Psicológica (2015), 36, 265-281.

Effects of numerical surface form in arithmetic word problems

Josetxu Orrantia[1], David Múñez[1], Sara San Romualdo[1] and Lieven Verschaffel[2]

[1]*Universidad de Salamanca, Spain*

[2]*Katholieke Universiteit Leuven, Belgium*

Adults' simple arithmetic performance is more efficient when operands are presented in Arabic digit (3 + 5) than in number word (three + five) formats. An explanation provided is that visual familiarity with digits is higher respect to number words. However, most studies have been limited to single-digit addition and multiplication problems. In this article, we examine to what extent format effects can be found in the context of arithmetic word problems, in which visual familiarity is reduced (Manuel had 3 marbles, and then Pedro gave to him 5). Participants with high and low arithmetic fluency solved addition and subtraction word problems in which operands were presented in both formats. The overall results showed an advantage for digits operands relative to words operands. In addition, the format effects were more evident for subtraction and low-skilled participants. These results were interpreted in terms of more rapid access of digits to numerical magnitude.

To what extent number format affects the representation of numerical information? The answer to this question is a central issue in the field of numerical cognition (Cohen Kadosh & Walsh, 2009). On the one hand, it is assumed that the numerical representation of quantity is abstract or amodal and independent of notation (e.g., Dehaene & Akhavein, 1995; Naccache & Dehaene, 2001; Schwarz & Ischebeck, 2000); on the other hand, it is suggested that such representation is not abstract, and it is mediated by modality-specific processes (e.g., Cohen Kadosh, 2008; Cohen Kadosh,

---

[1] This research was supported by grant PSI2011-27737 from the Spanish Ministerio de Ciencia e Innovación. Corresponding author: Josetxu Orrantia. Facultad de Educación. Pº de Canalejas, 169. 37008 Salamanca, Spain. Email: orrantia@usal.es

Henik, & Rubinstein, 2008). Much of this research has focused on exploring whether number format influences well-described effects related to quantity processing. Some of those effects are the "numerical distance effect" (discrimination between two numbers is faster when the distance between them increases; Moyer & Landauer, 1967), or the "SNARC effect" (responses to relatively larger numbers are faster for the right hand, those to smaller numbers for the left hand; Dehaene, Bossini, & Giroux, 1993). In this article, however, we analyze the influence of format in elementary arithmetic, a research area that has generated the same debate that involves the numerical representation of quantity and the idea of abstraction (Campbell & Metcalfe, 2009). Thus, several works (see Campbell & Epp, 2005, for a review) indicate that adults are slower and more error prone in simple arithmetic tasks when operands are displayed as number words (e.g., two + four) than when they are presented in Arabic digits (e.g., 2 + 4). This effect, the so-called "format effect", has led to different interpretations based on the above-mentioned two theoretical perspectives. One interpretation is that calculation relies on amodal representations that are abstracted from surface forms, whereas another different explanation suggests that representations depend on modality. These interpretations have arisen from different models describing the componential architecture of numerical processing and calculus, and how the various, involved processes are interrelated.

Thus, according to McCloskey´s *single-abstract-code* model (McCloskey, 1992; McCloskey & Macaruso, 1995; Sokol, McCloskey, Cohen, & Aliminosa, 1991), the format effect arises while operands are being encoded and does not affect the mechanisms of calculation. These mechanisms would operate on an abstract representation that does not rely on format. In other words, operands (whether they are numeric words or Arabic digits) are converted into a single abstract numerical representation before the calculation processes are triggered. This means that encoding and calculation processes are independent and additive. Meanwhile, Dehaene (1992; Dehaene & Cohen, 1995) proposed the *triple code* model. According to that model there would not be just a single abstract numerical representation but three internal codes to represent numbers: a verbal code used for verbal counting and arithmetic-facts retrieval, an visual Arabic code for multi-digit calculation and parity judgment, and an analog magnitude code used for numerical comparison and calculus estimation. However, this model does not predict format effects on the mechanisms of calculation since each format, number words or Arabic digits, is transferred to a verbal code before calculation starts (Noël, Fias, & Brysbaert, 1997), mainly for overlearned calculations such as single-digit addition and multiplication problems (but not for subtraction problems, in which a magnitude code is used). Therefore, coding and calculation processes remain independent. In a different vein, the *complex coding* model (Campbell, 1992, 1994; Campbell & Clark, 1988, 1992) suggests that operands´ format affects not only coding but also calculation. This model assumes that numbers are not represented in an abstract fashion but different numerical codes exist, which depends on modality. Therefore, calculation processes would be mediated by specific format encoding processes rather than by an abstract code. Campbell's model suggests that those processes interact.

