is similar in terms of polarity for *Large* and *Medium* (s×L more positive that L×s) problems, and opposite for the *Small* ones (L×s more positive that s×L); whereas in the previous behavioural experiments *Small* and *Medium* always showed a similar pattern with respect to order of operands.

The ANOVA on midline sites in the 300-400 ms reveals a main effects of both *Size* ($F(2,30)=3.94$, $\varepsilon_{GG}=0.78$, $p<0.05$) and *Order* ($F(1,15)=14.95$, $p<0.01$). Moreover, the ANOVA reveals an *Order* by *Longitude* interaction, $F(3,45)=4.99$, $\varepsilon_{GG}=0.52$, $p<0.05$. The ANOVA on lateral sites for the same interval shows an effect of *Order* ($F(1,15)=17.53$, $p<0.01$) and a marginal significant interaction between *Longitude* and *Size* ($F(4,60)=2.98$, $\varepsilon_{GG}=0.51$, $p<0.1$).

The *Size* effect is due to the positivity for *Large* problems with respect to *Small* and *Medium* problems described above, that emerged just as a tendency in the larger 300-500 ms time window. The interaction of *Order* and *Longitude* in the midline sites and the main effect of *Order* on lateral sites is due to the broad distributed large positivity for L×s problems with respect to s×L, which is more evident on the posterior areas of the scalp. The effect is numerically larger for *Small* and *Large* problems with respect to *Medium* problems, but the fact that there is no interaction involving *Order* and *Size* does not allow us to speculate on this point. Despite possible differences in the amplitude of the effect, all three levels of size show the same pattern that is not likely to be the frontal negativity due to inversion of operand discussed by Zhou and Colleagues (2007). Since does not mirror behavioural preferences in the order of operands and it is mainly posteriorly distributed. For these reasons the effect is more likely to be attributed to the processing of the second operand, that may differ as a function of the fact that the first operand is smaller or larger.

The ANOVA in the 400-500 ms interval on midline sites shows an effect of *Order*, ($F(1,15)=15.45$, $p<0.01$), a *Longitude* by *Order* interactions ($F(3,45)=4.36$, $\varepsilon_{GG}=0.55$, $p<0.05$), and a *Size* by *Order* interaction ($F(2,30)=5.48$, $\varepsilon_{GG}=0.99$, $p<0.01$). The ANOVA in the same latency interval on lateral sites gives an effect of *Order* ($F(1,15)=8.83$, $p<0.01$), and a four-way *Lateralization* by *Longitude* by *Size* by *Order* interaction ($F(4,60)=4.52$, $\varepsilon_{GG}=0.72$, $p<0.01$). In figure 5.6 the means in the different conditions for midline sites are reported showing that the *Order* effects are larger posteriorly, and the *Size* by *Order* interaction is due to the fact that for *Small* and *Medium* problems s×L are more negative that L×s, whereas for *Large* problem the opposite pattern emerge.

It is worthwhile to note that despite amplitudes in the L×s order is very similar for all conditions this does not allow to assume that the effect for *Large* problems has to be interpreted as a positivity for s×L conditions, since already in the previous time windows *Large* problems were globally more positive than *Medium* and *Small* and thus it is possible a superimposition of a long lasting order-independent positivity for *Large* problem and a negativity for the non-privileged order. Similar pattern emerges for the lateral sites (see figure 5.7 and 5.8). Despite the interpretation of a four-way interaction is always complex, it is likely to be due to the fact that both *Order* effect and the inversion of the effect for *Large* problems is more evident over the left hemisphere.

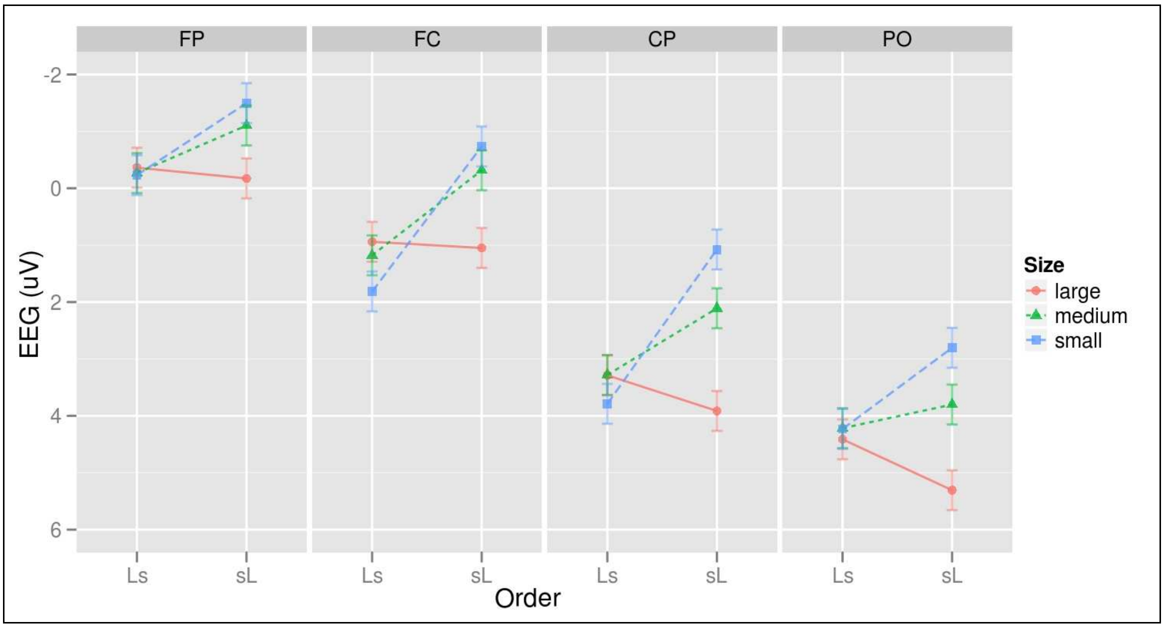


**Figure 5.6: mean voltage amplitudes in the different conditions for midline sites.**

Summarizing, the analysis aimed to study operands order effects for multiplication provides easy to interpret effects in the early development of the problem-cue interval, whereas later effects appear to be due to the interplay of different components that make the data very difficult to interpret. In the first interval we analyzed (300-400 ms) *Size* and *Order* do not interact and a large posteriorly distributed negativity for s×L order with respect to L×s order emerges, independently from *Size.* A broad distributed effect of *Size* is as well present with *Large* problems waveforms being more positive than *Medium* and *Small*. In the following time window (400-500 ms) a similar pattern persists for *Medium* and *Small* problems, for which the more positive potential is for the preferred order, whereas an inversion of the effect, especially on posterior left sites, emerges for *Large* problems.

