About the influence of the presentation format on arithmetical-fact retrieval processes

Marie-Pascale Noël^{a,*}, Wim Fias^b, Marc Brysbaert^b

^a*Cognitive Neuropsychology Department, Catholic University of Louvain, Louvain-La-Neuve, Belgium*

^b*Department of Psychology, University of Leuven, Leuven, Belgium*

Received 31 August 1996, final version 17 March 1997

Abstract

This article presents the results of two experiments. In Experiment 1, French-speaking participants were asked first to retrieve the product of two numbers presented in Arabic or verbal code, and then to perform a number-matching task on the same material to assess the encoding time difference between numerals in the two formats. Experiment 2 involved the same multiplication task with Dutch-speaking participants who name two-digit numbers in reverse order. The format effects obtained by Campbell and Clark (1992); Campbell (1994) for multiplication were replicated. However, several observations suggest that some of these effects may be due to encoding time differences between word and digit numerals. The same size-by-format interaction was found for the number-matching task as for the multiplication task, and the effect disappeared with practice in the multiplication task. Finally, despite the fact that the linguistic structure of number names differs between French and Dutch, the types of error produced in both groups were identical. The last result does not match with the hypothesis that operand intrusion errors are due to interference between reading processes and arithmetical-fact retrieval processes. Implications of these findings for the debate about the nature of arithmetical-fact retrieval are discussed.

Key words: Calculation; Arithmetical facts; Presentation format; Numbers

1. Introduction

1.1. General overview

The realization of simple addition or multiplication problems up to $9 + 9$ or 9×9 has been studied extensively. This research has converged into a global

* Corresponding author. E-mail: noel@neco.ucl.ac.be

consensus according to which most humans from about 8–10 years old retrieve arithmetical-facts from memory (see Cornet et al., 1988; McCloskey et al., 1991b; Ashcraft, 1992 for recent reviews, but see LeFevre et al. (1996a), (1996b) for a larger emphasis on the use of procedural strategies rather than on simple retrieval from memory). So, the resolution of simple mathematical operations has become widely studied not only to describe the internal organization of these facts but, more importantly, as a way of approaching more general questions about the elementary memory processes and the type of representations underlying the retrieval. This debate about the nature and the organization of mental representations is an old one in cognitive (neuro)psychology. It has opposed those postulating a unique and abstract semantic system (e.g., Anderson, 1978; Caramazza et al., 1990) and those defending the existence of format and/or modality specific representations (e.g. Paivio, 1986; McCarthy and Warrington, 1988; Shallice, 1987; Kosslyn, 1983).

Within the literature of mathematical cognition, three main theoretical positions can be distinguished: the abstract model of McCloskey et al. (1985), the encoding-complex view of Campbell and Clark (1988) and the triple-code model of Dehaene (Dehaene, 1992; Dehaene and Cohen, 1995).

According to McCloskey et al. (McCloskey et al., 1985; McCloskey, 1992), arithmetical processing operates upon a single and abstract format that constitutes the sole support for internal representations of numerical quantities. In such an architecture, orally solving a simple arithmetic problem such as 3×7 presented in the Arabic code requires the following sequence of operations: (1) transformation of the Arabic numerals into an abstract representation, (2) interpretation of the arithmetical sign by a sub-component of the calculation system, (3) retrieval of the result of the operation in an abstract network, and (4) transformation of the result into an adequate verbal output form. In other words, arithmetical-fact retrieval operates on abstract representations and is totally independent of the format used for the presentation of the problem or the production of the response. The authors' main support for this assumption (Sokol et al., 1991) is the study of a brain-damaged patient, PS, who suffered from a deficit in the retrieval of arithmetical-facts but was unimpaired with respect to the peripheral stages of recognition and production of numerals in all formats. PS was tested on single-digit multiplications and the format of stimuli and responses was manipulated. The error distributions (number and type) did not differ between stimulus or response formats, and the patient tended to err on exactly the same problems for each condition (the correlations of the error rates for individual problems across formats ranged between 0.83 and 0.90). However, although this profile supports the hypothesis of a single format mediating the retrieval of arithmetical-facts, it does not address directly the additional claim that these representations are abstract in nature.

In contrast to the above model, the encoding complex model of Campbell and Clark (Campbell and Clark, 1988; Clark and Campbell, 1991) emphasizes the importance of the presentation and the production format for arithmetical-fact retrieval. More specifically, Campbell and Clark assume that arithmetical-fact retrieval is mediated by format-specific representations, which are thought to be numerous: verbal (i.e., articulatory, auditory, orthographic, and motor codes for

various spoken and written number words, as well as unique codes for special populations, such as sign language) and non-verbal (i.e., visual and motor codes for digits, imaginal and other analog codes for magnitude, and combined visual-motor representations for various number-related activities, such as finger counting or using an abacus). The model thus upholds that numerical processing in general, and arithmetical-fact retrieval in particular, “may differ as a function of the presentation format because Arabic and verbal inputs, e.g., potentially provide independent pathways to number-fact representations” (Campbell, 1994, p. 6–7).

In support of their model, Campbell and Clark (1992) presented results showing that the performance of normal participants in a simple multiplication task differed significantly as a function of the presentation format (Arabic or written verbal code). They reported a greater increase of RTs and error rates due to problem difficulty in the verbal than in the Arabic presentation condition. Moreover, a qualitative analysis of the structure of the errors led the authors to conclude that the errors were not the result of difficulties located at the encoding stages but that they originated from the level of the retrieval stage (see also below). More recently, Campbell (1994) has extended the data by contrasting the realization of simple additions and multiplications presented in words and in digits. In this case, he observed that the effects of the presentation format interacted with the arithmetical operation, a result that is more difficult to explain by a model supposing the intervention of a unique common code for the retrieval of arithmetical-facts.

Another version of the multiple code hypothesis is the triple-code model of Dehaene (1992). This model looks like an attempt to accommodate McCloskey, Caramazza and Basili’s postulate of modularity with Campbell and Clark’s assumption of multiple and specific internal codes. Dehaene posits the existence of three different internal codes for mentally representing numbers: an auditory verbal code, a visual Arabic code and an analog magnitude code. Communication between these representations is made through translation paths. The model also specifies which code is used for which numerical operations: the verbal code is used for verbal counting and arithmetic-fact retrieval, the Arabic code mediates multi-digit calculation and parity judgment, and the analog magnitude code is used for numerical comparison or quantity approximation. In support of his theoretical position, Dehaene has provided single-case evidence from brain-damaged patients (e.g., Dehaene and Cohen, 1991) and group data from normal participants (e.g., Dehaene et al., 1993). However, so far he has not published an attempt to validate his assumption that arithmetical-fact retrieval is mediated by verbal representations.

Although Dehaene postulates the existence of specific and multiple internal codes, it should be made clear that he does not predict different performance patterns as a function of the presentation format. In the case of arithmetical-fact retrieval, for instance, he assumes that every type of input is first translated into the verbal code before arithmetical-fact retrieval takes place. The predictions from Dehaene and McCloskey’s model are thus the same regarding the absence of format-specific influences.

There have been several other attempts to find out whether format manipulations

influence retrieval of arithmetical-facts. The first of these studies was by Gonzalez and Kolers (1982, 1987). They used a verification task with simple additions presented either in Arabic code ($2 + 3 = 5$), Roman code ($\text{II} + \text{III} = \text{V}$), or in various combinations of the two formats ($2 + \text{III} = \text{V}$). They not only showed that RTs varied as a function of the presentation format and two classic arithmetic structural variables (the MIN and the SUM), but also that these two variables interacted. Finally, they observed an effect of the position of the Roman numerals: these were more disturbing when they stood in an addend position than when they represented the sum. A similar pattern of results was obtained in another experiment where participants had first been trained to name Roman numerals rapidly. Therefore, Gonzalez and Kolers concluded that participants did not translate Arabic and Roman numerals into an abstract code but that the equation-verification mechanism acted on format-dependent representations. However, they did not explain how the format-dependent representations could result in the specific pattern of findings they obtained. Gonzalez and Kolers's interpretation has been questioned by Noël and Seron (1992) (see also Sokol et al., 1991), who showed that the pattern of RTs could be explained by a peculiarity of the Roman code which results in that the encoding time is not constant for all numerals but varies as a function of their structure, with numerals presenting an analog (I, II or III) or a symbolic structure (V, X) being processed faster than those with a complex structure (IV, VI, VII, IX, ...).

The effect of presentation format has also been studied by looking at the automatic activation of the sum of two numbers. LeFevre et al. (1988) asked participants to decide whether a target number (e.g., 5), shown immediately after a pair of numbers, had been presented in the initial pair (the numbers of the initial pair could be tied or not by an arithmetic symbol, e.g., 2 5, or $2 + 5$). In such tasks, it is usually observed that the time to reject a number is longer when this number equals the sum of the numbers in the pair. LeFevre et al. found that the interference effect was the same for all formats (words or digits) of the items, which is more in line with the idea of a single, code-independent representation.

Bilinguals have also been studied to test the different hypotheses. In Marsh and Maki (1976), addition problems were presented in Arabic format to bilinguals who were instructed to respond in their first or second language. The number of additions per problem varied from one (e.g., $3 + 4$) to three (e.g., $3 + 5 + 4 + 2$). A similar paradigm was applied by McClain and Shih Huang (1982), but they presented the additions orally in either the first or the second language, and the participants had to respond in the same language. In both studies, the participants responded more rapidly in the first than in the second language, but the slope of the function relating the RTs to the number of operations was identical. This suggests that the retrieval stage was the same in both language conditions. Unfortunately, these results are not very helpful from a theoretical point of view. They can equally well be interpreted in terms of the multiple-code or the central-abstract view (Marsh and Maki, 1976; Noël, 1991; McCloskey, 1992). As for the former, one might say that the participants first transcribed the problems from the non-preferred language to the preferred language and then performed the arithmetic operation in that language. Thus, the slower response times in the

second language are due to the supplementary translation stage. According to the central-abstract code view, all problems are initially transcoded into a language-independent representation. The observation that this takes longer in the second language than in the first, can then be attributed to the fact that the transcoding is less automatic and more time-consuming in the second than in the first language.

Frenck-Mestre and Vaid (1993) published two other experiments with bilinguals. In the first one, participants were asked to add mentally two visually presented numbers and to decide whether the sum equalled a subsequently presented probe. This probe was either the correct sum, or a number that differed from the sum by a small or a large split. The presentation format of the probe and the pair of numbers was manipulated (digits, first and second language), as well as the SOA between the presentation of the numbers and the probe. It was found that bilinguals were slower to respond to problems presented in the second language than in the first language, and were slower for problems presented in word format than in digit format. The effect of the split (i.e., longer rejection times for small splits than for large splits) was observed in each of the three presentation formats. Finally, responses to true-sum probes were faster than responses to non-sum probes in the digit and the first language conditions but not in the second language condition (it should be noted, however, that there was only a trend for an interaction between type of probe and presentation format: $p < 0.08$).

Virtually the same procedure was used in the second experiment of Frenck-Mestre and Vaid (1993), except that this time there were both addition and multiplication problems and that the probe types were different. The proposed solution could either be the true answer to the problem—*true probes*—the correct solution for the other operation (i.e., the sum of the addends in case of a multiplication problem and the product of the addends in case of an addition problem)—*associatively related probes*—or mathematically unrelated to the problem—*neutral probes*. In addition to the main effect of presentation format, the authors found that the associative confusion effect (i.e., the difference between the rejection of associatively related vs. neutral probes) was present in the digit and the first language condition but not in the second language condition. However, the authors recognized that the overall longer processing times in the second language condition may have masked the appearance of the associative confusion effect, which is short-lived. This last remark illustrates how cautious one must be in the interpretation of results from bilinguals. In their present state, these data do not seem to be very conclusive.