An argument that is often used by Campbell and colleagues to support their model is that operands´ format interacts with the so-called "problem-size effect" (PSE). PSE has been found in numerous studies (see Ashcraft, 1992; Ashcraft & Guillaume, 2009; and Zbrodoff & Logan, 2005, for various revisions), and refers to the fact that reaction times to simple arithmetic problems are, in general, slower and more error prone if the operands and their correct solutions become numerically larger. Different explanations have been suggested to explain PSE. Thus, some models assume that the association between small numbers and response in memory is stronger (Ashcraft, 1982, 1987; Campbell & Graham, 1985; Siegler & Jenkins, 1989; Siegler & Shipley, 1995), probably, because small problems are more frequently practiced (Hamann & Ashcraft, 1986). Alternatively, it is possible that different procedures, others than memory retrieval for large problems, may contribute to PSE. In this sense, there is evidence that adults not only use retrieval procedures when solving simple arithmetic problems (e.g., Campbell & Xue, 2001; Hecht, 1999; LeFevre, Sadesky, & Bisanz, 1996; Thevenot, Fanget, & Fayol, 2007). For example, Lefevre et al. (1996) found that adults use a variety of strategies to solve single-digit addition, such as counting or transformations (e.g., 9 + 6 is 9 + 1 + 5). And, consequently, these strategies increased with problem size. Since strategies are slower than retrieval from memory, it is possible that PSE is mediated by its use. But, what is important in the context of this article is that different studies have shown that PSE is larger when operands are presented as number words than when presented in Arabic digits format (e.g., Campbell, 1994; Campbell & Alberts, 2009; Campbell & Clark, 1992; Campbell & Fugelsang, 2001; Campbell & Penner-Wilger, 2006; Jackson & Coney, 2007a; Noël, Fias, & Brysbaert, 1997).

In a novel study, Campbell and Clark (1992) asked participants to solve basic multiplications presented in Arabic digit format or number word format. Their results showed an interaction between problem size and format (longer response times and higher error rates for large problems presented in number word format). Since PSE can only occur during calculation, the cognitive cost of the number word format cannot be solely attributed to operands coding but it also impacts on calculation. However, McCloskey, Macaruso, and Whetstone (1992; see also Noël et al. 1997) argued that the interaction between format and problem size could be attributed to encoding differences between number words and Arabic digits since words demand more time to be coded than Arabic digits. Hence, those differences increase with numerical magnitude. In a later study, however, Campbell (1994) contrasted format effects in addition and multiplication. The author found that the interaction between format and problem size was larger for addition than for multiplication. This finding would be difficult to explain in terms of encoding processes since the same operands were used for both operations. More recently, Campbell et al. investigated the interaction between format and problem size by analyzing format effects in strategy choice to solve simple arithmetic problems (Campbell & Alberts, 2009; Campbell & Fugelsang, 2001; Campbell, Parker, & Doetzel, 2004). In these studies, participants reported their solution strategy (e.g. retrieval or calculation strategies such as counting and transformations). Results of these studies showed that participants reported more calculation strategies for stimuli presented as number words than for stimuli presented as Arabic digits. In addition, that difference increased with problem size. These findings support the idea that format affects the mechanisms involved in calculation and not just the encoding process.

Therefore, although some researchers believe that format effects work only during the encoding processes and, hence, do not affect calculation, others provide evidence that format directly affects calculation. The explanation offered by the latter is particularly relevant for the present study. Campbell and colleagues have repeatedly suggested that since simple arithmetic is presented rarely as number words, then, the visual familiarity of this format is lower compared to the familiarity of the Arabic format. Hence, the strength of association between operands and answer will be weaker (e.g., Campbell, 1994; Campbell et al., 2004) and will influence the use of strategies. Less familiarity promotes greater use of calculation strategies against direct memory retrieval (Campbell & Fugelsang, 2001). According to the *complex coding* model, numbers can be presented in different codes (e.g., visual, verbal, visuospatial...) that are used in different tasks (e.g., addition, reading, counting...). The experience and practice with a particular format, such as simple arithmetic in Arabic format, would optimize individuals' performance in that format. Based on terms of visual familiarity, this argument is supported by Colome, Bafalluy, and Noël

(2011), who reported that adults prefer Arabic numerals to run arithmetic operations.