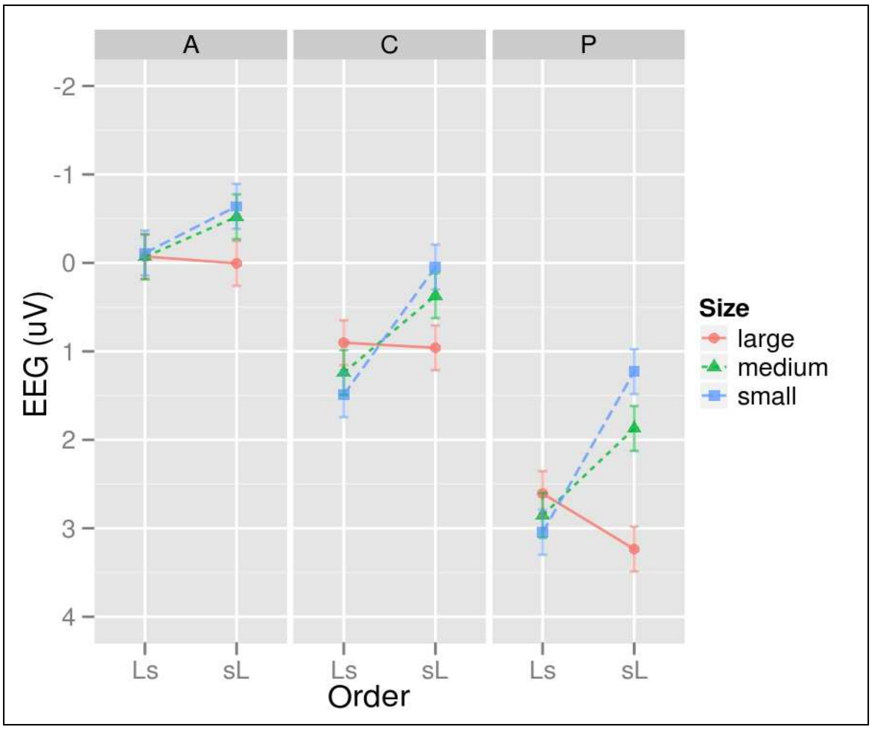


**Figure 5.7: mean voltage amplitudes in the different conditions for left sites.**

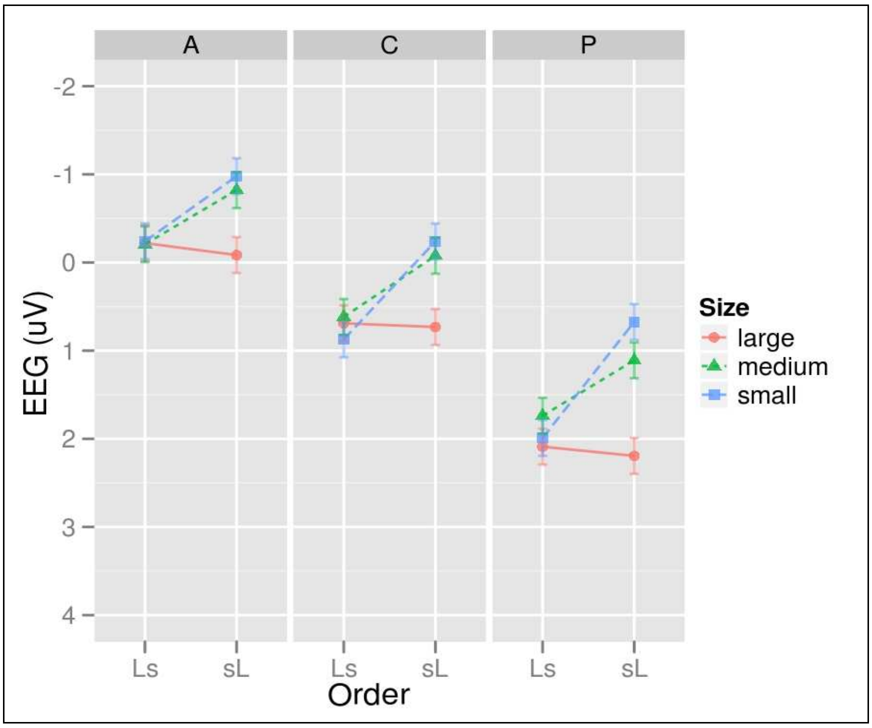


**Figure 5.8: mean voltage amplitudes in the different conditions for right sites.**

***Size and Order effects in addition***

In figure 5.9, 5.10, and 5.11 the grandaverages for the two operand orders (s+L in black and L+s in red) are reported in separate plots for the three *Size* levels (*Small, Medium, Large*). Both *Medium* and *Large* problems show a larger negativity frontally distributed after 300 ms for the non-privileged order (s+L), short lasting for *Large* and longer for *Medium*. On the other side *Small* problems show just a small negativity for L+s in an earlier time window. The two effects seem clearly different in latency and topography. Moreover, we know that additions with both operand equal/smaller than 5 are very easy and in the behavioural experiment reported in Chapter 2 no effect of order emerged. For these reasons we will analyze Order and Size effects only for *Medium* and *Large* problems, given no unique time interval can be chosen for all the three level of size encompassing similar effects[12]. The ANOVAs were performed in the 300-500 ms time interval, containing the whole time course of the anterior negativity, with two factors: *Order* and the *reduced-Size* (two levels: *Medium* and *Large*), hereafter *Size* for this paragraph.

The ANOVA in the 300-500 ms interval for midline sites shows a main effect of *Order*, ($F(1,15)=8.39$, $p<0.01$), an interaction *Longitude* by *Order* ($F(3,45)=4.00$, $\varepsilon_{GG}=0.49$, $p<0.05$), and by *Order* interactions ($F(3,45)=3.05$, $\varepsilon_{GG}=0.51$, $p<0.1$). Similarly for the lateral sites analysis we obtained an effect of Order ($F(1,15)=14.8$, $p<0.01$), an *Laterality* by *Longitude* by *Size* interaction ($F(2,30)=5.36$, $\varepsilon_{GG}=0.64$, $p<0.05$), and again just a tendency for the *Longitude* by *Size* by *Order* interaction ($F(2,30)=3.76$, $\varepsilon_{GG}=0.56$, $p<0.1$). Both analyses confirm the effect of order and its anterior distribution. Despite the effect for *Medium* is numerically larger and seem to last longer interaction with both *Order* and *Size* factor are

[12] At explorative level an analysis not reported here, restricted at the Small problem only, did not show any effect or interaction of *Order* in the interval 200-400 ms, thus the early negativity in the L+s waveform peaking at about 300 ms is likely to be just noise.

only marginally significant[13]. The ERPs correlate of the order of the operands preferences in addition can thus be described in terms of a frontal bilateral negativity for non-privileged order (s+L), evident only for problems with at least one operand larger than 5 (*Medium* and *Large*).

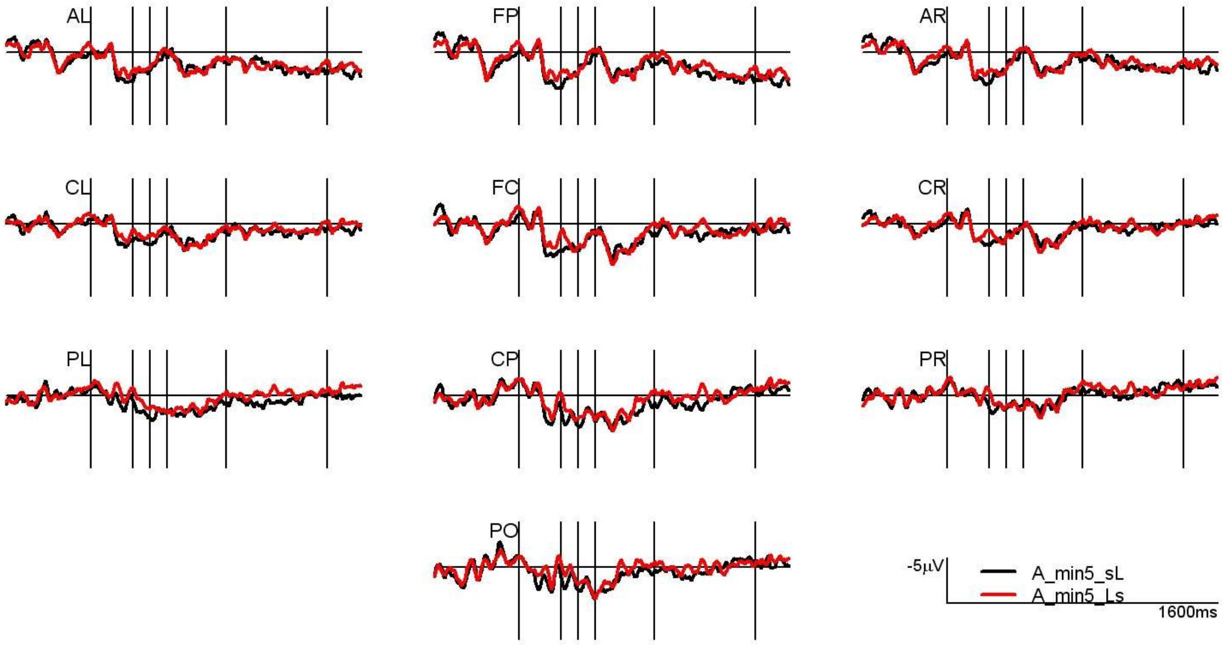


**Figure 5.9: Grandaverage plot for the two orders for *Small* problems: s+L (A_min5_sL) waveforms are plotted in black and L+s (A_min5_Ls) in red. Vertical lines are plotted at the onset of the second operand, and after 250,350,450,800,1400ms**

[13] In order to check for possible differences in the amplitude and time development we also tested separate time windows (300-400 ms and 400-500 ms) with other analyses not reported here, but we found no direct evidence in favour of a larger effect for *Medium* with respect to *Large* problems.