In summary, the really strong evidence against the hypothesis of a single representation underlying the retrieval of arithmetical-facts comes from the studies published by Campbell and Clark (1992); Campbell (1994). Because these data are so important to the debate, they will be the basis of the work presented here. However, before we explain the precise aim of our research, a detailed description of Campbell's findings and the criticisms raised against them will be given.

1.2. Campbell and Clark's experiments

As stated previously, Campbell and Clark (1992) found that performance was

worse (i.e., slower RTs and higher error rates) for multiplications presented in words than in digits. The disadvantage increased with the size of the problem, and the errors in the word problems tended to be more distant from the correct response than the errors in the digit problems. On the basis of these findings, Campbell and Clark argued that the Arabic-digit format was more efficient to activate number-fact representations than the word format because arithmetical problems are more often encountered in digits than in words. In contrast to this view, McCloskey et al. (1992) claimed that the format effects could also be due to encoding differences between words and digits. They argued that words need more encoding time than digits and that this difference increases with number magnitude because the function between frequency and magnitude has a steeper slope in the word format than in the digit format (see Dehaene et al., 1990). Accordingly, if people adopt a response deadline that limits the time interval before responding, then the remaining retrieval time is shorter for word problems than for digit problems, and the difference between both formats increases with the magnitude of the numbers involved in the problem (which is highly correlated with the size of the problem). The hypothesis of a reduced time for the retrieval stage thus predicts higher error rates for large-size problems and can also explain the type of errors produced by assuming that short time intervals promote errors more distant from the expected answer. In support of their interpretation, McCloskey et al. (1992) ran a regression analysis on the RT-differences between the word and digit formats with the number of characters occupied by the problem and the frequency of each addend as predictor variables. They showed that in this analysis the problem-size effect and the fan effect disappeared (the fan effect refers to the fact that problems which have the solution shared by other problems produce longer RTs).

Subsequently, Campbell (1994) contrasted addition and multiplication problems in word and digit formats and found different format $\times$ problem size interactions for both operations. At first, it seems difficult to account for this result in terms of operand encoding difficulties as the same operands were used for both operations. Yet, here too it is possible to propose an interpretation that does not require an influence of the presentation format on the arithmetical-fact retrieval stage itself. Such an explanation is actually provided by Campbell (1994) himself: “If encoding of word format problems uses more attentional resources compared to the digit format, then processing of the more attention-demanding larger problems could be disrupted more by the word format relative to smaller-number problems. The form of this format $\times$ problem-size interaction could differ for addition and multiplication if effects of problem size vary as a function of operation. For example, if larger addition problems were more likely to involve attention-demanding procedural strategies relative to multiplication (e.g., $9 + 5 = 10 + 5 - 1 = 14$), then larger addition problems presented in word format might be especially susceptible to interference due to reduced attentional resources” (p. 34). In other words, “the observed interactions of format, operation, and problem size, while not predicted by the abstract-modular model, do not necessarily demand the conclusion that arithmetic processing varied with format” (p. 34).

Beside the RTs analyses, Campbell also studied the errors produced. First, he

analysed positive *inter-trial error priming*, i.e., the fact that answers to previously tested problems have a greater-than-chance probability of appearing as the error response to subsequent problems (Campbell and Clark, 1989). In his study, Campbell (1994) observed a stronger positive error priming between problems presented in the same format than between problems presented in different formats (in his experiment digit trials alternated with word trials). According to Campbell, positive error priming requires explicit retrieval of a previous answer. This assumption is supported by two empirical facts. First, when alternating multiplication trials with product-naming trials, positive error priming is only observed in connection with previously retrieved answers (Campbell, 1991). Second, when alternating production and verification trials in simple multiplication, Campbell and Tarling (1996) also observed a much higher positive error priming in production errors by previous production trials than by previous (true) verification trials. From these results, Campbell concludes that positive error priming arises from number-fact retrieval processes, as opposed to more peripheral encoding or production processes. Consequently, the stronger positive error priming between same-format problems relative to different-format problems is interpreted as supporting the idea of format specific number-fact retrieval processes. A weak aspect of this reasoning, however, is that we do not know precisely what provokes the positive error priming effect. Indeed, Campbell's research has shown that the positive error priming effect is sensitive to every type of modification induced so far. In order to settle whether the effect is really specific to the number-fact retrieval process, and not due to some perceptive priming, it is necessary to show that the effect is insensitive to modifications which are not supposed to affect the retrieval process. For instance, one should show that the priming is not influenced by physical changes such as the type of font used (Times or New York, bold or italic, lower case or upper case, . . .) or by the spatial presentation of the items (horizontal vs. vertical) and so on. As long as these types of control have not been done, we cannot be confident about the real meaning of the positive error priming effect, and strong conclusions on the basis of this phenomenon seem hazardous.

The study of the inter-trial priming effect was followed by an analysis of the types of error produced by the participants. In this respect, the most challenging finding for McCloskey's model was the observation of a larger proportion of *naming* and *operand-intrusion* errors in the word problems than in the digit problems (in both Campbell and Clark, 1992; Campbell, 1994). Naming errors refer to cases where the entire incorrect response corresponds to one or both of the problem's operands (e.g., $4 \times 8 = 8$, 4, 48 or 84). Operand-intrusion errors are errors in which only one of the problem's operands intrudes in the response (e.g., $4 \times 8 = 42$, 28, . . .). According to Campbell, these errors are an indication of interactions between arithmetical-fact retrieval and reading processes. Indeed, "about 80% of the operand intrusions also were arithmetically related products (e.g., $7 \times 4 = 24$), as opposed to unrelated or miscellaneous responses (e.g., $7 \times 4 = 34$) which suggests that they resulted from priming or activation of number-fact representations" (Campbell, 1994, p. 8). Furthermore, these errors were not only sensitive to the presentation format but also to the position of the

intruder in the original problem: More often than expected by chance the intruder retained its position (e.g., $6 \times 9 = 63$ was more frequent than $9 \times 6 = 63$). A related phenomenon was reported for the correct response trials: Fewer errors and shorter RTs were observed when one of the operands matched the corresponding digit in the answer (e.g., $6 \times 4 = 24$) than when there was no such positional match (e.g., $4 \times 6 = 24$).

According to Campbell and Clark (1992) and Campbell (1992), this reveals an interaction between reading and number-fact retrieval processes. Faced with a multiplication problem, people would not only activate the corresponding product but also, unintentionally, reading-based associations with the left operand encoded as the decade number and the right operand encoded as the unit number (e.g., 6×9 activates 'sixty-nine'). Since words are often encountered in a reading context, word problems would activate these reading-based associations more strongly than digit problems, which explains the difference in rates of naming errors and intrusions errors between both presentation formats.

The fact that a central process such as arithmetical-fact retrieval is affected by the format and the position of the presented items seems difficult to reconcile with McCloskey's amodal perspective. Yet, McCloskey et al. (1992) argue that, although not directly predicted by their model, the word/digit difference in the rate of operand intrusions can nonetheless be explained within their theoretical perspective. They agree with Campbell and Clark that intrusion errors reflect an interference at the arithmetical-fact retrieval stage. Yet, they assume that this interference applies to semantic representations rather than to verbal internal codes. More specifically, they speculate that faced with a problem, "numeral comprehension processes not only encode the operands separately, but also attempt to treat the entire problem as a single numeral (7×6 or seven $\times$ six would generate a representation of $\{7\}10\text{exp}0$ and $\{6\}10\text{exp}0$ but also of $\{7\}10\text{exp}1, \{6\}10\text{exp}0$). This numeral representation may then interfere with the arithmetical-fact retrieval process, leading to activation of an answer representation with which it shares one or both quantities (especially in matching positions). (...) The strong tendency for operand intrusions to be table-related errors may be explained by the fact that an incorrect answer receiving activation from both the problem encoding and the interfering numeral encoding is especially likely to occur as an error". Furthermore, these authors explain the format difference by proposing that "word problems may have appeared more numeral-like than the digit problems (...). Thus the tendency to encode the problem as a single numeral may have been greater for the word format than for the digit format". (McCloskey et al., 1992, p. 527–528). The explanations provided by Campbell and Clark on the one hand and by McCloskey et al. on the other hand are thus amazingly similar, although they refer to different processes (reading vs. number comprehension) acting upon different internal codes (verbal vs. abstract).

In summary, the strongest evidence against the assumption that the arithmetical-fact retrieval process relies on abstract representations comes from Campbell and Clark (1992); Campbell (1994). Three major types of presentation format effects have been described: (1) the size effect is more important in word than in digit

problems; (2) naming errors and operand intrusion errors are more frequently produced in response to word than to digit problems; and (3) positive error priming is stronger within format than between formats. We have argued that the importance of the last finding is not clear given that the processes underlying the positive error priming effect are poorly understood. The interactions between presentation format and problem-size have been criticized by McCloskey et al. (1992) who proposed an alternative interpretation based on encoding peculiarities of the formats. However, this explanation is clearly ad hoc and cannot be considered as a direct prediction of the modular-amodal model. Therefore, we replicated Campbell and Clark (1992) and examined the possible effects of encoding factors on the interaction between format and problem-size (Experiment 1). In addition, we tested Campbell's interpretation of operand intrusions in table-related errors by looking at a language that reverses the position of units and tens in the number names (Experiment 2).

EXPERIMENT 1: REPLICATION AND CONTROL OF THE ENCODING TIME

This experiment was designed in the first place to see whether the format effects reported by Campbell and Clark (1992) could be replicated. Indeed, it should be mentioned that although the differences between word and digit formats were statistically significant, they were very small. One might even say that the most striking result of the original article was the very high similarity of response patterns in both conditions (this was true for error types and rates, as well as for RTs). The authors themselves noted that “the pattern of RTs across the 64 problems was very similar for digit and word formats ($r = 0.913$)” (Campbell and Clark, 1992, p. 464).

The second objective of this study was to verify experimentally the criticism of McCloskey et al. (1992) that most of the format effects are due to encoding processes: “word problems generally took longer to encode than digit problems and the magnitude of this encoding disadvantage for the word format varied across problems as a function of several (encoding) variables” (McCloskey et al., 1992, p. 511). This hypothesis was already considered by Campbell and Clark (1992) who recognized that only for word problems was there a length difference between small problems (i.e., with addends from 2 to 5, the mean number of characters = 4) and large problems (i.e., with addends from 6 to 9, the mean number of characters = 4.25). Yet, in a multiple regression analysis, they showed that the length of the problem contributed to, but did not explain all of the RT differences between word and digit problems (i.e., the size effect remained significant after controlling the length). However, McCloskey et al. (1992) responded that additionally introducing the frequency of the numerals in the regression analysis led to the disappearance of the size and fan effects (see above).

We do not intend here to replicate McCloskey et al.'s analyses, because trying to control all the variables which may play a role in the encoding processes is a very risky venture. One is never sure that all factors have been considered, nor that

their measure is valid enough. Our way of dealing with the problem of discrepancies in encoding time is different: in addition to the multiplication task, we ask participants to realize a task in which the same pairs of digits and words have to be identified without the requirement of a mathematical operation. If the format effects obtained in the multiplication task are due to encoding factors, then the same effects should be found for this second encoding task.

The critical point is of course, the choice of a good encoding task. In his paper, Campbell (1994) tried to address the problem of differences in encoding time by running a number naming task in which participants had to read all the numbers (words and digits) from 2 to 9. According to Campbell (1994), this task “provided an estimate of differences in encoding efficiency between individual digits and individual words” (p. 14). We agree that the results of this task may reflect the encoding times of individual digits and words, but we fear that they cannot be extrapolated to the encoding times of pairs of digits or words. Several experiments have shown that the differences in processing time between single digits and single number-words are very small (and often non-significant) whereas much larger differences are obtained when two stimuli have to be processed on a single trial (for a review, see Noël, 1991). This is probably partly due to the fact that digit pairs can be encoded in parallel whereas word pairs must be processed sequentially (Noël, 1991; McCloskey et al., 1992), and to the fact that the difference in length between the word and the digit format is larger for pairs than for single numerals (McCloskey et al., 1992).