However, most of the studies that have examined format effects in simple calculation have presented numbers in the context of simple arithmetic problems (i.e., 3 + 5 vs. three + five), where visual familiarity may intervene. Less is known about the effects of format in tasks where calculation is one of the components of the task. In the present study, we analyze the influence of format in the context of simple arithmetic *word* problems (e.g., Manuel had 3 marbles and someone gave him 5 more). Two reasons underlie our rationale. First, since calculation is a component of solving word problems, our task allows analyzing the processes underlying basic calculation (Orrantia, Múñez, Vicente, Verschaffel, & Rosales, 2014; Orrantia, Rodriguez, & Vincent, 2010; Orrantia, Rodriguez, Múñez, & Vincent, 2012). Second, and more important, our stimuli operands do not induce any activation of a "more-or-less-familiar" visual representation. Moreover, since our task has a strong reading component, it would be feasible that the number word format is more familiar in the context of an arithmetic word problem (e.g., Manuel had three marbles and someone gave him five more) than in simple arithmetic tasks (e.g., three + five).

The aim of this research was to analyze whether calculation processes are modulated by format when operands are presented in the context of an arithmetic word problem. In other words, we aim at exploring whether format effects also occur after the potential visual familiarity of Arabic digits format is removed. These effects were analyzed in terms of speed and accuracy. Although some studies have explored the influence of format in strategy choice (retrieval vs. strategies) by asking participants the strategies they used for, other authors suggest that the validity and accuracy of self-reports should be considered with caution (Kirk & Ashcraft 2001; Thevenot, Castel, Fanget, & Fayol, 2010). Moreover, recently Fayol and Thevenot (2012) have suggested that using some strategies can achieve such a degree of automaticity that a conscious access can be difficult, and often confused with memory retrieval. Assuming the theoretical interest that analyzing format effects on strategy choice may have, our aim was to study how format affects performance on calculation processes regardless of strategy choice. However, our study added to previous research by considering a component that plays an important role in arithmetic word problem solving and calculation, namely, arithmetic fluency. In this sense, we know that less skilled individuals are also less effective in arithmetic facts retrieval, calculation strategies and, in addition, they also show a larger problem size effect than their counterparts (e.g., Hecht, 1999; Jackson & Coney, 2007b; LeFevre, Bisanz, et al, 1996; LeFevre et al, 1996;

Thevenot et al, 2007). Thus, in order to fully understand format effects on arithmetic word problem solving, our sample included participants with different levels of competence in arithmetic fluency.

In light of these considerations, we designed arithmetic word problems with two types of operands or *format*, Arabic digits and number words, within two *problem-size* categories (large vs. small). According to Campbell et al.´s model the effect of format on arithmetic word problem solving would be validated by an interaction between format and problem size. Response times and error rate would be higher for stimuli presented as number words than for those presented as Arabic digits. And these differences would be more evident for large-size problems than for small-size problems. In addition to format and problem size, two variables were introduced in our design: *operation* (addition vs. subtraction) and participants´ *proficiency in arithmetic fluency* (high vs. low). Regarding the variable operation, most of the previous studies have used addition and multiplication to analyze format effects. To the best of our knowledge, only one study has used subtractions (Campbell & Alberts, 2009). The authors found longer response times for large size problems relative to small ones, when stimuli were presented as number words. This cognitive cost was similar for both, addition and subtraction operations. Accordingly, we would expect similar format effects for both operations in the context of arithmetic word problems. As regards proficiency in arithmetic fluency, we are not aware of any previous studies analyzing the influence of format according to this variable. However, since less skilled individuals are less proficient in calculation, a larger format effect would be expected for this group of participants.

METHOD

Participants. Sixty-seven undergraduate students from the University of Salamanca participated for course credit. The mean age of the students was 22.3 years (SD = 2.1), and they were ranked as a function of their score on an arithmetic fluency test. In order to maximize the chances to obtain two heterogeneous groups in arithmetic fluency, participants whose scores were nearest the median (i.e., above the 35th percentile and below the 65th percentile) were eliminated. The median score of the whole population was 18, and the low-skill group comprised 25 participants (19 females and 6 males) whose mean score on the arithmetic fluency test was 12.08 (SD = 2.27; range of 8–15), whereas the high-skill group comprised 25 participants (15 females and 9 males) whose mean score was 23.24 (SD = 1.88; range of 20–26) [$t(48) = 16.53, p < .0001$].