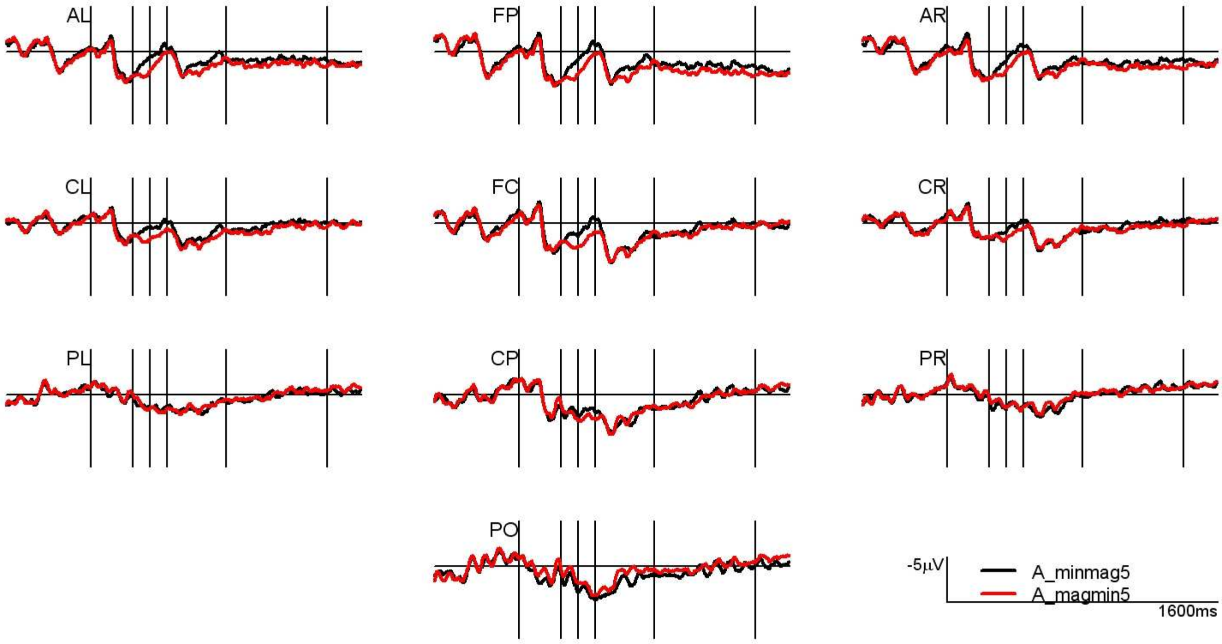


**Figure 5.10: Grandaverage plot for the two orders for *Medium* problems: s+L (A_minmag5_sL) waveforms are plotted in black and L+s (A_magmin5_Ls) in red. Vertical lines are plotted at the onset of the second operand, and after 250,350,450,800,1400ms**

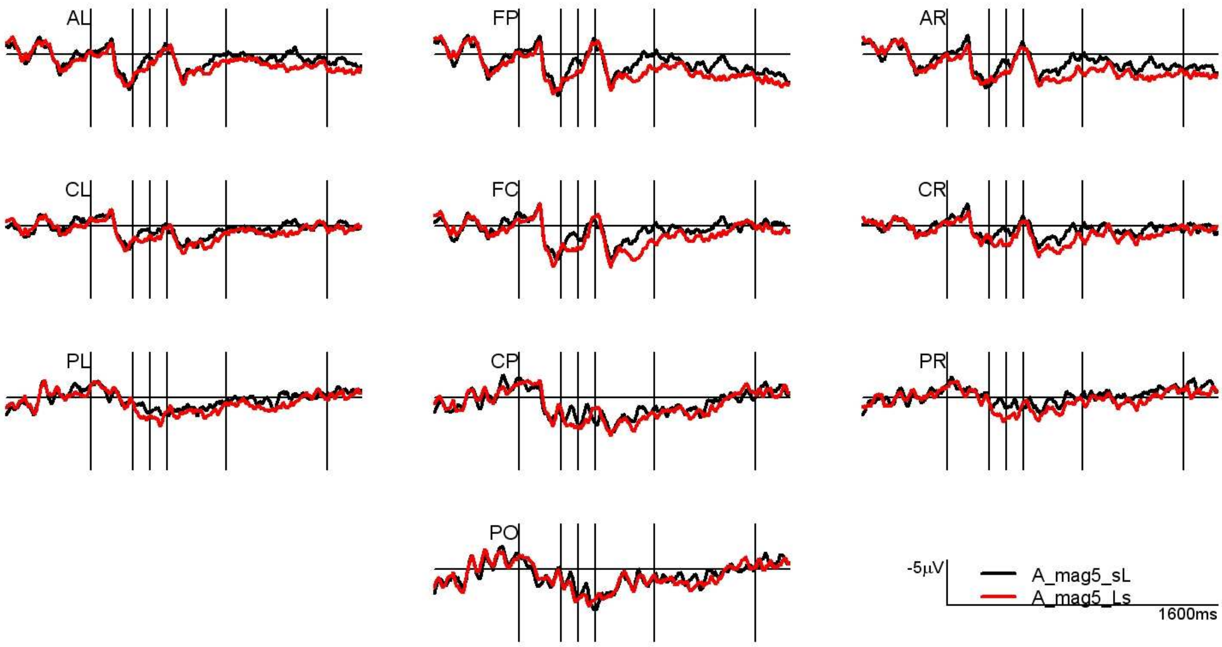


**Figure 5.11: Grandaverage plot for the two orders for *Large* problems: s+L (A_mag5_sL) waveforms are plotted in black and L+s (A_mag5_Ls) in red. Vertical lines are plotted at the onset of the second operand, and after 250,350,450,800,1400ms**

***Self-report based analysis***

We decided to consider also a different approach to analyze the data, by splitting the epochs not according to stimuli properties (the kind of problem) but based on the self report of participant about the procedures they reported to use. In particular, we decided to pool all the trials remaining, after artifact rejection procedure above described, where participants explicitly indicated they used the "*Inversion*" procedure (e.g., to solve the problem 3×7 by reordering it in the 7×3 order) to solve the problem. The "*Inversion*" trials have been compared with all the other trials. This approach, somewhat unusual in the ERPs research, but already efficiently used in an fMRI study on arithmetical cognition (Grabner et al., 2009), lead to a largely unbalanced design since size and difficulty of the problems are not balanced both in terms of the stimuli (the problems) and the other self-report information we collected during the structured debriefing. Moreover, since *Inversion* is not reported very frequently, we compared waveforms coming from a very different number of epochs and thus with different SNR (signal to noise ratio). However, ANOVAs are typically robust enough to deal with this kind of unbalance design (e.g., P3 studies typically compare waveforms elicited by frequent target with infrequent standards). The aim of this analysis is to compare the order effect we identified for additions and multiplications with the Zhou et al. (2007) findings in the Chinese population. In fact, in this analysis ERP signature can be attributed to an explicit and aware cognitive process of inversion of the operands. The frequency of reporting the use of inversion procedure to solve the problems was not homogeneous across participants. We selected for subsequent analysis only the participants for which more than 10 epochs for each type of problem (*Addition* and *Multiplication*) that the participants reported to solve with explicit inversion in the self-report. For multiplications all 16 participants fulfilled the criteria, the resulting averages were formed by an average number of 33.3 epochs (min=12, max=59) in the *Inversion* condition and by an average number of 124.3 epochs (min=68, max=164) in the *Non-Inversion* condition. For additions only 7 participants fulfilled the criteria, the resulting averages were formed by an average number of 36.6 epochs (min=11, max=65) in the *Inversion* condition and by an average number of 124.3 epochs (min=61, max=169) in the *Non-Inversion* condition. The data averaging, the statistical analysis, the topographical design, and the method were the same as in the previous analyses based on *Size* and *Order* classification. Before analyzing the ERPs it is interesting to see which problems forms the *Inversion* condition, that is which was the overall proportion of times for which the inversion procedure was used within the cells defined by the *Order* and *Size* factors used in the previous analyses (see table 5.9).

| order | size | epoch | count | prop | operation |
|---|---|---|---|---|---|
| s+L | medium | 158 | 336 | 0.47 | addition |
| L+s | medium | 6 | 336 | 0.02 | addition |
| s+L | large | 62 | 126 | 0.49 | addition |
| L+s | large | 7 | 126 | 0.06 | addition |
| s+L | small | 42 | 126 | 0.33 | addition |
| L+s | small | 0 | 126 | 0 | addition |
| Lxs | large | 92 | 288 | 0.32 | multiplication |
| sxL | large | 48 | 288 | 0.17 | multiplication |
| Lxs | medium | 56 | 768 | 0.07 | multiplication |
| sxL | medium | 291 | 768 | 0.38 | multiplication |
| sxL | small | 54 | 288 | 0.19 | multiplication |
| Lxs | small | 0 | 288 | 0 | multiplication |

**Table 5.9: the trials used to make the average in the self-report analysis. The column "order" reports the order of the problem; "size" the size of the problem; "epoch" the number of epochs used in the analysis in each *Order* and *Size* condition; "count" the total number of epoch in each *Order* and *Size* condition; "prop" the proportion of epoch used; "operation" the operations.**

Gradaverage plots for *Multiplication* and *Additions* are reported respectively in figure 5.12 and figure 5.13. The grandaverage of the condition for which participants reported to have used the procedure of *Inversion* is plotted in red and the *Non-Inversion* condition is plotted in black. For both *Multiplications* and *Additions* a large negativity is evident bilaterally at frontal site in the interval between 300 ms and 500 ms. The pattern is very similar for *Additions* and *Multiplications*, despite the differences in the processing of the two operations which ERPs correlate is mainly visible on posterior sites as already seen in the previous analysis. Given the participants considered for *Additions* and *Multiplications* are different, separate repeated measure analysis with *Inversion* only as factor have been reported.