To measure the encoding times for pairs of numerals presented in words or in digits, we used a number-matching task. Basically, participants were presented a pair of dot patterns followed, after two seconds, by a pair of digit or word numerals. The participant’s task was to decide as quickly as possible whether the quantities represented by the pair of digit or word numerals were the same as those represented by the dot patterns. How adequate is this task? To answer that question, a description of the sub-processes underlying the task should be given. However, such a description is necessarily influenced by the theoretical framework of reference.

According to the amodal theory of McCloskey et al. (1985), the pair of dots would activate the corresponding amodal semantic representations (through a still unspecified comprehension process) and these representations would be kept in memory. The presentation of the pair of digits or words would then be encoded (through the number comprehension systems) and activate the corresponding semantic representations. Finally, these semantic representations would be compared to those corresponding to the dot patterns. Since the comparison process is done on abstract representations, it is not influenced by the presentation format of the stimuli, and the difference between word and digit pairs should be a measure of their encoding time difference.

The situation is slightly more complex when the triple-code model of Dehaene (1992) is considered. In this perspective, one should draw a distinction between pairs of dots which are, or which are not, subitizable: the former would directly activate an analog representation of the corresponding quantity while the latter

would give rise to an activation of the verbal representations through counting procedures. Pairs of word or digit numerals would activate the corresponding verbal and visual-Arabic representations. The next step would consist in the matching judgment. If this step involves a true comparison of quantities, then it is supposed to be realized on analog representations. In that case, the verbal or visual-Arabic representations generated from the pair of numerals should be converted into the corresponding analog magnitude, and a similar transformation would be expected for the non-subitizable dot patterns (from the verbal representations into the corresponding analog representations). Accordingly, the difference between word and digit pairs in this task would correspond to differences in encoding time (i.e., activation of the corresponding verbal or visual-Arabic representations) plus the time needed for the activation of the corresponding analog representations. However, such a difference may not exactly correspond to the difference observed in the multiplication task since in this latter task, the difference between the two formats would correspond to the same encoding times plus the time needed for the activation of the corresponding verbal representations (when necessary, i.e., for digit problems only).

On the other hand, one could consider that the number-matching task is not based on a true comparison of quantities but rather on a comparison of verbal representations. Indeed, it has been repeatedly shown that short-term memory tasks with numbers as stimuli largely make use of auditory verbal codes (Pynte, 1974; Gielen et al., 1991, Experiment 1; Brysbaert, 1995, Experiment 2), which is in line with the idea that short-term memory makes use of an articulatory loop for rehearsal (Baddeley, 1986). This can be deduced from the finding that silent-reading times of (Arabic) numerals are linearly related to the syllable length of the number names. Assuming that phonological images in short-term memory can be matched on non-numerical grounds (as needed for words which are not numbers), it is quite feasible that the number-matching task addresses the verbal representations rather than the analog representations. In that case, the analog representations triggered by the dots would activate the corresponding verbal representations to enable short-term memory storage, and these verbal codes would be used for the matching judgment. If the above reasoning is correct, then the encoding task would be very similar to the multiplication task, as both would be based on verbal representations (either due to short-term memory storage or because arithmetical-fact retrieval is based on verbal codes).

In Campbell and Clark's theory, the comparison process is assumed to be realized on specific internal codes which might differ as a function of the presentation format of the stimuli (words or digits). The difference obtained in this task between a pair of digits and a pair of words would thus not only reflect encoding difference but also the difference between an unspecified comparison process realized on probably different internal codes.

In summary, whether the number-matching task can be considered as an adequate way of measuring the encoding time difference between pairs of digits and of words depends on the theoretical reference used. However, it should be stressed that the very different assumptions of the three current models probably

do not allow to find a task which is unanimously considered as a good way of evaluating encoding processes. It is our conviction that the task we devised is the closest we could generate. At least, it allows a straightforward interpretation within McCloskey's theory, which is what mostly matters for the purpose of this experiment. Indeed, Experiment 1 was designed to test the alternative account of Campbell's findings by McCloskey et al. (1992), and to see whether encoding time differences between the two presentation formats (as defined by McCloskey's theory) could explain the format effects obtained in a multiplication task.

2. Method

Participants: 24 students at the Psychology Department of the Catholic University of Louvain participated in the experiment for course credit. They were all native French speakers with normal or corrected-to-normal vision. There were 5 males and 19 females; ages ranged from 18 to 28 ($M = 20.7$, $SD = 2.3$).

3. Design and apparatus

All participants attended two one-hour sessions on two different days (usually within a one-week interval). The experiment took place in a quiet small room in the presence of the experimenter. Stimuli appeared on the screen of a Macintosh Performa, which was placed approximately 50 cm in front of the participant. Each session started with the multiplication task followed by the number-matching task. Half of the participants started the first multiplication session with the Arabic format and the second session with the word format; the other half were given the reverse order. For each of these sub-groups, half of the participants pressed the left key when there was a match in the number-matching task and the right key when there was a mismatch, and vice versa for the other half.

3.1. Stimuli and procedure for the multiplication task

Fifty-six multiplication problems were used: all the problems from 2×2 through 9×9 , but not the 'ties' (2×2 , 3×3 , ... 9×9). Problem operands appeared as digits or as lower-case French number words. Stimuli were displayed horizontally at the centre of the screen using the Times 24-point bold font. They were typed in black on a white font screen. Digit problems occupied 5 character spaces or approximately 2 cm (blank spaces included, e.g., ' 3×5 '). The length of word problems varied from 10 (e.g., 'six $\times$ huit') to 14 characters (e.g., 'trois $\times$ quatre') or 4.5 to 5 cm.

Participants were given 8×56 multiplication problems in digit format and 8×56 problems in word format for a total of 896 trials. Presentation format alternated every two sets of 56 problems with a resting period in-between. Participants were always informed about the presentation format of the next series

of stimuli. Within each block, problems were presented randomly, with an order that differed from one participant to the other. Each session started with 10 training trials (i.e., 5 problems written in digits and 5 presented in words) and lasted for about half an hour.

Trials began with a fixation point (*) appearing for 1 s in the center of the screen and directly followed by the problem which remained visible until the participant responded. The inter-stimulus interval was 1 s. No feedback about accuracy or speed was given.

Participants wore a lapel microphone that triggered a voice-activated relay. Measurement of the response time started from the appearance of the problem on the screen and was accurate to 1 ms. The experimenter recorded on a sheet of paper the answers produced by the participant as well as failures of the voice-activated relay.

Instructions emphasized both speed and accuracy; however, participants were told that under speed pressure conditions, occasional errors could be expected.

3.2. Stimuli and procedure for the number-matching task

In this task, participants were first shown two sets of dark dots presented in a canonical pattern (see Fig. 1 for a few examples). The number of dots presented to the left and to the right were never equal and varied from 2 to 9. Each dot had a diameter of 0.5 cm. The full stimulus comprising the two sets of dots was about 11 cm wide which allowed a space of about 3 cm between both sets. The dot stimuli

Fig. 1. Examples of pairs of dot arrays.

were presented in the centre of the screen for 2 s¹ and were then replaced by either a pair of digits or a pair of number words separated by a '§' symbol flanked by a blank space on each side (e.g., 'trois § deux'). The numbers were written in the same font as the one used for the multiplication problems (i.e., Times 24 points bold font, typed in black on a white screen).

The participant's task was to decide whether or not the two digits or the two words expressed the same numerosities as the ones represented by the dots in the first stimulus. Two keys of the number pad (the '1' and the '2') were used for the responses.

The number pairs of the number-matching task corresponded to those used in the multiplication task. Each of the 56 pairs of numbers was presented four times in digits and four times in words, with half of the trials in each format corresponding to a match. In case of a mismatch, one of the numbers (randomly left or right) differed from the corresponding dots by ± 1 unit (e.g., pair of 5 and 3 dots followed by '4 § 3' or 'cinq § deux'). In each of the experimental sessions, participants were presented with the 56 problems presented once in digits and once in words, once as a match and once as a mismatch (i.e., $56 \times 4 = 224$ items per session; 448 for the whole task). Items were presented in a pseudo-random order which respected the three following criteria: (1) no four successive pairs with a match or a mismatch, (2) no four successive pairs presented in the same format (digits or words), and (3) never two pairs in a row with the same two numerosities, no matter what the presentation order or format was (e.g., never '2 § 3' after 'trois § deux').

Each trial started with a 1 s attention signal (!) followed 500 ms later by the presentation of the pair of dots. This stimulus stayed on the screen for 2 s and was then replaced by the pair of numerals until the participant responded. The ISI was fixed at 1 s. A whole session lasted for about 15 min and was preceded by a set of 10 training trials.

Participants were instructed to make their decision as quickly and accurately as possible.

4. Results

4.1. Multiplication task²

4.1.1. Error rates

A total of 460 errors were produced in response to word problems ($M = 19.17$,

¹ Orally producing the number of dots presented in a canonical order takes about 1300 ms for large sets of 9 dots (Mandler and Shebo, 1982, Experiment 4) but the simple production time of a verbal numeral in the usual sequential order averages 300 ms (Mandler and Shebo, 1982, Experiment 1). This suggests that 1 s is a sufficient delay to perceive the quantity of one large set of 9 dots. Since two sets of dots appeared simultaneously, we opted for a presentation delay of 2 s.

² Due to a technical problem, the problem 'neuf $\times$ quatre' has not been presented during the experiment. Consequently, the data for the corresponding problem in the Arabic format (9×4) have not been considered in the analyses.

$SD = 11.6$, 4.35%) and 370 to digit problems ($M = 15.42$, $SD = 11.2$, 3.5%). In order to see whether the tendency to make an error was influenced by the size of the problem, we used the same criteria as Campbell (1994): problems were classified as small if both operands were smaller than or equal to 5. Large problems were all the other cases. Our set of 56 pairs thus involved 12 items from the small group and 44 from the large one.

A 2×2 ANOVA on the percentage of errors with format and problem size as repeated measures revealed that more errors were made in response to word problems than to digit problems ($F(1,23) = 6.73$, $p < 0.02$, $MSe = 1.25$) and more on large problems than on small ones ($F(1,23) = 50.81$, $p < 0.0001$, $MSe = 6.74$; with $M = 0.97\%$ and 4.75% for small and large problems, respectively). More importantly, although the interaction of format and size only approached significance ($F(1,23) = 3.82$, $p = 0.063$, $MSe = 1.34$, see Table 2), a paired t -test indicated a reliable difference between both formats for large problems ($t(23) = 7.28$, $p < 0.0001$) but not for small ones ($t(23) = 0.48$).

4.1.2. Error types

Following the classification proposed by McCloskey et al. (1991a) and extended by Campbell and Clark (1992), a distinction was made between the following types of errors:

1. *Operation error*: when the incorrect response corresponds to the correct response of another arithmetical operation (e.g., $4 \times 8 = 12$).
2. *Naming error*: when the incorrect response corresponds either to the name of one of the operands (e.g., $4 \times 8 = 8$) or to the name of the number resulting from the combination of the two operands (e.g., $4 \times 8 = 48$).
3. *Table-related error*: when the response belongs to the table of one of the operands (e.g., $4 \times 8 = 36$).
4. *Table-unrelated error*: when the response belongs to the single-digit multiplication table but not to the tables of the operands (e.g., $4 \times 8 = 35$).
5. The remaining responses which could not be classified into any of the preceding categories were labelled '*miscellaneous*' (e.g., $4 \times 8 = 52$).