Materials

Arithmetic fluency. As a preliminary test, participants completed an arithmetic fluency test (adapted from the Thurstone's Primary Mental Abilities Test), which consists of 50 addition problems involving two or three single digits (e.g., 9 + 6; 7 + 5 + 8) or double digits (e.g., 43 + 72; 67 + 58 + 45) presented vertically. All participants were instructed to solve the problems as fast and accurately as possible. The total arithmetic score was the number of problems correctly solved in 1 min.

Experimental stimuli. Each experimental trial consisted of an elementary word problem without the usual question. All the word problems belonged to Change 1 (addition) and Change 2 (subtraction) according to the classification scheme for one-step addition and subtraction word problems of Riley, Greeno, and Heller (1983). The problems were always of the form "Pedro had X marbles and then gave/removed him Y", and they differed in the name of the protagonists, the nature of the objects, and the size of the numbers. The operands used in the word problems consisted of 56 pairs of numbers corresponding to all possible pairings of the digits 2 through 9 when ties (e.g., 3 + 3) are excluded. For subtractions, each number of the pair became the subtrahend. To enable testing of the problem size effect, problems composed from pairs with a product ≤ 25 were classified as small, whereas those with a product ≥ 25 were classified as large.

Procedure. Participants received four blocks of 56 trials with operation (i.e., addition or subtraction) and format (i.e., digits or words) alternating across blocks. Block order was counterbalanced, and problem order in each block was randomized with the constrain that the same operands or result did not appear in succession. Participants were informed of the characteristics of the stimuli and that they should respond to each stimulus to the question "How many objects have the character at the end". The stimuli were presented using SuperLab software, which ran on computers with 15-inch monitors. Word problems were presented word by word (Geneva 36-point font), at a fixed pace, using rapid serial visual presentation (RSVP). Each word was exposed for 300 ms, and there was a 50-ms interval between words. Each trial began with the participants pressing the space bar when a 1,000-ms ready signal ("***") appeared in the centre of the screen. The words of the word problem then appeared one at a time in the centre of the screen. The last word was the second number of the corresponding operation, and it remained on the screen until the participant responded. The participants were instructed to solve the word problem correctly as quickly as possible. A microphone connected to a voice-activated relay and interfaced with the computer registered the latency of the responses. Response timing began when the second number of the operation appeared and stopped when the microphone detected the verbal response. To familiarize participants with the procedure, they performed five practice trials prior to each block.

RESULTS

Mean response times (RT) and percentage of errors (see Table 1) were analysed in a 2 (skill level: high, low) x 2 (problem size: small, large) x 2 (operation: addition, subtraction) x 2 (format: digits, words) analysis of variance (ANOVA) with the first factor as a between-subject measure and the last three factors as repeated measures.

Response times

A total of 9.1% of RTs were excluded because they were incorrect responses, microphone failure, or outliers deviating more than 3 standard deviations from the participant's mean, with no significant differences between conditions respect to the last two aspects. All main effects were significant. Mean RT was slower in low-skilled (2115 ms) than in high-skilled (1186 ms) participants [$F(1, 48) = 183.29$, $p < .0001$, $\eta^2_p = .79$], slower for subtraction (1841 ms) than for addition (1460 ms) [$F(1, 48) = 100.16$, $p < .0001$, $\eta^2_p = .68$], slower with large (1967 ms) than with small (1335 ms) problems [$F(1, 48) = 230.69$, $p < .0001$, $\eta^2_p = .83$], and slower for words (1750 ms) than for digits (1552 ms) stimuli [$F(1, 48) = 97.21$, $p < .0001$, $\eta^2_p = .67$]. The skill level factor interacted with the other factors, such that the operation effect was larger in low-skilled (+568 ms) than in high-skilled (+194 ms) participants [$F(1, 48) = 24.21$, $p < .0001$, $\eta^2_p = .33$], the size effect was larger in low-skilled (+976 ms) than in high-skilled (+287 ms), and the format effect was larger in low-skilled (+266 ms) than in high-skilled (+129 ms) [$F(1, 48) = 11.76$, $p < .001$, $\eta^2_p = .20$]. The operation by size interaction was significant [$F(1, 48) = 30.73$, $p < .0001$, $\eta^2_p = .39$], reflecting a larger size effect for subtraction (+804 ms) than for addition (+461 ms), and this was larger in low-skilled than in high-skilled participants [$F(1, 48) = 6.68$, $p < .02$, $\eta^2_p = .13$]. There was also a significant interaction between format and size [$F(1, 48) = 14.57$, $p < .0001$, $\eta^2_p = .24$], due to greater word-format costs for large (+262 ms) than for small (+133 ms) problems. And this interaction was modulated by the skill level [$F(1, 48) = 5.63$, $p < .03$, $\eta^2_p = .11$], due to the greater word-format cost for large problems was more pronounced in low-skilled participants. Despite the operation by format was not significant ($F < 1$), the three-way interaction of operation, format, and size reached the significance $F(1, 48) = 4.51$, $p < .04$, $\eta^2_p = .09$], because the greater word-format cost for large problems was larger for subtraction. Lastly, the four-way interaction approached the standard levels of significance [$F(1, 48) = 3.47$, $p = .07$, $\eta^2_p = .07$]. However, given our interest in how the format effect is modulated by size and operation as a function of skill level, RTs were analysed separately for each skill level using a 2 (problem size: small, large) x 2 (operation: addition, subtraction) x 2 (format: digits, words) repeated measures ANOVA. Taking into account the most interesting interactions, for high-skilled participants only the format by size interaction approached the significance [$F(1, 24) = 3.93$, $p = .059$, $\eta^2_p = .14$], whereas for low-skilled participants both the format by size [$F(1, 24) = 11.04$, $p < .003$, $\eta^2_p = .32$] and the format by size by operation [$F(1, 24) = 4.97$, $p < .04$, $\eta^2_p = .17$] interactions reached the significance, reflecting greater word-format costs for large than for small problems, being this effect larger for subtraction than for addition.