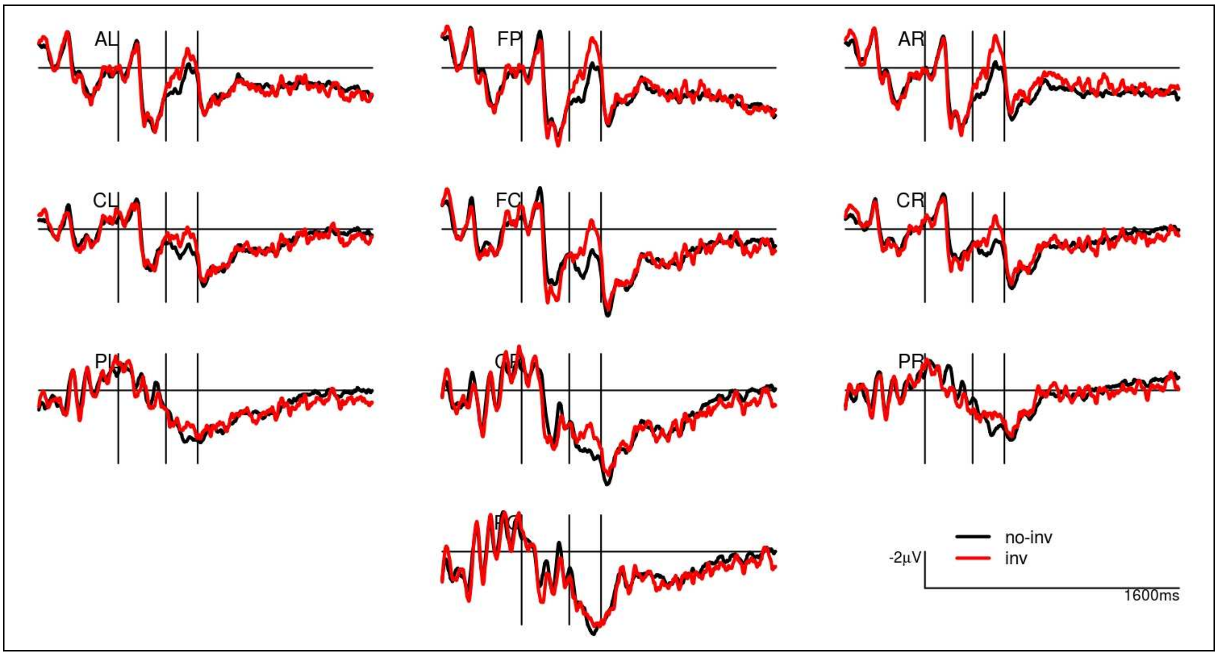


**Figure 5.12: Grandaverage plot for the two *Inversion* condition for multiplication: the waveforms of the trials where the participants reported to use inversion (inv) are plotted in red and the waveforms of the trials where the participants did not report to use inversion (no-inv) are plotted in black. Vertical lines are plotted at the onset of the second operand, and after 300 and 500ms.**

ANOVA on midline sites in the 300-500 ms for *Multiplications* shows an *Longitude* by *Inversion* interaction, $F(3,45)=5.87$, $\varepsilon_{GG}=0.68$, $p<0.01$. The analysis in the same time interval on lateral sites shows a main effect of *Inversion* ($F(1,15)=6.06$, $p<0.05$) and a marginal *Longitude* by *Inversion* interaction ($F(2,30)=4.11$, $\varepsilon_{GG}=0.59$, $p<0.1$). The analysis on the same time interval for *Additions* shows a marginal significant *Longitude* by *Inversion* interaction, $F(3,18)=3.87$, $\varepsilon_{GG}=0.48$, $p<0.1$ in the midline analysis, whereas no significant effect emerges from the analysis on lateral sites. Given the small number of participants in the *Addition* condition (7) and the fact that the effect that we attribute to the cognitive process of explicit reordering of the operands is very similar to that of *Multiplications*, we think the interpretation of the marginal significant effect is very likely to be confirmed with a larger participant pool. We also want to notice that the numerical amplitude of the effect (more than 1µV on the average on the whole time window for both *Additions* and *Multiplications*) overcame all the other effect previously described (different operations, size of the problem).

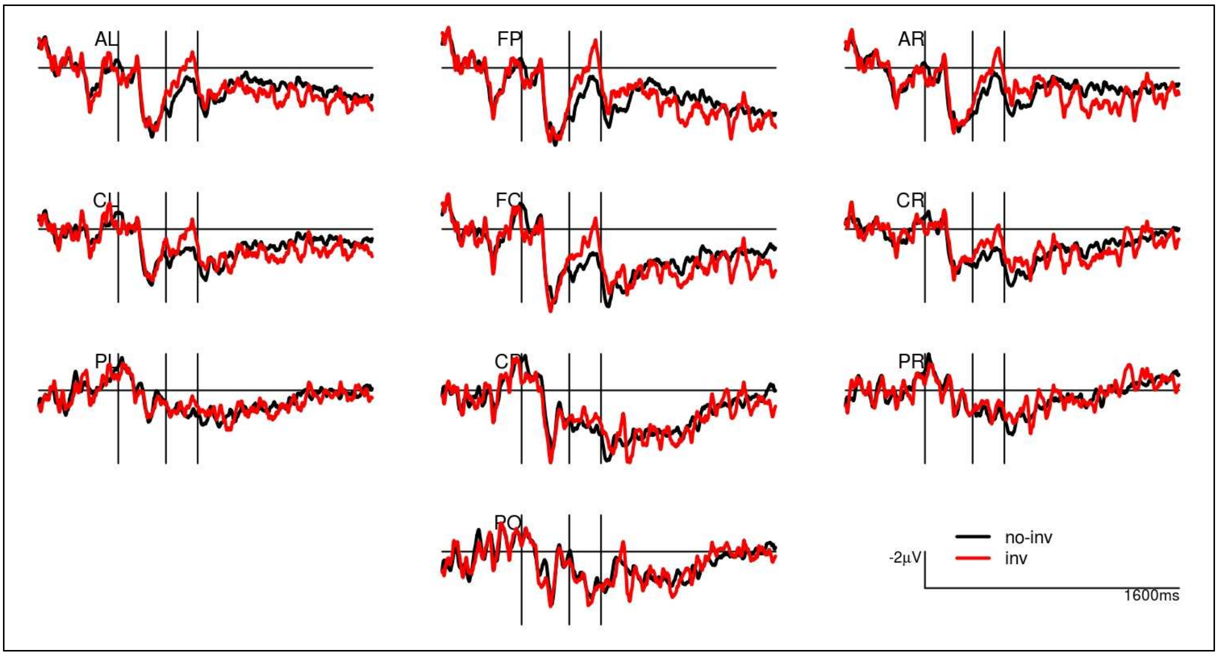


**Figure 5.13: Grandaverage plot for the two *Inversion* condition for addition: the waveforms of the trials where the participants reported to use inversion (inv) are plotted in red and the waveforms of the trials where the participants did not report to use inversion (no-inv) are plotted in black. Vertical lines are plotted at the onset of the second operand, and after 300 and 500ms.**

## 5.4 DISCUSSION

The behavioural analysis showed that in the verbal response condition the participants responded faster than in the manual condition for both operation. Moreover, in multiplication the size effect was significant, despite this can be due just to the fact that larger results are harder to represent and produce both verbally and manually, it is also possible that participants did not solve the problems before the presentation of the response cue. No order effect or interactions with size emerged for multiplications, while in additions a very small order effect emerged since participants responded faster for the L+s order than with the s+L order, again suggesting that participants could have waited the cue to complete the calculation.