Furthermore, within each of the last three categories, a supplementary distinction was made between errors which did, or did not, involve the intrusion of one of the problem's operands (e.g. $4 \times 6 = 36$ vs. $4 \times 6 = 30$ for the table-related errors). Thus, we had 8 categories of errors: (1) operation, (2) naming, (3) table-related with an intrusion, (4) table-related without intrusion, (5) table-unrelated with an intrusion, (6) table-unrelated without an intrusion, (7) miscellaneous with an intrusion, and (8) miscellaneous without intrusion.

Within the data collected, no operation error was observed. Furthermore, all but one of the naming errors could equally well be considered as table-related with intrusion errors (e.g., $3 \times 5 = 35$ or $4 \times 8 = 48$, $4 \times 2 = 24$). Yet, for the subsequent analyses, these were considered as belonging to the naming category.

To compare the error distributions in the two formats (see Table 1), paired t -tests were calculated for each error type on the proportions of errors for words

Table 1

Error distribution in the French-speaking sample as a function of the presentation format (number and mean of the proportions)

Type of error	Word condition	Digit condition	Two-tailed $t_{(23)}$	p
Naming	19 (0.041)	6 (0.011)	2.39	0.025
Table-related and intrusion	202 (0.450)	126 (0.335)	2.31	0.030
Table-related	161 (0.312)	167 (0.435)	– 2.47	0.021
Table-unrelated and intrusion	20 (0.046)	8 (0.043)	0.09	NS
Table-unrelated	8 (0.017)	12 (0.048)	1.30	NS
Miscellaneous and intrusion	21 (0.066)	22 (0.055)	0.50	NS
Miscellaneous	28 (0.068)	29 (0.071)	– 0.09	NS
Total	460 (1)	370 (1)		

and digits. Significant differences between the presentation conditions were obtained for three types of errors: naming, table-related with intrusions and table-related without intrusions. Naming errors were produced more frequently in response to word problems than to digit problems. The same was true for table-related errors with intrusion, but not for table-related errors without intrusion, which returned the opposite pattern of findings. Overall, the proportion of table-related errors was very similar for both formats ($t(23) = -0.197$, NS). It should also be mentioned that the interaction between format and intrusion was only present for table-related errors and not for table-unrelated and miscellaneous errors. Finally, the larger proportion of miscellaneous errors in response to word problems reported by Campbell and Clark was not replicated here (even after combining miscellaneous errors with and without intrusion, $t(23) = 0.13$, NS).

Finally, a measure of the distance of the errors produced in the two formats was calculated. Campbell and Clark (1992); Campbell (1994) had observed that table-related errors tended to be closer to the expected answer when they were produced in response to a digit problem than to a word problem. To see whether our data replicated this effect, an ANOVA with format and distance (± 1 operand vs. ± 2 or more operands³) as repeated measures was computed on the number of table-related errors. More errors were produced in response to the word problems ($F(1,23) = 5.68$, $p < 0.026$, $MSe = 9.25$) and these errors tended to be close to the correct answer ($F(1,23) = 5.23$, $p < 0.032$, $MSe = 61.05$). The interaction of format and distance was only marginally significant ($F(1,23) = 3.46$, $p < 0.076$, $MSe = 9.11$). Despite this weak tendency, paired t -tests were calculated and revealed that the number of close errors did not differ between the formats ($M = 8.5$ and $M = 8.8$ in the digit and word condition, $t(23) = 0.39$, NS) whereas the number of more distant errors was less important in the digit condition than in

³ When an error was related to both the first and the second operand, the distance was calculated on the basis of the larger operand so as to minimize it.

the word condition ($M = 3.7$ and $M = 6.3$, $t(23) = 2.96$, $p = 0.007$). In other words, we replicated the effect observed by Campbell. However, it is possible that we were confronted with a spurious effect. Indeed, given that word problems gave rise to more intrusion errors, it is possible that this type of error tended to be more distant from the correct response than errors without intrusions. Similarly, we have shown that the difference in the number of errors between the two formats was particularly important for large problems. Here too, one could imagine that errors to small problems tend to be closer to the target whereas errors to large problems would be more distant. Such a state of affairs would then predict more distant errors in response to word problems than to digit problems. The subsequent analyses aimed at testing these hypotheses.

First, an analysis with format, distance (± 1 vs. ± 2 or more), and intrusion (present or absent) was calculated on the number of errors. The interaction between these three variables was not significant.

Second, as the difference between the two formats is known to increase with the size of the problem, we wanted to know whether this could affect the format-by-distance interaction observed previously. Unfortunately, the error rate for small problems was not large enough for a statistical analysis, but an interaction between format and distance seemed unlikely given that 5/18 errors produced in response to small digit problems were close to the target versus 6/20 errors in the word condition. For the large problems, the format-by-distance interaction was not very different from the one obtained in the global analysis (as might be expected from the low number of errors in the small problems). There was a main effect of format ($F(1,23) = 5.04$, $p < 0.035$, $MSe = 9.56$) and of distance ($F(1,23) = 6.07$, $p < 0.022$, $MSe = 61.95$), together with a marginal interaction between these two factors ($F(1,23) = 3.76$, $p < 0.065$, $MSe = 8.07$). A t -test indicated a format difference in distant errors ($t(23) = 2.99$, $p = 0.007$, $M = 5.75$ and $M = 3.21$) but not in close errors ($t(23) = 0.74$, $M = 8.58$ and $M = 8.29$, respectively). Thus, the effect of format on the distance of table-related errors was true for large problems only.

Another phenomenon we wanted to replicate was the tendency for intrusion errors to keep the position they occupied in the original multiplication format. To test this, a $2 \times 2 \times 2$ ANOVA with repeated measures was run on the number of table-related errors with intrusion. The three repeated measures were: format, positional match (same position of the intruder in the original problem and in the answer, or different position), and the lexical class of the response to which the intruder belonged (decade or unit). As expected, the format effect was significant ($F(1,23) = 17.2$, $p < 0.0001$, $MSe = 1.75$) showing more errors in response to word problems ($M = 8.42$) than to digit problems ($M = 5.25$). In addition, intrusions were more frequent in the unit position than in the decade position ($F(1,23) = 20.67$, $p < 0.0001$, $MSe = 11.11$; with $M = 11.21$ and 2.46 , respectively) and they tended to respect the position they occupied in the original problem ($F(1,23) = 10.97$, $p = 0.003$, $MSe = 2.07$; $M = 8.21$ and $M = 5.46$, respectively). Finally, the interaction of format and lexical class was significant ($F(1,23) = 15.0$, $p < 0.001$, $MSe = 1.9$): Intrusions in the unit position were more frequent in words

than in digits (respectively, $M = 7.17$ and $M = 4.04$) but a similar number of intrusions appeared in the decade position for both formats ($M = 1.25$ for words and $M = 1.21$ for digits).

4.1.3. Reaction times

Median reaction times were calculated for each participant on the correct responses. Due to the malfunctioning of the voice-activated relay, 235 additional trials in the digit condition and 283 in the word condition were unreliable. So, a total of 7.0% of the RTs in the word condition and 5.9% in the digit condition had to be disregarded. As observed by Campbell and Clark (1992), the RTs measured for each problem were highly correlated between the two presentation formats ($r = 0.983$).

A first ANOVA was calculated with presentation format (digit/word) and size (small/large) of the problem as repeated measures (see Table 2). This analysis revealed faster responses for digit problems than for word problems ($F(1.23) = 141.79$, $p < 0.00001$, $MSe = 17479.08$; with, respectively, $M = 970$ ms and $M = 1291$ ms), as well as shorter RTs for small problems than for large ones ($F(1.23) = 65.38$, $p < 0.00001$, $MSe = 32968.67$; $M = 980$ ms for small problems and $M = 1280$ ms for large ones). More importantly, the interaction between these two factors was significant ($F(1.23) = 9.53$, $p = 0.0052$, $MSe = 1498.74$) and indicated a larger size effect in the word condition than in the digit condition.

Let us note, however, that if the absolute RT difference between small and large problems is higher in the word condition than in the digit condition, the reverse tends to be true if one considers the percentage of decrease between small and large problems in each format. Indeed, a mean decrease of 24.1% is measured in the word condition while the decrease observed in the digit condition is 26.8% (two-tailed paired t -test: $t(23) = 1.97$, $p = 0.061$).

In summary, we were able to replicate Campbell's results. More errors were observed in response to word problems than to digit problems, and the difference increased for larger multiplications. The type of error produced also differed as a function of the presentation format with more naming errors and table-related errors with intrusion in response to word problems. Furthermore, in the latter type of error, the intruding operand tended to keep the position it held in the original problem. Finally, table-related errors produced in response to large digit problems were closer to the target than those produced in response to large word problems.

Table 2
Performance (RTs and percentage of errors) as a function of format and size of the problem in the French-speaking sample

	RTs (ms)			% of errors		
	Large	Small	L-S	Large	Small	L-S
Words	1453	1129	324	5.28	1.04	4.24
Digits	1108	832	276	4.23	0.91	3.32
W-D	345	297		1.05	0.13	

Regarding RTs, we were able to replicate the main finding obtained by Campbell, i.e., an interaction between the presentation format of the problem and the size effect. However, such a difference in the size effect between the two formats did not hold when comparing the percentage of decrease between small and large problems in each presentation condition.

4.2. Number-matching task

4.2.1. Errors

A total of 3.1% errors was produced in this task. These were similarly distributed among the two presentation formats ($M = 3.35\%$ for digits and $M = 3.62\%$ for words) and among the match and mismatch trials ($M = 3.35\%$ and 3.62% , respectively). The small rates do not allow statistical analysis.

4.2.2. Reaction times

A first ANOVA was run on the median response times for the correct responses with format (digit/word) and type of trials (match/mismatch) as within-subject variables and response key as between-subjects variable. Match trials were verified more rapidly than mismatch trials ($F(1,22) = 54.68$, $p < 0.0001$, $MSe = 3900.35$; respectively, $M = 790$ ms and 884 ms), digit pairs were processed faster than word pairs ($F(1,22) = 117.23$, $p < 0.0001$, $MSe = 3083.75$; respectively, $M = 775$ ms and 898 ms), and these two factors interacted with one another ($F(1,22) = 4.68$, $p = 0.04$, $MSe = 567.41$). The difference between word and digit pairs was larger in the match trials than in the mismatch trials; this probably refers to the fact that, in the case of a match trial, the two numerals had to be encoded whereas in the case of mismatch, half of the time encoding the first numeral of the pair was sufficient to make the decision (given the assumption that it is a self-terminated process)⁴.

A second analysis was conducted only on the match trials with format and size of the items as within-subject factors. This analysis confirmed the effect of format observed previously ($F(1,23) = 128.52$, $p < 0.00001$, $MSe = 3059.65$). Furthermore, small pairs took less time than large ones ($F(1,23) = 8.13$, $p = 0.009$, $MSe = 3427.69$; $M = 767$ ms and 801 ms, respectively), and the interaction between format and size was significant ($F(1,23) = 6.17$, $p = 0.021$, $MSe = 1219.43$, see Table 3) with a larger size effect for word pairs than for digit pairs.

⁴ This interpretation is supported by the results obtained in a repeated-measure ANOVA restricted to the mismatch trials with format and location of the difference (first vs. second numeral of the pair) as factors. Beside the format effect ($F(1,23) = 66.7$, $p < 0.00001$, $MSe = 4325$) this analysis indicates an effect of the location of the difference ($F(1,23) = 7.29$, $p < 0.013$, $MSe = 3592$) which interacts with the format ($F(1,23) = 7.73$, $p < 0.011$, $MSe = 3094$). The location of the difference has nearly no influence on the RTs of digit trials ($t(23) = -0.1$, *NS*, $M = 827$ ms and 828 ms when the differences lies in the first vs. second numeral) whereas for word trials, RTs are much faster when the difference lies in the first numeral than when it lies in the second numeral ($t(23) = -3.24$, $p < 0.005$, respectively, $M = 905$ ms and $M = 970$ ms). These results also support the hypothesis of parallel encoding processes for pairs of digits but sequential processing for word numerals.