Table 1. Mean (and standard error) response times and percentage of errors as a function of skill level, operation, problem size, and format.

	High skill				Low skill			
	Addition		Subtraction		Addition		Subtraction	
	Large	Small	Large	Small	Large	Small	Large	Small
	RT							
Digits	1098 (57)	923 (41)	1409 (99)	1058 (49)	2047 (57)	1375 (41)	2789 (99)	1717 (49)
Words	1277 (71)	1060 (42)	1537 (99)	1129 (57)	2339 (71)	1561 (42)	3237 (99)	1769 (57)
% Errors								
Digits	7.3 (1.9)	1.4 (0.9)	10.4 (2.4)	2.0 (1.0)	12.4 (1.9)	3.4(0.9)	19.0 (2.4)	4.1 (1.0)
Words	5.9 (1.6)	1.3 (0.6)	11.7 (2.3)	3.9 (0.9)	10.9 (1.6)	2.9(0.6)	20.3 (2.3)	4.9 (0.9)

Errors

There were 7.6% of incorrect responses. The ANOVA indicated that low-skilled participants made more errors than high skilled participants (9.7% vs. 5.5%) [$F(1, 48) = 9.81$, $p < .003$, $\eta^2_p = .17$]; there were more errors for subtraction than for addition (9.5% vs. 5.7) [$F(1, 48) = 23.51$, $p <$

.0001, $\eta^2_p = .33$]; and more errors for large than small problems (12.2% vs. 3%) [F(1, 48) = 92.96, p < .0001, $\eta^2_p = .66$]. The size by skill level interaction was significant [F(1, 48) = 7.19, p < .005, $\eta^2_p = .13$], due to a larger problem size effect in low-skilled than in high-skilled participants. Furthermore, the size by operation interaction was significant [F(1, 48) = 4.47, p < .04, $\eta^2_p = .09$], reflecting a larger problem size effect for subtraction than for addition. The only significant effect related to the format condition was the size by format interaction due to greater word-format costs for subtraction relative to addition operation. Campbell and colleagues (Campbell et al., 2004; Campbell & Alberts, 2009) also found no effects of format on errors. According to these authors, format effects on RT, errors, or both would depend on participants' emphasis on speed versus accuracy.

DISCUSSION

The present study aimed at exploring format effects in calculation when operands are presented within the context of arithmetic word problems. Results showed better performance when stimuli were presented as Arabic digits as opposed to those displayed as number words. This effect was more evident for large problems respect to small ones. This result replicates the interaction between format and size that has been found in previous studies. Furthermore, the interaction was more evident for subtraction than for addition. Findings also revealed that less-skilled participants showed larger effects.

Since the context of the task reduces the visual familiarity of the Arabic format, then, how to explain these effects? A prior issue to be considered is how operands and sign were presented. In contrast to previous studies, where both, operands and sign, used to appear simultaneously, in the present study operands and sign were sequentially presented. In this sense, format effects would be associated to the processing of the second operand since the experimental condition allows enough time for the first operand to be processed regardless of format. Therefore, our sequential presentation would reduce the potential format effects, however, this was not true. Campbell (1999) already manipulated the presentation of operands, simultaneous vs. sequential presentation (right operand presented 800 ms after the left operand). His design also reduced the visual familiarity of the Arabic format. Campbell's results showed that the interaction between format and size was similar for both conditions.