The ERPs elicited by addition and multiplication problems in the interval between the problem presentation (i.e. second operand onset) and the response cue show a long lasting left lateralized positivity for multiplications with respect to additions mainly distributed on central-parietal sites with an onset at about from 350 ms after the problem presentation. The left lateralization can suggest a larger involvement of verbal memory in the retrieval of multiplication facts with respect to additions. Differently from the hypothesis we developed on the basis of the literature we did not find a frontal negativity for multiplications with respect to additions.

The following comparisons of problem size and ties showed a clear differences of the waveforms elicited by ties with respect to all other problems (independently from size) for both additions and multiplications. Even if the effect appear slightly different (numerically larger in case of multiplications and with some difference in the topography of the early stage of the effect) for both multiplications and additions, the effect can be described as a broad distributed positivity between 300 ms and 500 ms after the onset of the second operand of the problem. This effect is not easy to be interpreted but can be classified either as an extremely suppressed N400 or as a P3. In both cases it is difficult to decide whether this effect has to be attributed to the processing of the second operand per se or to the fact that ties are stored in a separate repository within the arithmetic facts memory as some models assumes (*network interference model*, see chapter 1). The results of the experiment reported in Chapter 4 suggest that the tie problems are stored separately. However, the fact that the same effect shows up for both operations favours a simpler interpretation of the effect in terms of an extremely suppressed N400, reflecting the easier processing of the second operand when it is the same of the first operand. This effect could be similar to a repetition priming effect that largely facilitates the recognition of the  number and the activation of all the relevant nodes within the arithmetic fact memory.

Problem size effects on ERPs waveforms are rather smaller than one could have been expected on the basis of the literature and of the large differences that typically emerge in behavioural experiments. For multiplication a large bilateral broadly distributed positivity is present mainly for *Large* problems with respect to both *Medium* and *Small* with onset at about 500 ms post-stimulus and lasting for some hundreds on ms; for addition a smaller effect between 700 ms and 900 ms over the right hemisphere with *Small* problems more negative that *Medium* and *Large* ones. With addition the effect we found, even if smaller than in literature, is consistent with more demanding processes required to solve *Medium* and *Large* additions compared to *Small* ones. Within this frame it is possible that the right-sided negativity for small additions is likely to be due to the fact that the response is early selected given the problem is very easy, releasing memory load at around 700 ms from problem presentation. The more surprising result is that, while we basically replicated the slow positive wave effect for large additions, also for multiplications a similar pattern emerges instead of the expected frontal negativity (Jost et al., 2004a; 2004b).

The mismatch between our findings and the literature is very likely to be due to the specific paradigm we used: delayed production task with a modality response cue. This mismatch could be due to different macro-processes involved in the two tasks. In fact, production and verification have been supposed to be performed by means of two different macro-processes that rely on the same “knowledge base” (Zbrodoff and Logan, 1990; Zbrodoff and Logan, 2000). The different ERPs patterns could therefore be due to the different neural sources involved in these two macro-processes. More than the difference between production and verification, given that frontal negativities were found also with implicit production tasks (Jost et al., 2004a), the fact the participants did not know which will be the response modality makes it even less likely that they complete the calculation before the cue appears. In fact, if they select the response in a given format (arabic versus verbal) it is possible that the response has to be then converted in the requested modality. Behavioural data also enforce this explanation: RTs of post-cue response still show large effects of size in multiplications. Despite larger problems have larger results that could require more time for programming a motor response both in the typing and the vocal trials, differences of about 295 ms between large and small multiplication problems may make suppose that at least for some operations the selection of the response was not performed in the problem-cue interval but only after the presentation of the cue. Therefore, the frontal negativity expected for large multiplications could be diluted in time windows following the interval we analysed or even after the cue presentation. Despite all these observations that makes it difficult to drive strong theoretical implication for the findings on size effects with the present paradigm it is worthwhile to comment the effects we found.

The paradigm of delayed response with a response-modality cue, that we implemented in this study, is very useful to have clean psychophysiological data since the fact the participants does not know which response modality will be required in each trial guarantees that the ERPs elicited by the problem are not affected by motor response preparation potentials. This preparation potentials can be a strong confound especially when large RT differences between experimental cells are present. On the other side our delayed paradigm could lead the participants to postpone the selection of the response. However given the aims of this study that are more inclined to the study of the processing of the problem than the selection of the response the lack of replication of strong size effects could be regarded as an advantage rather than a problem.

With respect to the main findings of this experiment regarding order of operands, two kinds of analysis were conducted. One using the traditional approach that splits problems on the basis of their intrinsic properties (*Order* and *Size*), the other that defined the experimental cells on the basis of self-reports (*Inverted* and *Non-Inverted*). Despite possible confound and problems in terms of unbalance of stimuli the latter analysis gave more consistent and clearly interpretable results and will be thus discussed as first.

A bilateral central-frontal effect was found associated both to trials in which the participants report to adopt the "*Inversion*" procedure in both multiplications and additions. The effect is compatible in terms of polarity and topography with what predicted on the basis of the limited literature on this topic (Kiefer and Dehaene, 1997; Zhou et al, 2007). The effect was similar to what found in the traditional analysis (see below) for additions presented in the non-privileged order. Therefore, this effect could be due to an explicit, somewhat aware, process of operands reordering in working memory. The frontal effect could be produced by a process of reordering of the operands that involve abstract representation of the numbers and requires a working memory load. Interesting the effects has a more clear time development than in the previous studies, being concentrated in the 300-500 ms interval, returning promptly to baseline. The Zhou et al. (2007) effect was more long lasting and with a variable topographic development and for Kiefer and Dehaene (1997) there was also an inversion of polarity in an earlier time window. As Zhou et al. (2007) also admit it is likely that the effect in the previous studies were spurious given the superimposition of different neural sources (and thus of different cognitive processes).

In the traditional analysis based on intrinsic properties of problems two time intervals were analysed. In the 300-400 ms interval we found a large posteriorly distributed positivity for s×L order with respect to the L×s order, independently from size that thus, given our hypothesis cannot be interpreted as the correlate of reordering that reverses for large problems. We interpret this effect as due to process that compare the size of the operands before the reordering process. In fact, the operands prior to be reordered have to be compared to establish the relative size. The fact that this effect is not present in addition could be explained with the fact that the privileged order change across the size for multiplication but not for addition. In addition problems the L+s order is always privileged with respect to the s+L order, whereas in multiplication the privileged order change across the size. In both operation the comparison process has to detect the larger and the smaller operands. However, in addition it is sufficient to identify the larger operand and to order it in first position, whereas in multiplication the identification of the larger operand is not sufficient to reorder the problem in the stored order. In fact, with multiplication the reordering of the operand is based not only on the comparison of the operands but also on the size of the problems. The operands have to be reordered in the L×s order if the multiplication is small or large, whereas they have to be reordered in the s×L order if the multiplication is large. This further problem size checking could explain why in the 300-400 ms time window we found an effect only for multiplication but not for addition.

In the 400-500 ms time window we found and *Order* by *Size* interaction in multiplication in line with our predictions, namely an inversion for large problems with respect to both medium and small ones. For *Small* and *Medium* problem the s×L (non-privileged) order was more negative than the L×s order, whereas for *Large* problems the L×s (non-privileged) order was more negative than the s×L order. This effect was most pronounce on the posterior left sites. Differently, for *Large* and *Medium* additions we found in the same 300-500 ms time windows that the s+L order (non-prileged) was more negative than the L+s order. This effect involved the frontal bilateral sites and was very similar to the effect found in the self-report based analysis and thus can be interpreted in terms of an explicit reordering process.

We interpret these results above presented as due to two distinct effects. The first was left posterior between 400 and 500 ms, and was present only for multiplication. The second one was bilateral frontal between 300 and 500 ms, and was present for addition and in the self-report analysis for both operation. It is worthwhile to note that in multiplication both effects are elicited with similar latencies (300-500 vs 400-500 ms), whereas in addition only the frontal effect is present (again with a similar latencies).