Table 3

Mean RTs of match trials as a function of the presentation format and the size of the items in the number-matching task

Format	Size		
	Large	Small	L–S
Words	874	822	52
Digits	728	712	16
W–D	146	110	

Note that the format-by-size interaction could not be attributed to the fact that the item length increased with number size in words but not in digits. Indeed, contrary to English, French has longer word names for the numbers 2–5 (mean of 4.75 characters) than for 6–9 (mean of 3.75 characters). The length of the stimuli should thus play in the opposite direction than the one observed here.

4.3. Effects of formats in multiplication: Contribution of encoding factors

The aim of adding the number-matching task to the multiplication task was to have a paradigm that did not imply the retrieval of arithmetical-facts but that nevertheless required the encoding of pairs of digits or words identical to those presented in the multiplications. The hypothesis was that if the interaction between format and size obtained in the multiplication task was mostly due to encoding factors, then a similar interaction should be found in a non-arithmetical task that involved the same encoding processes. The analysis run on the number-matching task alone indicated the presence of such a format-by-size interaction. Now, by combining the RTs of both tasks in a single analysis, we wanted to compare the strength of these interactions.

A repeated-measure ANOVA was computed with three within-subject factors: task (multiplication vs. number-matching⁵), format (word vs. digit) and size (small vs. large). If the interaction of format and size obtained in the multiplication task is due to encoding factors, one should expect the same interaction in the other task, i.e., the second-order interaction of task, format and size should not be significant.

The analysis revealed a main effect of task ($F(1,23) = 82.95$, $p < 0.00001$, $MSe = 69641.49$), format ($F(1,23) = 159.77$, $p < 0.00001$, $MSe = 15163.24$), and size ($F(1,23) = 77.28$, $p < 0.00001$, $MSe = 17298.26$), as well as significant first-order interactions between these factors (see Tables 2 and 3 for the means). The size-by-format interaction, already obtained for each of the tasks separately, was also present here ($F(1,23) = 12.40$, $p < 0.002$, $MSe = 1715.50$). Two new interactions involving the task variable were observed. First, the size effect was stronger in the multiplication task than in the numerical comparison task (difference between small and large pairs equalled 300 ms and 34 ms, respectively; $F(1,23) = 44.33$, $p < 0.00001$, $MSe = 19095.27$). This might be due to the fact that part of

⁵ For that task, only the RTs for match trials were considered.

the size effect is due to encoding factors (e.g., frequency), and thus plays equally in both tasks, whereas another part is task-specific (e.g., products of large problems require calculation or counting procedures more often than small problems). Second, the format effect was more important in the multiplication task than in the number-matching task (difference of 321 ms and 128 ms, respectively; $F(1,23) = 83.53$, $p < 0.0001$, $MSe = 5373.44$). However, the second-order interaction of the three variables was not significant ($F(1,23) < 1$) which indicates that the format-by-size interactions were identical in the multiplication and in the number-matching tasks.

5. Discussion

This first experiment replicates the major format effects obtained by Campbell and Clark (1992); Campbell (1994) with French-speaking participants: Digit problems were answered faster and more accurately than word problems and the difference was more pronounced for large multiplication problems than for small ones. Furthermore, the type of error produced was affected by the format. More naming errors and table-related errors with intrusion were observed in the word condition than in the digit condition and table-related errors produced in response to large digit problems were numerically closer to the target than errors in large word problems.

Some findings suggest, however, that the format-by-size interaction measured in the RTs might be due to factors playing at the encoding level rather than at the arithmetical-fact retrieval stage. Indeed, very much the same format-by-size interaction was obtained in the number-matching task where participants had to judge whether a pair of digit or word numerals expressed the same quantity as those represented by previously shown dot patterns. Such a similarity between the RTs obtained in the two tasks adds further credit to the deadline hypothesis proposed by McCloskey et al. (1992) for the interpretation of format effects on errors. Indeed, since encoding a pair of words is more time consuming than encoding a pair of digits (in particular for large numbers), the time available for the retrieval of the product is smaller for word problems than for digit problems. Such a state of affairs is likely to provoke more premature interruptions of the search process, and hence more errors in the word condition than in the digit condition, especially for large problems. Finally, given the greater time constraints, the errors produced for words would tend to be more distant from the target than those produced for digits (both in terms of numerical distance and in terms of error types).

However, this interpretation does not allow a direct explanation of two phenomena: (1) the observation of a larger format effect in the multiplication task than in the number-matching task, and (2) the fact that more table-related errors involved the intrusion of one (or both) of the operands in the word condition than in the digit condition.

As regards the format-by-task interaction, different accounts may be considered. First, some presentation formats may entertain a privileged link with some numerical tasks, as assumed by Campbell and Clark (1988) or by Dehaene (1992). For instance, the verbal code may underlie the retrieval of arithmetical-facts, as supposed in the triple-code model. In that perspective, two possibilities are still plausible. On the one hand, one could consider that the number-matching task privileges the verbal representations (used in short-term memory). Accordingly, exactly the same format effects are expected to be measured in both the number-matching task and the multiplication task. On the other hand, one may assume that the number-matching task is a true magnitude comparison task based on analog representations, in which case the format effect observed in the multiplication task should be smaller than the one obtained in the number matching task⁶, which is the opposite of the actual profile obtained.

A second explanation for the format-by-task interaction might be that, on a number of trials, numerals were encoded more than once. So far it has been assumed that the encoding differences between the formats were caused by a single encoding of the numerals in a display. However, Gielen et al. (1991), who registered eye movements in a number reading task and a number comparison task with a total of three numbers, found that participants quite frequently did not progress steadily from left to right, but made a regression to a previously seen number. This re-encoding probably occurred because the input signal of the number had decayed too much to be still reliable. Given the long response times in the current tasks, such re-encoding may have been needed as well, especially in the multiplication task with verbal numerals where response times took more than 1 s. If so, this may explain why the format effect is magnified in the multiplication task.

Finally, a third explanation of the format-by-task interaction might be related to interference effects playing at the output level. Indeed, many models of word recognition (e.g. Coltheart, 1978) assume that visually presented words can activate representations in the phonological output buffer through asemantic direct connections between the input graphemic lexicon and the phonological output lexicon or through non-lexical grapheme–phoneme conversion rules. Such a direct activation of phonological representations from visual input has also been proposed for word numerals by McCloskey (1992), but seems less likely for digits. Indeed, it has been shown that semantic magnitude information is accessed when participants have to transcode Arabic numerals into spoken verbal numerals but not when transcoding word numerals into phonemes. (Thus, Brysbaert (1995) (Experiment 4) obtained a number magnitude priming effect in Arabic numeral

⁶ Indeed, according to the triple-code model of Dehaene (1992), word multiplication problems would already be presented in the appropriate code for arithmetical-fact retrieval whereas digit problems would require first to be transcoded into the verbal corresponding representation before activating the arithmetical-fact network.

naming; Fias et al. (1996) found a SNARC effect⁷ in a phoneme monitoring task with Arabic input but Fias et al. (submitted) did not replicate this observation with written verbal numerals). Furthermore, when a verbal and an Arabic numeral are presented simultaneously, in a Stroop-like paradigm, an interference effect of the verbal numeral is reported on the naming of the Arabic numeral, but there is no effect of the Arabic numeral on the naming of the verbal numeral (Fias et al., submitted), a result which suggests a privileged access to phonological information for written verbal numerals but not for Arabic numerals.

In our experiment, the phonological representations activated by the word stimuli would not interfere equally in the two tasks. Indeed, the type of response required in the two tasks was not identical: a spoken verbal response had to be produced in the multiplication task but a simple key-press was used in the number-matching task. Consequently, one may expect that the phonological representations activated by the written word stimuli would interfere with the retrieved answer in the multiplication task but not in the number-matching task, while such phonological representations would be only weakly activated by the digits and would thus interfere only slightly with the response, even in the multiplication task.

This line of reasoning could also account for the larger proportion of intrusion errors in word multiplications than in digit multiplications. For instance, the presentation of the word problem 'seven $\times$ eight' would activate the phonological representations of 'seven' and 'eight' to a much higher degree than the corresponding digit problem would. During the retrieval process of the arithmetical-fact, the correct product, but also other associated products are activated (on an abstract or a specific code, the question is still open). Since the answer has to be spoken out loud, the phonological representations of the correct and close answers would be activated in the phonological output system. For instance, the phonological representations of 'fifty-six' but also of 'forty-eight' (6×8), 'sixty-four' (8×8), 'forty-nine' (7×7) and 'sixty-three' (7×9) would reach a certain level of activation. The response actually produced would be the one which is the most activated in the phonological output system. In the hypothesis of a unique, code-independent fact retrieval process, this level of activation would not be influenced by the presentation format of the equation, but would only depend on the degree of activation of the answers in the arithmetical-fact network. However, in the case of word problems (but not in the case of digit problems), an extra source of activation would influence the activation level of the phonological representations in the output system since the phonological representations

⁷ The SNARC effect refers to a spatial-numerical association between response codes in a parity judgment task in which small numbers were reacted to faster with the left hand than with the right hand and the reverse being true for large numbers. According to Dehaene et al. (1993), this effect indicates that magnitude information, which is represented as a position on a left-right oriented mental number line, was accessed during parity judgment.

corresponding to the addends of the problem are also activated at that level. As a consequence, (correct or incorrect) responses involving one of the addends would receive supplementary activation from the grapheme–phoneme conversion of the addends (or the direct asemanic lexical route) and might be more likely to be produced as the answer. In the example above, ‘forty-eight’ would receive extra activation from the addend ‘eight’ and might thus obtain more activation than ‘fifty-six’, thus leading to a table-related error with intrusion. Table-unrelated or miscellaneous errors with intrusion would be less likely to occur because these answers are less strongly activated in the network and thus their corresponding phonological representation would hardly be activated in the phonological output system.

Campbell and Clark (1992) have dismissed the possibility that intrusions would be due to interference in the verbal output system because they assume that only the solution with the highest activation in the arithmetical-fact network is forwarded to the phonological output buffer. If instead one assumes that all solutions which reach a certain degree of activation in the arithmetical-fact network (i.e., the correct as well as the associated solutions) automatically dissipate to the phonological output buffer to a degree relative to their activation level, then our explanation is plausible. Our hypothesis is similar to the one defended by Caramazza and Hillis (1990) who argue that several entries in the phonological (or orthographic) lexicon are activated in parallel from the semantic system in proportion to the degree of their semantic activation. The difference between the conceptions of Campbell and Clark on the one hand, and our own on the other, reflects the dichotomy between models assuming a discrete-sequential vs. a parallel or in cascade transmission of information between representational levels (Humphreys et al., 1988).

Interestingly, the interpretation we propose leads to different predictions than Campbell’s explanation for the position effect of the intrusion errors. Experiment 2 intends to contrast these different assumptions.

EXPERIMENT 2: MULTIPLICATIONS WITH DUTCH-SPEAKING PARTICIPANTS

According to Campbell, naming and operand intrusion errors are caused by an interference between fact-retrieval and reading processes. During the encoding stage, reading-based associations are activated unintentionally. These associations consist of a tendency to encode the multiplication problem as a two-digit number, with the left operand as the decade number and the right operand as the unit number. This gives rise to the position effect in the intrusion errors (Campbell, 1994).