The most likely explanation for this result might be that numbers´ semantic processing (i.e., magnitude) involves processes that are format-specific. Despite some studies suggest that processing Arabic digits and number words can follow a non-semantic route that avoids accessing conceptual codes (e.g., Herrera & Massif, 2012), a different group of studies claim that Arabic digits and number words are differentially processed when the task requires some kind of semantic elaboration (e.g., Campbell, 1994; Campbell & Clark, 1992; Damian, 2004; Dehaene et al, 1993). In other words, accessing the numerical magnitude is more efficient with Arabic digits than with number words. For example, Damian (2004) found that when participants had to judge whether a number between one and nine was larger or smaller than five response times for numbers in Arabic digits format were faster than for numbers presented as number words. Likewise, Campbell et al. (Campbell, 1994; Campbell & Clark, 1992; Campbell, Kanz, & Xue, 1999) found that the word format led to errors numerically far from the correct answer relative to digit format. According to the interpretation by the authors, miscalculations with number words seem to be less fixed by the semantic distance, suggesting that the word format activates a weaker magnitude representation than the digit format. This explanation, coupled with the fact that magnitude discrimination becomes more difficult when increasing the numerical size (see Dehaene, 1992, 2001), would allow to interpret the above mentioned interaction between format and problem size in terms of a weaker magnitude activation by number words. Given that the representation of numerical magnitude may be activated in the context of an arithmetic word problem (Orrantia & Múñez, 2013), the effects of format on calculation that have been found in the present study could be explained by a weakest contribution of magnitude information in the word format.

That interpretation would also explain the findings regarding the variables operation and level of competence. The difference between formats for large problems regarding small problems was more evident for subtraction problems. This result might be explained in terms of magnitude processing since subtraction requires further magnitude manipulation than addition (Dehaene, Piazza, Pinel, & Cohen 2003). Although some subtraction facts (especially with small operands) can be solved by memory retrieval, many are solved by calculation procedures (Campbell & Xue, 2001; LeFevre, DeStefano, Penner-Wilger, & Daley, 2006; Seyler, Kirk, & Ashcraft, 2003; but see Fayol & Thevenot, 2012). Those procedures are semantically mediated, that is, they require some kind of magnitude manipulation. In this sense, the expected format effects in calculation would be larger for subtraction than for addition since accessing the magnitude representation is less effective in word format and, in addition, further magnitude manipulation is required for subtraction. However, our results were different from those found by Campbell and Alberts (2009), where the interaction between format and size produced similar results for both operations. One plausible explanation of such discrepancies relates to the experimental conditions of our study. Firstly, ours was a sequential presentation whereas Campbell and Alberts´ (2009) was simultaneous; and secondly, our presentation was blocked by format while Campbell and Alberts´ (2009) stimuli were not. As Damian (2004) notes, the differences between Arabic digits and number words may be obscure when stimuli are interspersed. And those differences may emerge when stimuli are blocked because participants can possibly adopt different response criteria for each condition. Especially, we would say, in a sequential presentation where participants have enough time to process the first operand before the second operand is displayed.

The explanation in terms of accessing the numerical magnitude could also explain the results found according to participants´ level of competence. Results showed, for the very first time, that the interaction between format and size was more evident for less competent participants (η^2_p = .32 vs. .14 for low-skilled and high -skilled participants respectively), and the format by size by operation interaction was only significant for low-skill participants. As suggested by response time data, one possible interpretation is that this group of participants makes greater use of computational procedures against retrieval. In this sense, it would be feasible to consider that the interaction format by size affected, on the one hand, memory facts retrieval for most competent participants; and, on the other hand, the use of procedures that require greater semantic processing for less competent participants. However, as Fayol and Thevenot (2012) suggest, it is also possible that participants with higher level of competence also used procedures, but in a more efficient manner than their counterparts. According to these authors, individuals who are more competent in arithmetic fluency can get to use "compacted" procedures that can be quickly and automatically implemented. In this regard, we would assume that these procedures require less semantic processing than those under conscious control, which are possibly developed by less competent individuals.