The left posterior effect for multiplication has a topography similar to that of size effect. The effect at centro-parietal sites have been associated with the difficulty and the size of the problems in additions (see for example Kong et al. 1999; Núñez-Peña et al., 2006; Núñez-Peña & Escera, 2007; Núñez-Peña et al. 2011; El Yagoubi et al., 2003) and typically interpreted in terms of selection of retrieval versus non-retrieval procedures. This interpretation was however developed mainly on the basis of additions problems in which non-retrieval procedures are probably more frequently used than for multiplications. We thus prefer in this case to adopt a more generic interpretation without excluding the effect could arise from a stronger competition between the nodes in the associative network that encodes arithmetic facts and that this can cause the cognitive system to choose a non-retrieval procedure. Therefore, we propose to interpret this effect as a correlate of a difficulty in accessing to the arithmetic facts memory when the problem is presented in the non-privileged order. Moreover, the frontal effect could be associated with less automatized processes, whereas the posterior effect could be associated with an automatized process of retrieval of the result (Pauli et al., 1994). The fact that two different processes are involved in the operand order effect is consistent with the *encoding complex model* which assumes than various representations are involved in the arithmetic solving process (see Campbell, 1992; 1994; Campbell and Clark, 1988; 1992).

Interesting the two analyses for multiplications order of operand effects allows us to individuate two distinct effects that are carried out contemporaneously: the negativity effect from self-report and the posterior positivity for large non-privileged order both arise around 400 ms. This is not contradictory since, from a methodological point of view, one analysis (self-report) mediates across different sizes and problem difficulty, while the other (intrinsic property) mediates across different solution procedures. On the other side, from a theoretical point of view, multiple mechanisms may contribute to the size and order effect we described in this thesis.

The result of ERPs study presented in this Chapter are consistent with the assumption that the order of the operands affect the process require to solve an multiplication or addition problem. We found a complex ERPs pattern associated with two different effects that suggests that different processes, different neural areas, and different representations are involved in the operands-order effect. Future researches are needed to better characterize the effects we found by using both the traditional stimulus-based way to analyse ERPs data and the method of defining experimental cells on the basis of self-reports.

# Chapter 6

## General Discussion

The starting point of this thesis was to use the commutative property of additions and multiplications as a tool to study some of the properties of the memory system that encodes the arithmetic facts, namely to evaluate if the arithmetic facts memory is organized according to the commutative property, and then if the order of the operands has an effect on the performance in solution of arithmetic problems.

In the first experiment, a production task where participants were presented with one-digit additions and multiplications (Chapter 2, Experiment 1), we found an effect of the order of the operands for additions and an interaction between the order of the operands and the size of the problem for multiplications. This indicated that commutative pairs are processed differently for the two operations. For additions the problems in which the first operand is the larger (L+s) were solved faster than the commuted problem (s+L). For multiplications we found an extremely surprising result, that could not have been predicted by any of the current models of arithmetical cognition: larger-first problems (L×s) were easier to solve than smaller-first problems (s×L) only for small and medium size problems (with at least one of the operands below 5) while the opposite pattern of preference emerged for large problems (when both operands are above 5). This interaction may also explain why in the literature there is no strong evidence of order effect in multiplication given that if the effects are not in the same direction for different sizes they could have been difficult to be detected.

We offered two possible explanations of the effect based on two very different assumptions within the models of the arithmetic facts memory. Some of these models in fact assume that both orders are encoded as independent nodes in the associative network that stores arithmetic facts, others that only one of the two commuted problems is represented. The two explanations were the *reorganization hypothesis* and the *asymmetry hypothesis*.

The *reorganization hypothesis* capitalizes on Butterworth et al. (2003) proposal of a reorganization of the memory during the childhood. Butterworth and Colleagues suggested than during the acquisition of the multiplication problems the use of non-retrieval procedures and the comprehension of the commutative principle contribute to organize the multiplication facts memory so that only the L×s order is stored as arithmetic facts in the adulthood. According to the Authors "the child learning multiplication facts is not passive, simply building associative connections between problems and solutions as they are experienced in recitation or in problem presentation. Rather, the facts in memory seem to be reorganised in a principled way that takes account of a growing understanding of the commutativity, and perhaps other properties of multiplication" (Butterworth et al., 2003, p. 15 of the pre-publication draft of the 1999). In Italy the children learn the problems in the s×L order before the ones in the L×s order, then the s×L order should have an advantage with respect to the L×s order. However, following Butterworth et al. (2003), the use by children of repeated additions procedure should favour, in later acquisition stages, the L×s order since it correspond directly to the more convenient way of using repeated addictions (7×3 is easier to be transformed in 7+7+7 than 3x7 that is more likely to be represented as 3+3+3+3+3+3+3). The Butterworth et al. (2003) empirical data was limited to small and medium problem since they came from a study that was conducted during first years of schools, however the Authors concluded an overall L×s preference should arise in the adult Italian population. Our data can be however explained by assuming that reorganization does not happen when both operand are larger than 5. For these problems in fact the procedure of repeated additions is not efficient at all and thus the preference for s×L order due to primacy during learning remains.

The *asymmetry hypothesis* explains the interaction in a completely different way, by assuming that both orders are stored in memory and that the effect arises from an asymmetrical spreading of activation from the operands and the problem nodes to the correct and incorrect results. All models in fact assume, on the basis of strong empirical evidence, that during the solution of a problem other related multiples of the operands are also activated. A number of models assume that the strength of association between nodes is shaped by experience (e.g. the frequency at which a problem is solved by means of non-retrieval procedures) and that result nodes can transmit activations one each other. Our pattern of data can be explained by assuming an asymmetric spreading in the multiplication table around two key points: the beginning of the table (N×1) and the tie (N×N) that is typically much easier to solve than other problems of similar size. The asymmetry should favour the transmission of activation in the direction of larger multiples than in the direction of smaller ones, given that tables are typically recalled serially from the smaller multiple to the larger.

The other experiments reported in this thesis were planned to answer to two basic questions that arise from the above discussion. The first is, clearly, to discriminate between the two explanations we offered, the second to better understand the nature and the locus of the RT differences due to the interplay of order and size in the solution of multiplications.

To answer to the first question we first replicated the production study with English participants that learn multiplication problems in the opposite way than Italians. Learning order should be irrelevant for the *asymmetry hypothesis* but for the *reorganization hypothesis* it predicts the absence of an order by size interaction for English. This experiment however gave a null result that we discussed in Chapter 2. Independently from the possible reasons of the absence of any order effect in this experiment we believe that an order effect should emerge also for English speakers and thus further studies are necessary to clarify this point.

After this failure we decided we needed to replicate the result with Italians. To this end we adopted a different strategy to disentangle between the two hypotheses: to directly verify the key assumption of the *asymmetry hypothesis* that is the asymmetric spreading of activation around ties. The verification task in Chapter 4 allowed us to replicate the interaction and at the same time by analyzing non-matching trials to exclude the *asymmetry hypothesis* since no directional effect in the activation of multiples around the ties emerged. This replica is also informative with respect of the second main question of this thesis, the locus of the effects. A verification task is in fact more likely to be performed by using retrieval procedures only, with respect to production for which non-retrieval procedures are more frequently used. By finding the same effect in production and verification is in favour of the hypothesis that the effect is not likely to arise only because of explicit inversion or other non-retrieval procedures.

The experiments reported in Chapter 3 were aimed to see if asymmetries between the more active multiples and size can develop also outside the domain of problem resolution and thus be clearly attributed to the internal organization of the arithmetic fact memory. Despite some methodological problems with the first task we used, a paradigm (matching task) that did not require the multiplication knowledge, we did not find any systematic size-dependent asymmetry in the activation of multiples in both that matching task and the second one (multiple verification task). The second task did require the knowledge of multiplication facts. However, it did not present problems but just a number for which the participants had to decide if another number was a multiple. For this reason only a subset of the associative network that store multiplication knowledge could be activated, namely only the associations between operands and multiples but not the whole problem nodes. The absence of systematic asymmetries in this task suggests that, in case the effect we found in the production and verification task have to be attributed to a longer retrieval time, the asymmetries are driven by the activation of the whole problem in the arithmetic facts memory.