Following Campbell’s explanation, speakers of a language with a reversed mapping of operand position and lexical class in number names should produce a different position effect in the intrusion errors. Two such languages are Dutch or German, which have the name of a 2-digit number started with the unit. For

instance, 48 is pronounced 'acht-en-veertig' (literally, 'eight-and-forty'). If the position effect in the intrusion errors is indeed due to reading-based associations between operand and lexical class, then speakers of these languages should present the following position effects. The word problem ' $\text{vier} \times \text{acht}$ ' (4×8) should not activate the verbal representation corresponding to 'forty-eight' but the representation corresponding to 'vier en tachtig' (84). Thus, a positional match between fact-retrieval and reading-based associations would be reflected in cases such as ' $\text{vier} \times \text{acht} = \text{vier-en-twintig}$ ' ($\text{four} \times \text{eight} = \text{twenty-four}$), which correspond to mismatches in French and English. A similar pattern does not need to be found for problems presented in Arabic format, as the Arabic notation is the same for Dutch and German as for French and English. In summary, Campbell's hypothesis would predict that Dutch- and French-speaking participants may present the same order effect in the digit condition but not in the word condition. An interaction between position effect, presentation format and language is thus to be expected.

If, on the contrary, intrusions are due to interference arising at the level of the phonological production system then different predictions are derived. We assume that word problems lead to the activation of phonological representations corresponding to each addend. Since addends used in our experiment are single digits or unit names, we expect interference, and thus intrusions, to appear mainly in the unit position of the answer rather than in the decade position (which is what was observed both by Campbell and in Experiment 1), and this is irrespective of the subject's language. Furthermore, since we read from left to right, the left operand is read before the right operand and the activation decay of the phonological representation of the first addend should be more important than the one for the second word operand. Consequently, interference should be stronger from the second operand than from the first one and this is irrespective of the subject's language. In summary, we expect more intrusions to appear in response to word problems than to digit problems, more intrusions in the unit position than in the decade position and more intrusions of the second word operand than of the first one, and this is irrespective of the subject's language. The same positional match is thus expected in both the Dutch- and French-speaking groups due to the combination of two factors: the fact that intrusions appear in the unit position and the fact that the intruder is more often the second addend of the problem.

To test these contrasting predictions, the multiplication task of Experiment 1 was repeated with Dutch-speaking participants from the northern half of Belgium (French is spoken in the southern half of Belgium).

6. Method

Participants: 24 students or research assistants from the Department of Psychology of the Catholic University of Leuven participated in the experiment. They were all native Dutch-speakers with normal or corrected-to-normal vision. Half were males. Mean age was 25.3 ($SD = 5.4$).

7. Procedure

The experiment took about 45 to 60 min and was finished in a single session. The stimuli were the same as those of the multiplication task of Experiment 1, except that the words were written in Dutch rather than in French (twee, drie, vier, vijf, zes, zeven, acht, negen: mean of 4.125 letters) and occupied 10–13 characters spaces. The procedure was also identical to the one used in Experiment 1. Half of the participants started with two sets of 56 items presented in digits, whereas the other half started with two sets of word problems. Presentation format changed every two sets and there was a brief rest period in-between. A longer rest period was given in the middle of the session (i.e., after the first eight sets of 56 multiplications).

8. Results

8.1. Analysis of the Dutch results

8.1.1. Error rates

A total of 760 commission errors were produced in response to word problems (7.08%) compared to 552 for digit problems (5.12%). As in the previous experiment, a first analysis examined the effects of format and size. Small problems were defined as those involving two operands inferior to or equal to 5; large problems contained at least one operand larger than 5. A higher percentage of errors was observed in response to word problems than to digit problems ($F(1,23) = 14.4$, $p = 0.001$, $MSe = 3.9$), and on large problems than on small problems ($F(1,23) = 44.3$, $p < 0.0001$, $MSe = 22.7$). The interaction between both factors was marginally significant ($F(1,23) = 3.4$, $p = 0.077$, $MSe = 3.6$), suggesting a larger difference in error rate between small and large problems in the word condition than in the digit condition (the format effect was slightly stronger for large problems ($t(23) = 3.19$, $p = 0.0041$) than for small problems ($t(23) = 2.29$, $p = 0.031$), see Table 4).

Table 4

Performance (RTs and percentage of errors) as a function of format and size of the problem in the Dutch-speaking sample

	RTs (ms)			% of errors		
	Large	Small	L–S	Large	Small	L–S
Words	1185	946	239	8.62	1.43	7.19
Digits	918	717	201	6.34	0.61	5.73
W-D	267	229		2.28	0.82	

8.1.2. Error types

The same typology of errors was used as in Experiment 1. Only three operation errors occurred. Two of them (i.e., $2 \times 4 = 6$ in the digit condition) could be interpreted as table-related errors and were classified in that category for subsequent analyses; the other one (i.e., $6 \times 8 = 48$ in the word condition) was included in the miscellaneous category. Furthermore, 35 errors in the word condition and 9 in the digit condition could equally well be considered as naming errors or table-related errors with intrusion (e.g., $4 \times 8 = 48$). They were analysed as belonging to the former category.

Paired *t*-tests calculated separately on the proportion of errors of each type (see Table 5) showed a significant format effect for table-related errors as well as a marginally significant effect for naming errors. The probability of table-related errors with intrusions was higher in the word condition than in the digit condition, whereas the reverse was true for errors without intrusion. Furthermore, naming errors tended to be more frequent in response to word problems than to digit problems. For the miscellaneous errors, no format effect was observed, even when the data with and without intrusion were combined (actually, in this analysis there was a tendency towards a larger proportion of errors in the digit format, $t(23) = -1.91$, $p = 0.069$). Overall then, very much the same pattern of errors was observed in the Dutch-speaking sample as in the French-speaking sample.

An ANOVA was calculated on the table-related errors with format and distance (± 1 operand vs. ± 2 operands or more) as repeated measures to see whether we could replicate the interaction reported by Campbell et al. and which indicates that errors on word problems are more distant to the target than errors on digit problems. This analysis gave rise to a main effect of format ($F(1,23) = 7.30$, $p < 0.013$, $MSe = 25.61$) and of distance ($F(1,23) = 6.75$, $p < 0.016$, $MSe = 40.51$) as well as an interaction between both factors ($F(1,23) = 18.60$, $p = 0.0003$, $MSe = 15.43$). An equal rate of close errors was observed in both formats ($M = 12.62$ for digits and $M = 11.96$ for words), but distant errors were twice as numerous in the word format ($M = 12.04$) than in the digit format ($M = 5.79$). The interaction did not seem to depend on the presence of an intrusion, as an ANOVA

Table 5

Error distribution in the Dutch-speaking sample as a function of the presentation format (number and mean of the proportions)

Type of error	Word condition	Digit condition	Two-tailed $t_{(23)}$	p
Naming	55 (0.085)	19 (0.052)	1.99	0.059
Table-related and intrusion	369 (0.484)	196 (0.321)	3.30	0.003
Table-related	207 (0.262)	247 (0.448)	-4.44	0.0001
Table-unrelated and intrusion	41 (0.047)	20 (0.036)	0.91	NS
Table-unrelated	13 (0.017)	11 (0.017)	-0.14	NS
Miscellaneous and intrusion	27 (0.022)	10 (0.020)	0.88	NS
Miscellaneous	49 (0.082)	48 (0.106)	-1.21	NS
Total	760 (1)	552 (1)		

with intrusion (present vs. absent) as a third factor did not reveal a reliable interaction between format, distance and intrusion ($F(1,23) = 2.91$, $p = 0.1$).

8.1.3. Response times

Median response times were calculated for trials with a correct response and an adequate functioning of the voice-activated relay (overall, 8.8% of the trials had to be discarded in the digit condition and 9.3% in the word condition). Once more, the correlation obtained between problems presented in the two formats was high ($r = 0.80$, $p < 0.001$).

The data were submitted to a repeated-measure ANOVA with format and size of the problem as variables. Both main effects, as well as the interaction, were significant (see Table 4 for the means). Word problems gave rise to slower RTs than digit problems ($F(1,23) = 185.59$, $p < 0.00001$, $MSe = 7956.26$), large problems were answered more slowly than small problems ($F(1,23) = 66.12$, $p < 0.0001$, $MSe = 17581.17$), and the difference between the two formats was bigger for large problems than for small ones ($F(1,23) = 6.18$, $p < 0.021$, $MSe = 1329.87$).

As the participants were given all problems in a single session, we could examine whether the format-by-size interaction was stable over time. Therefore, an ANOVA was computed with three repeated measures: format, size and rank number of the series (4 levels). As in the global analysis, the format ($F(1,23) = 176.92$, $MSe = 34452.78$, $p < 0.0001$), size ($F(1,23) = 60.13$, $MSe = 77632.66$, $p < 0.0001$) and format-by-size interaction ($F(1,23) = 8.76$, $MSe = 4612.07$, $p = 0.007$) were significant. The rank number of the series had a main effect on the RTs ($F(3,69) = 35.94$, $MSe = 6627.32$, $p < 0.0001$), and interacted with both format ($F(3,69) = 3.82$, $MSe = 3239.58$, $p = 0.021$) and size ($F(3,69) = 9.60$, $MSe = 2192.67$, $p < 0.0001$). These interactions were due to the fact that: (1) the effect of format tended to decrease over practice, except in the third series which came just after the long pause (the format effect was 275 ms, 267 ms, 341 ms and 228 ms for series 1, 2, 3 and 4), and (2) the effect of problem size decreased from the first to the second series (the difference between large and small problems was 265 ms, 209 ms, 205 ms and 203 ms for series 1, 2, 3 and 4). More importantly, the second-order interaction of the three variables was significant ($F(3,69) = 3.43$, $MSe = 1355.51$, $p = 0.028$). Thus, the critical interaction of format and size was not stable over time. To understand the meaning of the second-order interaction, separate ANOVAs (including size and format) were run on each of the series. These analyses indicated that although the main effect of format and size remained highly reliable throughout the whole experiment, their interaction was decreasing over time (except just after the pause) down to a non-significant level in the last series (see Table 6 for the RTs means and the effects).

In summary, the experiment with Dutch-speaking participants replicated the format effects observed by Campbell. Yet, we also discovered that the format-by-size interaction measured on the RTs was not stable over time and did actually decrease down to a non-significant level in the last series.

Table 6

Mean RTs as a function of the size and format of the problem in each series for the Dutch-speaking sample and statistics associated with the effect of these variables

Format	Size	Series 1 (ms)	Series 2 (ms)	Series 3 (ms)	Series 4
Word	Large	1311	1191	1151	1132 ms
Word	Small	1008	963	925	925 ms
Digit	Large	998	905	892	900 ms
Digit	Small	771	716	708	701 ms
Statistics: $F_{(1,23)}$ (and associated p value)					
Format effect		128.68 (0.0001)	159.01 (0.0001)	150.90 (0.0001)	121.12 (0.0001)
Size effect		51.24 (0.0001)	57.21 (0.0001)	62.01 (0.0001)	59.04 (0.0001)
Interaction		9.33 (0.006)	5.14 (0.033)	9.21 (0.006)	0.17 (NS)

8.2. Comparison of the two populations

This experiment intended to detect the influence of language on the multiplication task, and in particular, on the position effect observed in operand intrusions (both for errors and correct responses). But before going into the analyses of the intrusions, we would like to mention a first striking difference appearing in the miscellaneous errors (see Table 7). Whereas the French-speaking participants made a considerable number of errors in which they reported a decade-name only, the Dutch-speaking participants almost never did. The reverse was true for wrong responses in which only a unit-name was reported. Although remarkable, this language difference is not really surprising and is probably an artefact of the structure of the language. As the French name of a number starts with the decade, French-speaking participants are more likely to pronounce the decade than the unit before they realize that their answer is wrong. The opposite is true for Dutch-speaking participants. Therefore, chances are low that this effect of language on miscellaneous errors is revealing for the processes underlying fact-retrieval.