This explanation to interpret individual differences is also supported by studies showing that accessing the magnitude representation is directly related to individual differences in arithmetic, both in children (e.g., De Smedt, Verschaffel, & Ghesquière, 2009; Holloway & Ansari, 2009;

Vanbinst, Ghesquière, & De Smedt, 2012) and adults (e.g., Castronovo & Göbel, 2012; Lyons & Beilock, 2011; but see Sasanguie & Reynvoet, 2014). Thus, higher arithmetic competence relates to higher proficiency in accessing the magnitude representation through symbolic quantities. This fact is reflected, among others, by the "numerical distance effect" (Moyer & Landauer, 1967). This effect relates to the fact that discrimination between two numbers is faster when the distance between them increases (e.g., 2-3 vs. 2-8). Since two numbers that are numerically close have more overlap of the tuning curves compared to numbers that are numerically far apart, it is harder to discriminate them (Dehaene, 1997). The distance effect would reflect the activation of magnitude representation. A larger distance effect would indicate less accuracy in accessing numerical magnitude. In this context, Vanbinst et al. (2012) found that 9-year-old children with better access to the magnitude representation (i.e., a smaller numerical distance effect) performed faster during memory facts retrieval tasks and were more effective to implement calculation procedures. Castronovo and Göbel (2012) showed that adults with higher arithmetic fluency had a smaller numerical distance effect than those with lower fluency. In this sense, numerical magnitude processing skills would explain our findings respect to the variable level of competence. If accessing the numerical magnitude is less efficient with number words than with Arabic digits format, and subtraction requires further magnitude manipulation than addition, those with more difficulties in accessing numerical magnitude (i.e., less competent participants) would be more influenced by variables format and operation. However, this is a tentative hypothesis that needs further investigation since individual differences in accessing the magnitude representation were not directly measured in the present work.

In summary, the present study explored the effects of numerical format on calculus in the context of arithmetic word problems. Our findings are consistent with others that suggest that the format in which numbers are presented affects the calculation mechanisms and not just the encoding processes. Format effects can be explained by different mechanisms. In this study we assume an interpretation based on accessing the representation of numerical magnitude through symbolic numbers.

RESUMEN

Efectos del formato numérico en problemas aritméticos. Los adultos calculan más eficazmente cuando los operandos se presentan en formato arábigo (3 + 5) que cuando se presentan en formato palabras numéricas (tres + cinco). Una explicación ofrecida es la mayor familiaridad visual de los dígitos relativo a las palabras numéricas. Sin embargo, la mayoría de los estudios se han limitado a operaciones simples de cálculo con sumas y multiplicaciones. En el presente trabajo analizamos hasta qué punto se produce el efecto del formato en el contexto de un problema aritmético, en el que la familiaridad visual se elimina (Manuel tenía 3 canicas y le dieron 5). Participantes con diferente nivel de competencia en fluidez aritmética resolvieron problemas de suma y resta con los operandos en ambos formatos. Los resultados mostraron un efecto del formato, con mayor rapidez en formato dígitos que en palabras numéricas. Además los efectos fueron más evidentes en la operación de resta y en los participantes menos competentes en fluidez aritmética. Estos resultados fueron interpretados en función de una mayor eficacia del formato dígitos para acceder a la semántica del número.

REFERENCES

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review, 2*, 213–236.

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

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75–106.

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

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.

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

Campbell, J. I. D. (1999). The surface form x problem size interaction in cognitive arithmetic: evidence against an encoding locus. *Cognition, 70*, B25–B33.

Campbell, J. I. D., & Alberts, N. A. (2009). Operation-specific effects of numerical surface form on elementary calculation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 999–1011.

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

Campbell, J. I. D., & Clark, J. M. (1992). Cognitive number processing: An encoding-complex perspective. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 457–492). Amsterdam: Elsevier.

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

Campbell, J. I. D., & Fugelsang, J. (2001). Strategy choice for arithmetic verification: Effects of numerical surface form. *Cognition, 80*, B21–B30.

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., Kanz, C. L., & Xue, Q. (1999). Number processing in Chinese–English bilinguals. *Mathematical Cognition, 5*, 1–39.

Campbell, J. I. D., & Metcalfe, W. S. (2009). Numerical abstractness and elementary arithmetic. Commentary to Cohen Kadosh & Walsh: Numerical representation in the parietal lobes. *Behavioral and Brian Sciences, 32*, 313-373.

Campbell, J. I. D., Parker, H. R., & Doetzel, N. L. (2004). Interactive effects of numerical surface form and operand parity in cognitive arithmetic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 51–64.

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.

Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *Plos One, 7*, e33832.

Cohen Kadosh, R. (2008) Numerical representation: Abstract or non-abstract? *The Quarterly Journal of Experimental Psychology 61*, 1160–1168.