As we noted above, the behavioural results so far discussed cannot be explained by any model of the arithmetic facts memory. The models that are more suited to be adapted to explain our results are the ones that assume only one order is represented in the associative network that encodes arithmetic facts. Despite the RT costs can be either interpreted in terms of longer time to retrieval or to the use of non-retrieval procedures, the order for which the performance is the best has clearly to be assumed as the one that is represented in the arithmetic facts memory. The *interacting neighbors model* (Verguts and Fias, 2005) assumes that only half of the table is stored. Despite the model assumes the L×s order is the one represented, the Authors do not make a strong claims about this since they assume reordering process can be performed with no behavioural cost and thus predictions of the model would be the same even if the other order would be represented. First of all our results show that reordering do have a cost and also exhibit clear ERPs correlates, moreover the fact that order and size interacts make it necessary that the comparison of the size of the operands is not sufficient to decide whether the problem has to be inverted or not but an evaluation of the whole problem size has also to be considered. Clearly our effect can be explained by other models that as well would need similar or even larger extra assumptions to explain our complex pattern of data.

If we feel confident in the answers we gave to the first of the two main questions of this thesis (explanation of the interaction order by size in multiplications), much less clear is the interpretation of the locus of the effect and the mechanism that favours the performance for one of the two orders. The main problem both in interpreting behavioural data and even more crucially in interpreting the ERPs components is whether these can be attributed to retrieval only procedures, to non-retrieval procedures or both of them. This problem plagues all empirical literature on arithmetical cognition. Just as an example, in most of the modelling literature the assumption that adults solve one digit problems by means of retrieval only is widely diffused, on the other side most ERPs papers interpret the effects in terms of selection/activation of non retrieval procedures. The truth, as usually happens, seems to stay in the middle. Both self-reports, that clearly patterned chronometric data in Chapter 2, and especially our ERPs results suggest that the order preferences and their interaction with size cannot uniquely be attributed to either retrieval or non-retrieval procedures and that it is likely that both explicit inversions and implicit reordering contribute to the effects we found. Our ERPs study however clearly showed that the paradigm we used, with a delayed response with a modality cue, can be useful to study the early stages of arithmetic problem parsing. Moreover the use of both self reports and stimulus based analyses can furnish complementary data and thus we think that by applying this methodology could successfully lead of a better understanding of the role and the mechanisms of retrieval and non-retrieval procedures used to solve arithmetical problems.

## Acknowledgments

I would like to thanks Luca Polonio, Luigi Lombardi, Elena Rusconi, Giovanni Galfano, Brian Butterworth, Marinella Cappelletti, Bahador Baharami, André Knops, Manuela Piazza for the helpful comments and inputs they gave me in the development of this project.

# APPENDIX

non-matching trials

| | Tie+1 | Tie-1 | Neutral+1 | Neutral-1 | Filler |
|---|---|---|---|---|---|
| cues | 4 | 5 | 4 | 5 | 33 |
| probe | 20 | 20 | 38 | 38 | 51 |
| cues | 5 | 6 | 5 | 6 | 44 |
| probe | 30 | 30 | 22 | 22 | 65 |
| cues | 6 | 7 | 6 | 7 | 62 |
| probe | 42 | 42 | 34 | 34 | 39 |
| cues | 7 | 8 | 7 | 8 | 69 |
| probe | 56 | 56 | 52 | 52 | 46 |
| cues | 8 | 9 | 8 | 9 | 75 |
| probe | 72 | 72 | 66 | 66 | 58 |

matching trials

| | Cue balancing +1 | Cue balancing -1 | Probe balancing Multiple | Probe balancing Neutral | Filler |
|---|---|---|---|---|---|
| cues | 4 | 5 | 20 | 38 | 33 |
| probe | 4 | 5 | 20 | 38 | 33 |
| cues | 5 | 6 | 30 | 22 | 65 |
| probe | 5 | 6 | 30 | 22 | 65 |
| cues | 6 | 7 | 42 | 34 | 39 |
| probe | 6 | 7 | 42 | 34 | 39 |
| cues | 7 | 8 | 56 | 52 | 46 |
| probe | 7 | 8 | 56 | 52 | 46 |
| cues | 8 | 9 | 72 | 66 | 75 |
| probe | 8 | 9 | 72 | 66 | 75 |

**Appendix 1: stimuli adopted in the matching task (see chapter 3).**

| **Cue** | **Tie-1** | **Tie** | **Tie+1** | **Below Tie-1** | **Below sotto Tie** | **Above Tie** | **Above Tie+1** |
|---|---|---|---|---|---|---|---|
| 4 | 12 | 16 | 20 | 10 | 14 | 18 | 22 |
| 5 | 20 | 25 | 30 | 18 | 23 | 27 | 32 |
| 6 | 30 | 36 | 42 | 27 | 33 | 39 | 45 |
| 7 | 42 | 49 | 56 | 39 | 46 | 52 | 59 |
| 8 | 56 | 64 | 72 | 52 | 60 | 68 | 76 |

**Appendix 2: stimuli adopted in the multiple task (see chapter 3).**

## REFERENCES

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

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 527–538.

Butterworth, B., Marchesini, N., & Girelli, L. (2003). Multiplication facts: Passive storage or dynamic reorganization? In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills* (pp. 189–202. ). Mahwah, NJ: Erlbaum

Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. R. (2001). Storage and retrieval of addition facts : The role of number comparison. *The Quarterly Journal Of Experimental Psychology Section A*, 54, 1005–1029.

Campbell, J. I. D. (1987a). Network interference and mental multiplication. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 109–123.

Campbell, J. I. D. (1987b). Production, verification, and priming of multiplication facts. *Memory & Cognition*. 15, 349–364.

Campbell, J. I. D. (1987c). The role of associative interference in learning and retrieving arithmetic facts. In J. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 107–122). London: Oxford University Press.

Campbell, J. I. D. (1992). In defense of the encoding-complex approach: Reply to McCloskey, Macaruso & Whetstone. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 539–556). Amsterdam: Elsevier Science.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition*, 53, 1–44.

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

Campbell, J. I. D. (1999). Division by multiplication. *Memory & Cognition*, 27, 791–802.

Campbell, J. I. D., & Austin, S. (2002). Effects of response time deadlines on adults' strategy choices for simple addition. *Memory & Cognition*, 30, 988–994.

Campbell, J. I. D., & Clark, J. M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, & Goodman (1986). *Journal of Experimental Psychology: General*, 117, 204–214.

Campbell, J. I. D., & Clark, J. M. (1989). Time course of error priming in number-fact retrieval: Evidence for excitatory and inhibitory mechanisms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 920–929.

Campbell, J. I. D., & Clark, J. M. (1992). Numerical cognition: An encoding-complex perspective. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 457–491). Amsterdam: Elsevier Science.

Campbell, J. I. D., Dowd, R. R., Frick, J. M., McCallum, K. N., & Metcalfe, A. W. S. (2011). Neighborhood consistency and memory for number facts. *Memory & cognition*, 39, 884–93.

Campbell J. I. D., & Epp, L. J. (2005), Architectures for arithmetic, In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 347–360). NY: Psychology Press.

Campbell, J. I. D., Fuchs-Lacelle, S., & Phenix, T. L. (2006). Identical elements model of arithmetic memory : Extension to addition and subtraction. *Memory & Cognition*, 34, 633–647.

Campbell, J. I. D., & Graham, D. J. (1985). Mental multiplication skill: Structure, process, and acquisition. *Canadian Journal of Psychology*, 39, 338–366.

Campbell, J.I.D., & Gunter, R. (2002). Calculation, culture, and the repeated operand effect. *Cognition*, 86, 71–96.

Campbell, J. I. D., & Oliphant, M. (1992). Representation and retrieval of arithmetic facts: A network-interference model and simulation. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp.331–364). Amsterdam: Elsevier.

Campbell, J. I. D., & Penner-Wilger, M. (2006). Calculation latency: the mu of memory and the tau of transformation. *Memory & cognition*, 34, 217–226.

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

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44, 1–42.

Dehaene, S., & Cohen, L. (1995). Toward an anatomical and functional model of number processing. *Mathematical Cognition*, 1, 83–120.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuit for number processing. *Cognitive Neuropsychology*, 20, 487–506.

Delorme, A., & Makeig, S. (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics, *Journal of Neuroscience Methods*, 134, 9–21.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (1999). The contribution of working memory resources in the verification of simple arithmetic sums. *Psychological Research*, 62, 72–77.

De Rammelaere, S., Stuyven, E., & Vandierendonck, A. (2001). Verifying simple arithmetic sums and products: Are the phonological loop and the central executive involved? *Memory and Cognition*, 29, 267–273.

El Yagoubi, R., Lemaire, P., & Besson, M. (2003). Different brain mechanisms mediate two strategies in arithmetic: Evidence from event-related brain potential. *Neuropsychologia*, 41, 855–862.