More interesting findings were expected for the position effect of intrusions. According to Campbell, the tendency for operand intrusion errors to preserve the position they were holding in the original problem, reflects an interaction between reading and calculation processes. A similar explanation was provided for the

Table 7

Distribution of miscellaneous errors^a as a function of the group, the presentation format and the lexical structure of the response produced

Format	Word		Digit	
	French-speaking	Dutch-speaking	French-speaking	Dutch-speaking
Structure				
Unit	1	34 (35 ^b)	0	33 (34 ^b)
Decade	27	1	26	0
Decade-Unit	21	38 (39 ^b)	25	23 (24 ^b)

^a The operation error is not considered in this data set.

^b Error rates when the item '9 × 4' and 'negen × vier' are included.

finding that production latencies are shorter for answers that have a digit in the same position as the original problem (e.g., $6 \times 4 = 24$) than for answers that have a digit in a different position (e.g., $4 \times 6 = 24$). If Campbell's explanation is correct, then the fact that Dutch reverses the names of units and decades, should change the direction of these effects in the word format.

8.2.1. Analysis of positional match in the table-related errors with intrusions

The analyses of the table-related errors with intrusion were run to see whether the tendency of the intruding operand to keep the same position in the response as in the problem, was influenced by the participant's language (French or Dutch). Such a comparison is meaningful only for those numbers that have a reversal of units and decades in their names. As this is not the case for the teens below 17, analyses were restricted to responses greater than or equal to 17. They were run on a total of 832 errors (299 in the French-speaking group and 533 in the Dutch-speaking group⁸).

A first analysis was run on the percentage of table-related errors with intrusion for which the intruder respected the position held in the problem. Three participants (two French-speaking and one Dutch-speaking) were not included in this analysis because they did not produce table-related errors with intrusion in the Arabic format condition. A mixed ANOVA was run with group as a between-subjects variable and format as a within-subject variable. It only indicated a main effect of format ($F(1.43) = 4.32$, $p = 0.044$, $MSe = 662.93$) with a higher percentage of 'in-positions' in the Arabic format than in the word format (respectively, $M = 64.9\%$ and $M = 53.6\%$). This factor did not interact with the group ($F(1.43) < 1$). The rate of 'in-position' was significantly higher than expected by chance both in the digit condition (one-tailed t -test, $t(44) = 3.49$, $p < 0.0006$) and, to a lesser degree, in the word condition ($t(44) = 1.76$, $p < 0.043$).

A subsequent analysis was computed on the number of errors with the group as a between-subjects variable and three within factors: the format, the intruding operand (first vs. second) and the lexical class in which the intrusion appeared (decade or unit). Errors were significantly more numerous in the Dutch-speaking sample than in the French-speaking sample ($F(1.46) = 5.53$, $p = 0.023$, $MSe = 25.80$, respectively, $M = 22.2$, $SD = 18.5$ and $M = 12.5$, $SD = 8.4$)⁹ and in response to word problems than to digit problems ($F(1.46) = 32.68$, $p < 0.00001$, $MSe = 4.07$, with $M = 11.0$ for words and 6.3 for digits). The interaction of format and group only approached significance ($F(1.46) = 3.32$, $p = 0.075$, $MSe = 4.07$) suggesting a larger effect of the format in the Dutch-speaking group than in the French-speaking group. The operand effect was significant ($F(1.46) = 16.67$, $p < 0.005$, $MSe = 1.89$) but was modulated by the format ($F(1.46) = 7.76$, $p <$

⁸ The errors produced in response to the item ' 9×4 ' in both the digit and the word format were excluded from the analyses in both groups (since, due a technical problem, these were not available for the French-speakers).

⁹ This effect is not specific for the type of error analysed here: the Dutch-speakers were generally faster than the French-speakers but they produced more errors.

0.008, $MSe = 2.04$): in the digit condition, intrusions concerned equally the first and second operand (respectively, $M = 3.15$ and $M = 3.17$) whereas in the word format, intrusions were less frequent for the first ($M = 4.69$) than for the second operand ($M = 6.33$). The lexical class was significant ($F(1.46) = 13.16$, $p = 0.0007$, $MSe = 20.3$) and interacted with the format ($F(1.46) = 11.93$, $p = 0.0012$, $MSe = 3.04$): intrusions appeared more frequently in the Unit position than in the Decade position; such a difference however, was stronger in the word format ($M = 3.21$ in the Decade position and $M = 7.79$ in the Unit position; $t_{(47)} = 3.85$, $p = 0.0004$, 2-tail) than in the digit format (respectively, $M = 2.10$ and $M = 4.21$; $t_{(47)} = 2.94$, $p = 0.005$, 2-tail). Finally, the interaction between the operand and the lexical class was significant ($F(1.46) = 18.06$, $p = 0.0001$, $MSe = 1.68$). This interaction (see Table 8 for the means) showed that intrusions in the Unit position were more frequent when the intruder held the same position in the problem (i.e., when it was the second operand, e.g., $4 \times 8 = 28$) than when it did not ($8 \times 4 = 28$) ($t_{(47)} = 4.2$, $p = 0.0001$, 2-tail). However, such a difference was not statistically significant when intrusions appeared in the decade position ($t_{(47)} = 1.01$, *NS*).

In summary, the two groups did not differ in terms of the tendency of intrusion errors to preserve the position they were holding in the problem.

8.2.2. Analysis of problems with one of the operands in the correct answer

The second effect which seemed interesting to compare across languages, was the finding that multiplications with one of the operands in the 'correct' position of the answer are responded to faster than multiplications with one of the operands in the 'wrong' position (e.g., $7 \times 5 = 35$ vs. $5 \times 7 = 35$; Campbell and Clark, 1992). Campbell and Clark also found that the effect was somewhat stronger in the word condition than in the digit condition. Median RTs for correct responses with a positional match were compared to those with a mismatch for both presentation formats and both groups. These data were based on 6 items repeated 8 times in each format. The analysis involved one between-subjects variable (group) and two within-subject measures: format and position (match vs. mismatch).

The Dutch-speaking participants were faster than the French-speaking ($F(1.46) = 9.56$, $p = 0.0034$, $MSe = 205434.12$); digit problems were answered faster than word problems ($F(1.46) = 200.10$, $p < 0.0001$, $MSe = 22403.48$, $M = 974$ ms and 1280 ms), and responses were accelerated by a positional match ($F(1.46) = 4.34$, $p = 0.043$, $MSe = 13154.88$, $M = 1109$ ms for a match and $M = 1144$ ms for a mismatch) but the last factor did not interact with either format ($F(1.46) < 1$) or group ($F(1.46) < 1$).

Table 8
Mean rates of table-related errors with intrusions as a function of the operand and the lexical class

Lexical class	Operand	
	First operand	Second operand
Decade	2.81	2.51
Unit	5.03	6.98

The same analysis was repeated on the number of errors for these problems (a total of 392 errors). It indicated a higher number of errors in the Dutch-speaking group than in the French-speaking group ($F(1.46) = 4.11$, $p = 0.048$, $MSe = 13.18$, $M = 10.29$ and $M = 6.04$) and an interaction between format and positional match ($F(1.46) = 4.62$, $p = 0.037$, $MSe = 1.80$), showing that items with a positional match gave rise to less errors than items with a positional mismatch in the word condition (one-tail paired $t_{(47)} = -1.90$, $p = 0.032$, $M = 1.83$ and $M = 2.37$) but not in the digit condition (paired $t_{(47)} = 1.04$, NS). (Similar results are found with non-parametric statistics). Here again, the positional match was not influenced by the group (positional match*group: $F(1.46) = 1.24$, $p > 0.1$; positional match*format*group: $F(1.46) < 1$).

Finally, analyses were run after excluding the problems '6 × 2' and '5 × 3' which have a response below 17 (see above). For these problems, the ANOVA of the RTs returned a main effect of group ($F(1.46) = 14.33$, $p = 0.0004$, $MSe = 305603.79$) and of format ($F(1.46) = 190.09$, $p < 0.00001$, $MSe = 28172.30$). The positional match did not reach statistical significance ($F(1.46) = 2.18$, $p > 0.1$) and did not interact with the group ($F(1.46) < 1$). However, a non-parametric analysis of these same data showed that problems with a positional match were answered faster than those with a positional mismatch (Wilcoxon test, $Z = 2.71$, one-sided $p = 0.0034$, $M = 1175$ ms for match trials and $M = 1207$ ms for mismatch trials). With respect to the analysis of the number of errors, there was only a main effect of group ($F(1.46) = 4.18$, $p < 0.05$, $MSe = 10.78$) (this was also true for non-parametric statistics). It should be underlined, however, that the number of errors for these problems was quite low and thus that the power of the test is not high (174 errors in the word condition and 163 in the digit condition for a total of 48 participants).

9. Discussion

This second experiment replicated the finding of Campbell and Clark (1992) of format effects on speed and accuracy of multiplication, with a Dutch-speaking population. However, the critical format-by-size interaction measured on the RTs was shown to be unstable over time and decreasing down to a non-significant level at the end of the experimental session.

More importantly, this experiment aimed at studying the impact of the structure of the language on the position effect observed by Campbell et al. on intrusions in both correct responses and table-related errors. Two hypotheses were contrasted. According to Campbell et al., a reversal of the position effect should have been observed for the Dutch-speaking subjects confronted with word problems. On the contrary, the interpretation in terms of interference in the phonological output system predicted similar position effect in the two samples for both presentation formats. The results obtained in the different analyses indicate that the position effects are absolutely identical in both samples. Furthermore, as predicted by the

verbal interference hypothesis, the two samples showed more intrusion errors in response to word problems than to digit problems, more intrusions of the second operand than of the first one and more intrusions appearing in the unit position of the answer than in the decade position.

GENERAL DISCUSSION

The purpose of this article was to find out how arithmetical-facts are mentally represented and, more precisely, whether the retrieval process of these facts is affected by the input notation. Two experiments were conducted in which participants had to multiply pairs of numbers presented in digit or word formats. Although different languages were used (i.e., French and Dutch instead of English) and formats were blocked rather than alternated from trial to trial, we replicated the format effects reported by Campbell and Clark (1992); Campbell (1994). Problems presented in words were more difficult than problems presented in digits, and the difference between word and digit trials was more pronounced for large problems than for small problems.

According to Campbell et al., the interaction between format and problem size indicates that the input format influences the retrieval of arithmetical-facts. However, two results of our experiments argue against this interpretation. First, a similar format-by-size interaction was obtained in a number-matching task that did not involve the retrieval of arithmetical-facts. Second, the format-by-size interaction observed in the multiplication task was unstable with time: During the session, it decreased down to a non-significant level in the last series. These two findings suggest that the size-by-format interaction is not due to the retrieval process of arithmetical-facts but rather to factors that come into play during the encoding of a pair of digit or word numerals. McCloskey et al. (1992) showed that the difference in frequency between word and digit numerals increased with their numerical size. Given that frequency affects numeral recognition time (Brysbaert, 1995), this factor would also predict an increasing encoding time difference between digits and words as the quantity they represent becomes larger.

Regarding errors, their rate was higher in response to word trials than to digit trials and this was especially true for large problems. Furthermore, errors on word problems tended to be more distant from the target than errors on digit problems. However, if we follow the deadline hypothesis proposed by McCloskey et al. (1992), these two format effects can be the consequence of encoding time differences between words and digits. If participants limit the period of time they are allowing themselves before giving a response, then the time available for retrieval (after encoding the problem) is smaller for words than for digits and this handicap increases with the size of the numbers.