Cohen Kadosh, R., Henik, A. & Rubinsten, O. (2008) Are Arabic and verbal numbers processed in different ways? *Journal of Experimental Psychology: Learning, Memory and Cognition 34*, 1377–1391.

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brian Sciences, 32*, 313-373.

Colomé, A., Bafalluy, M. G., & Noël, M.-P. (2011). Getting to the source: a questionnaire on the learning and use of arithmetical operations. *Psicológica, 32*, 223-253.

Damian, M. F. (2004). Asymmetries in the processing of Arabic digits and number words. *Memory & Cognition, 32*, 164 –171.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 10*, 469–479.

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

Dehaene S (1997) *The number sense*. Oxford: Oxford University Press.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*, 16-36.

Dehaene, S., & Akhavein, R. (1995) Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition 21*, 314–326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*, 371–396.

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 circuits for number processing. *Cognitive Neuropsychology, 20*, 487– 506.

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

Hamann, M. S., & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction, 3*, 173–192.

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

Herrera, A., & Macizo, P. (2012). Semantic processing in the production of numerals across notations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 40–51.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*, 17–29.

Jackson, N., & Coney, J. (2007a). Simple arithmetic processing: Surface form effects in a priming task. *Acta Psychologica 125*, 1–19.

Jackson, N., & Coney, J. (2007b). Simple arithmetic processing: Individual differences in automaticity. *European Journal of Cognitive Psychology, 19*, 141-160.

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

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

LeFevre, J.-A., DeStefano, D., Penner-Wilger, M., & Daley, K. E. (2006). Selection of procedures in mental subtraction. *Canadian Journal of Experimental Psychology, 60*, 209–220.

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

Lyons, I. M., & Beilock, S. L. (2011) Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 12*, 256–261.

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

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

McCloskey, M., Macaruso, P., & Whetstone, T. (1992). The functional architecture of numerical processing mechanisms: defending the modular model. In J. I. D. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 493–538). Amsterdam: Elsevier.

Moyer, R. S., Landauer, T. K. (1967) Time required for judgements of numerical inequality. *Nature, 215*, 1519–1520.

Naccache, L. & Dehaene, S. (2001). Unconscious semantic priming extends to novel unseen stimuli. *Cognition 80*, 223–237.

Noël, M.-P., Fias, W., & Brysbaert, M. (1997). About the influence of the presentation format on arithmetical-fact retrieval processes. *Cognition, 63*, 335–374.

Orrantia J., & Múñez D. (2013). Arithmetic word problem solving: Evidence for a magnitude-based mental representation. *Memory & Cognition. 41*, 98–108.

Orrantia J., Múñez D., Vicente, S., Verschaffel, L., & Rosales, J. (2014). Processing of situational information in story problem texts. An analysis from on-line measures. *Spanish Journal of Psychology, 17,* e8, 1-14.

Orrantia, J., Rodríguez, L., Múñez, D., & Vicente, S. (2012). Inverse reference in subtraction performance: An analysis from arithmetic word problems. *The Quarterly Journal of Experimental Psychology, 65*, 725–738.

Orrantia, J., Rodríguez, L., & Vicente, S. (2010). Automatic activation of addition facts in arithmetic word problems. *The Quarterly Journal of Experimental Psychology, 63*, 310–319.

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

Sasanguie, D., & Reynvoet, B. (2014). Adults' arithmetic builds on fast and automatic processing of Arabic digits: Evidence from an audiovisual matching paradigm. *Plos One, 9*, e87739.

Schwarz, W. & Ischebeck, A. (2000) Sequential effects in number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1606–1621.

Seyler, D. J., Kirk, E. P., & Ashcraft, M. H. (2003). Elementary subtraction. Journal of Experimental Psychology: Learning, Memory and Cognition, 29, 1339–1352.

Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

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: Lawrence Erlbaum Associates, Inc.

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

Thevenot, C., Castel, C., Fanget, M., & Fayol, M. (2010). Mental subtraction in high and lower-skilled arithmetic problem solvers: Verbal report vs. operand-recognition paradigms. *Journal of Experimental Psychology: Learning, Memory & Cognition, 36*, 1242–1255.

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

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2012). Numerical magnitude representations and individual differences in children's arithmetic strategy use. *Mind, Brain, and Education, 6*, 129-136.

Zbrodoff, N. J., & Logan, G. D. (2005). What everyone finds: The problem size effect. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 331–346). New York: Psychology Press.

(Manuscript received: 13 June 2014; accepted: 27 January 2015)