French, J. W., Ekstrom, R. B. , Price, L. A. (1963). *Manual for kit of reference tests for cognitive factors*. Educational Testing Service, Princeton, NJ.

Galfano, G., Mazza, V., Angrilli, A., & Umiltà, C. (2004). Electrophysiological correlates of stimulus-driven multiplication facts retrieval. *Neuropsychologia*, 42, 1370–1382.

Galfano, G., Penolazzi, B., Vervaeck, I., Angrilli, A., & Umilta, C. (2009). Event-related brain potentials uncover activation dynamics in the lexicon of multiplication facts. *Cortex*, 45, 1167–1177.

Galfano, G., Rusconi, E., & Umiltà, C. (2003). Automatic activation of multiplication facts: Evidence from the nodes adjacent to the product. *The Quarterly Journal Of Experimental Psychology Section A*, 56, 31–61.

Galton, F. (1894). Arithmetic by smell. *Psychological Review*, 1, 61–62.

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia*, 47, 604–608.

Graham, D. J. (1987). An associative retrieval model of arithmetic memory: How children learn to multiply. In J. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 123–141). London: Oxford University Press.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, 79, 329–343.

Hecht, S. A. (1999). Individual solution processes while solving addition and multiplication math facts in adults. *Memory & cognition*, 27, 1097–1107.

Hunt, E., & Agnoli, F. (1991). The Whorfian Hypothesis: A Cognitive Psychology Perspective. *Psychological Review,* 98, 377–389.

Jost, K., Beinhoff, U., Hennighausen, E., & Rosler, F. (2004). Facts, rules, and strategies in single-digit multiplication: evidence from event-related brain potentials. *Cognitive Brain Research*, 20, 183–193.

Jost, K., Hennighausen, E, & Rosler, F. (2004). Comparing arithmetic and semantic fact retrieval: Effects of problem size and sentence constraint on event-related brain potentials, *Psychophysiology*, 41, 46–59.

Kiefer, M., & Dehaene , S. (1997). The time course of parietal activation in single-Digit multiplication: Evidence from event-related potentials, Mathematical Cognition, 3, 1–30.

Kirk, E. P., & Ashcraft, M. H. (2001). Telling stories: The perils and promise of using verbal reports to study math strategies. *Journal of*

*Experimental Psychology: Learning, Memory, and Cognition*, 27, 157–175.

Kong, J., Wang, Y., Shang, H., Wang, Y., Yang, X., Zhuang D., (1999). Brain potentials during mental arithmetic - effects of problem difficulty on event-related brain potentials. *Neuroscience Letters*, 260, 169–172.

Kutas, M., Van Petten, C., and Kluender, R. (2006). Psycholinguistics electrified II: 1994-2005. In: M. Traxler & M.A. Gernsbacher (Eds.), *Handbook of Psycholinguistics* (pp. 659–724), 2nd Edition, New York: Elsevier.

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

LeFevre J. A, Bisanz J, and Mrkonjic L. (1988). Cognitive arithmetic: evidence for obligatory activation of arithmetic facts. *Memory and Cognition*, 16, 45–53.

LeFevre & Liu (1997), The Role of Experience in Numerical Skill: Multiplication Performance in Adults from Canada and China, *Mathematical Cognition*, 3, 31–62

LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996b). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 216–230.

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

McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of Arithmetic Fact Retrieval: An Evaluation in Light of Findings From Normal and

Brain-Damaged Subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 377–397.

McCloskey, M., & Macaruso, P. (1995). Representing and using numerical information. *American Psychologist*, 50, 351–363.

Metcalfe, A. W. S., & Campbell, J. I. D. (2011). Adults' strategies for simple addition and multiplication: verbal self-reports and the operand recognition paradigm. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 37, 661–672.

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

Niedeggen, M., & Rosler, F., & Jost, K. (1999). Processing of incongruous mental calculation problems: Evidence for an arithmetic N400 effect. *Psychophysiology*, 36, 307–324.

Núñez-Peña, M. I., Cortinas, M., Escera, C., (2006). Problem size effect and processing strategies in mental arithmetic. Neuroreport, 17, 357–360.

Núñez-Peña, M I., & Escera, C. (2007). An event-related brain potential study of the arithmetic split effect. *International Journal of Psychophysiology*, 64, 165–173.

Núñez-Peña, M. I., Gracia-Bafalluy, M., & Tubau, E. (2011). Individual differences in arithmetic skill reflected in event-related brain potential. *International Journal of Psychophysiology*, 80, 143–149.

Pauli, P., Lutzenberger, W. Rau, H., Birbaumer, N., Rickard, T. C., Yaroush, R. A., Bourne Jr., L. E. (1994). Brain potentials during mental arithmetic: Effects of extensive practice and problem difficulty. *Cognitive Brain Research*, 2, 21–99.

Parkman, J. M. (1972). Temporal aspects of simple multiplication and comparison. *Journal of Experimental Psychology*, 95, 437–444.

Parkman, J. M., & Groen, G. J. (1971). Temporal aspects of simple addition and comparison. *Journal of Experimental Psychology*, 89, 335–342.

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W. Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37, 127–152.

Prieto-Corona, B., Rodríguez-Camacho, M., Silva-Pereyra, J., Marosi, E., Fernández, T., & Guerrero, V. (2010). Event-related potentials findings differ between children and adults during arithmetic-fact retrieval. *Neuroscience Letters*, 468, 220–224.

Rickard T. C. (2005), A revised identical elements model of arithmetic fact representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 250–257.

Rickard T. C., Bourne L. E. Jr (1996), Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1281–1295.

Rickard T. C., Healy A. F., Bourne L. E. Jr (1994), On the Cognitive Structure of Basic Arithmetic Skills: Operation, Order, and Symbol Transfer Effects, *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 20, 1139–1153.

Robert, D. N., & Campbell, J. I. D. (2008). Simple addition and multiplication: No comparison. *European Journal of Cognitive Psychology*. 20, 123–138.

Rusconi, E., Galfano, G., Rebonato, E., & Umiltà, C. (2006). Bidirectional links in the network of multiplication facts. *Psychological Research*. 70, 32–42.

Rusconi, E., Galfano, G., Speriani, V., & Umiltà, C. (2004). Capacity and contextual constraints on product activation : Evidence from task-irrelevant fact retrieval. *The Quarterly Journal of Experimental Psychology*, 57A, 1485–1511.

Seitz, K. & Schumann-Hengsteler, R. (2000) Mental multiplication and working memory, *Journal of Cognitive Psychology*. 12, 552–570.

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*. 117, 258–275.

Siegler, R. S., & Shipley, E. (1995). Variation, selection and cognitive change. In G. Halford&T. Simon (Eds.), *Developing cognitive competence: New approaches to process modelling* (pp. 31–76). Hillsdale, NJ: Erlbaum.

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale NJ: Erlbaum.

Smith-Chant, B. L., & LeFevre, J.-A. (2003). Doing as they are told and telling it like it is: self-reports in mental arithmetic. *Memory & Cognition*, 31, 516–528.

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

Thevenot, C., Fanget, M., & Fayol, M. (2007). Retrieval or nonretrieval strategies in mental arithmetic? An operand recognition paradigm. *Memory & cognition*. 35, 1344–1352.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology*. 47A, 631–650.

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

Zbrodoff, N. J., & Logan, G. D. (1990). On the relation between production and verification tasks in the psychology of simple arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 16, 83–97.

Zbrodoff, N. J., & Logan, G. D. (2000). When it hurts to be misled: a Stroop-like effect in a simple addition production task. *Memory & cognition*. 28, 1–7.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem-size effect. In J. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–345). NY: Psychology Press.

Zhou, X., Booth, J. R., Lu, J., Zhao, H., Butterworth, B, Chen, C, & Dong, Q. (2011). Age-Independent and Age-Dependent Neural Substrate for Single-Digit Multiplication and Addition Arithmetic Problems. *Developmental Neuropsycholog*y, 36, 338–352.

Zhou, X., Chen, C., Dong, Q., Zhang, H., Zhou, R., Zhao, H., Chen, C., Qiao, S., Jiang, T., & Guo, Y. (2006). Event-related potentials of single-digit addition, subtraction, and multiplication. *Neuropsychologia*. 44, 2500–2507.

Zhou, X., Chen, C., Zhang, H., Chen, C., Zhou, R., & Dong, Q. (2007). The operand-order effect in single-digit multiplication: An ERP study of Chinese adults. *Neuroscience Letters*. 414 41–44.