At first sight, some format effects reported by Campbell et al. and replicated in our experiments seem difficult to explain by the encoding difference hypothesis. For instance, significantly more table-related errors with intrusion of the problem's operand were observed in the word format than in the digit format. (A similar

finding was obtained for naming errors but it should be kept in mind that nearly all of these errors could also be classified as table-related errors with intrusion).

According to Campbell (1994), these errors witness the interaction of two processes: problem reading and fact retrieval. More explicitly, he assumes that

1. Arabic and verbal inputs provide independent pathways to number-fact representations;
2. the pathway to the arithmetical-fact network is stronger (faster) from the digit input than from the written verbal input;
3. reading associations from the verbal inputs to the auditory-verbal representations are stronger than those from the visual Arabic code;
4. numeral-reading and number-fact retrieval involve a common representational system, i.e., the verbal-code system; and
5. the interferences are due to reading-based associations which involve a left-to-decade and a right-to-unit association.

On the basis of these assumptions, Campbell explains why verbally presented problems take longer to solve and why they are more likely to include errors which are due to competition between reading-based associations and number-fact retrieval. In addition, the idea of reading-based associations leads to the prediction that operand intrusions will tend to keep their position as in the original problem formulation.

Yet, if numeral-reading and number-fact retrieval involve a common verbal representational system, it seems odd to assume that the associations from verbal input to this verbal representational system are less strong than those from Arabic input in the case of fact retrieval, but stronger in the case of numeral reading. Furthermore, direct access from digits to the verbal system is not supported by the data of Fias et al. [see Section 5 of Experiment 1, in particular, the presence of a SNARC effect in a phoneme monitoring with Arabic stimuli (Fias et al., 1996) and the absence of interference of digits in a naming task of verbal numerals (Fias et al., submitted)]. Finally, Campbell's interpretation of intrusion errors in terms of reading-based left-to-decade and right-to-unit associations is contradicted by our results. Indeed, it should predict a reversal of the errors in a language in which the names of two-digit numbers start with the units (i.e., 48 in Dutch is named 'eight and forty') but our data showed exactly the same pattern of intrusions for both Dutch-speaking and French-speaking participants in the digit condition as well as in the word condition. Therefore, we think that intrusions cannot be due to interference of reading and fact-retrieval processes.

Instead, we propose that operand intrusion errors are caused by interferences at the phonological output system rather than at the arithmetical-fact retrieval stage. These interferences are stronger in the word condition than in the digit condition because phonological representations are automatically activated from the graphemic representation of words but not from the ideographic representation of digits, as shown by Fias et al. (see Section 5 of Experiment 1). We assume that all

activation in the arithmetical-fact network automatically dissipates to the phonological output buffer. That is, the numerical system does not wait until one single—correct—solution has been selected before it starts to trigger the output buffer so that the correct response, as well as, to a lesser degree, closely associated responses, will be activated in the output buffer. In the case of word problems these activations will add up with the activations coming from the automatic transcoding of the written words. As a consequence, arithmetically-related responses with phonological components which correspond to the problem's operands will receive a double source of activation in the phonological output system and thus give rise to the production of semantically-related intrusion errors. Within this perspective, the position effect encountered in the intrusion errors should present two characteristics. First, irrespective of the language, the second operand has more chances of intruding because it is read last and its activation has decayed less at the time solutions start to trigger the response buffer. Second, again irrespective of the language, intrusions will mainly occur in the unit position, as the operands are numbers below ten. Both characteristics were present in our experiments.

If reading interference plays a role at the level of the phonological output system rather than at the level of fact retrieval, how can we explain why so few intrusions gave rise to table-unrelated or miscellaneous errors? To answer this question, it may be instructive to examine which system design would be prone to such errors. This would be the case if the activation of tens-names and unit-names in the phonological output system were independent. Then, a word problem like '*seven $\times$ eight*' would trigger the arithmetical-fact '*fifty-six*' which in turn would activate /*fifty*/ and /*six*/ in the output buffer. Together with the activation of the operands /*seven*/ and /*eight*/, table-unrelated errors like /*fifty-eight*/ and (to a lesser extent) /*fifty-seven*/ would be expected.

However, such a conception is only realistic for a discrete-sequential model, where a single response is forwarded from the fact-retrieval system to the phonological output system. Otherwise, many more table-unrelated and miscellaneous errors should be expected, even in the digit condition where nearly no interference from reading occurs. Let us explain that with an example. For instance, in a cascade model, presentation of the digit problem '*7 $\times$ 8*' would give rise to a strong activation of the phonological representation of the correct response /*fifty-six*/ but also to a weaker activation of associated responses such as /*sixty-four*/ (*8 $\times$ 8*), /*forty-eight*/ (*6 $\times$ 8*), /*forty-nine*/ (*7 $\times$ 7*) and /*sixty-three*/ (*7 $\times$ 9*). If the representations of decade-names and unit-names were activated in complete isolation, then quite often the decade-names /*forty*/ and /*sixty*/ would be pronounced because these names are triggered by two different solutions. This would lead to miscellaneous errors of the type /*forty-six*/ and /*sixty-six*/, which, as we have seen, are very rare in normal individuals. Therefore, we have to recognize that when a response is activated in the arithmetical network, it does not lead to an independent activation of the decade- and unit-name in the phonological output system, but rather it activates the representation of a combined decade-unit

structure. A similar hypothesis has actually been proposed by Cohen and Dehaene (1994)¹⁰.

If we accept the idea of decade-unit structures activated in the phonological output system, there is no problem any longer to explain the near absence of miscellaneous or table-unrelated errors. Let us come back to our problem but in the word format, i.e., '*seven × eight*'. The same responses as those activated for the digit format are triggered in the arithmetical-fact network and dissipate to the phonological output system. However, now we have the additional activation of the phonological representations of the addends /seven/ and /eight/. This state of affairs would not give rise to responses such as /fifty-eight/ because this decade-unit structure did not receive any activation from the arithmetical-fact network and so would only be primed by the operand /eight/. Similarly, table-unrelated errors such as /eighty-one/ would nearly never occur because such answers receive virtually no activation in the arithmetical-fact system. On the other hand, what we should expect are responses such as /forty-eight/ which are triggered by both the number-fact system and one of the operands. So, if the notion of connected decade-unit structures triggered by the arithmetical-fact retrieval is accepted, our interpretation correctly predicts that intrusion errors mostly correspond to table-related errors rather than to table-unrelated or miscellaneous errors.

In summary, the two experiments presented here strongly suggest that the format effects obtained by Campbell and Clark (Campbell and Clark, 1992; Campbell, 1994) are the result of encoding differences between word and digit numerals and interference effects at the final output stage. So, there is no strong evidence left that shows an impact of the presentation format on the realization of simple calculations. Such a state of affairs argues for the existence of a single representation medium underlying arithmetical-fact retrieval. A similar conclusion was reached by other authors studying the influence of notational input on arithmetical-fact retrieval (e.g., LeFevre et al., 1988; Marsh and Maki, 1976; Noël and Seron, 1992; Sokol et al., 1991).

Two further issues need to be addressed: First, what is the nature of the underlying representation; and second, is this medium identical for all kinds of number processing? According to McCloskey et al. (1985), a single amodal base-ten representation is used for all tasks. A different position was taken by Dehaene (1992) who proposed that an auditory-verbal representation is used for the retrieval of arithmetical-facts but not for tasks such as magnitude comparison or parity judgment. Our results do not allow to solve these questions definitely but some indications can nevertheless be pointed out.

Regarding the first question, we have found several pieces of evidence against Dehaene's idea of an auditory-verbal basis of arithmetical-fact retrieval. First, if

¹⁰ These authors examined a patient who could not retrieve the answer of multiplications such as 6×9 but who could nevertheless decide which of two numerals, e.g., 54 and 55, belonged to the multiplication table. Faced with these results, these authors suggested that there might be an output lexicon containing compound words corresponding to the results of simple multiplications.

digit and word problems were transcoded into a common auditory-verbal representation before accessing the arithmetical-fact store, we would expect the same reading interference for both formats. Second, access to the auditory-verbal (phonological) representations is faster from verbal numerals than from Arabic numerals (Fias et al., submitted), whereas we observe the reverse pattern in our multiplication task. Finally, if the retrieval of arithmetical-facts were mediated through verbal representations, the format differences between words and digits should have been smaller in the multiplication task than in the number-matching task. In fact, we found that the format effect was twice as large in the multiplication task than in the number-matching task. In summary then, the hypothesis of a verbal code underlying the retrieval of arithmetical-facts is not well supported by our data, which are more in line with the hypothesis of an amodal (or maybe even a visual-Arabic) representation underlying the retrieval of arithmetical-facts.

Regarding the second question, i.e., the commonalities of the code underlying the retrieval of arithmetical-facts in the other types of numerical tasks, the major indications come from experiments in which the presentation formats and the numerical tasks were manipulated simultaneously. Such a design has been used by Campbell (1994) as well as by Vorberg and Blankenberger (1993); Blankenberger and Vorberg (1997).

In their Experiment 1, Vorberg et al. compared numbers displayed in digits, words, dots or dice in three tasks: naming, incrementing or decrementing the presented number. In those conditions, an interaction between format and operation was obtained: naming a number-word was about as fast as naming a digit but increment costs were heavier for number-words (125 ms) than for the three other formats (86 ms for digits, dice and dots combined). What do these results tell us? First, one could conclude that there is a privileged link between, on the one hand, the word format and the naming task and, on the other hand, the digit, dot and die format and the increment operation. However, one could also interpret those results simply by arguing that incrementing or decrementing a number requires some access to the semantic system whereas for naming, this semantic access is probably necessary for the digits, dots and dice but not for the words, for which direct transcoding mechanisms have been largely shown in the field of language processing.

The three other studies which have obtained a task-by-format interaction have all compared additions with multiplications under different presentation formats. Thus, Campbell (1994) has compared word and digit problems, Vorberg et al. (Experiment 2) has used digits and dice whereas Blankenberger et al. (Experiment 2) has contrasted digits, words and dice. However, the pattern of results obtained is not very informative with respect to the issue of the possible associations between tasks and underlying internal code. Thus, in Campbell's study, an advantage of addition over multiplication was observed both for digits (advantage of 89 ms) and for words (advantage of 99 ms)¹¹. On the contrary, Blankenberger et

¹¹ RTs calculated on the whole set of trials, and not simply averaged on small and large problems since large problems are much more numerous than small ones.

al. only found an advantage of addition over multiplication in the digit condition (7 ms). In the word condition, there was an advantage of multiplication (34 ms, means obtained in the visual condition where the interaction appeared). Finally, the advantage of addition in the digit condition was not replicated by Vorberg et al. who reported a 28 ms advantage of multiplication. It may be interesting to note that these discrepancies cannot be due to the range of stimuli used (Vorberg and Blankenberger always used numbers from 2 to 6, whereas Campbell used the range from 2 to 9). If we consider only the results for the small problems in Campbell's experiment (i.e., those in the range of 2 to 5), the same profile remains, namely an advantage of addition in both formats (22 ms for digits and 66 ms for words). It seems to us that such a variability in the results does not lead to any clear conclusion. Furthermore, up to now, experiments devised to test task-by-format interactions, have not been designed on the basis of *a priori* hypotheses about the association between tasks and internal codes. Indeed, the only theory which has specified such privileged associations is that of Dehaene (1992). However, in this model, additions and multiplications are both supposed to be mediated by auditory-verbal representations and therefore no task-by-format interaction is expected for these two operations. (We are currently trying to test the task-code associations assumed by the triple-code model by measuring the effect of presentation format in different numerical tasks expected to rely on different internal codes).

In conclusion, the results described in the present paper support the hypothesis of a single medium underlying the retrieval of multiplicative facts. This medium is probably not verbal, at least not if it is required that every input is transcoded into that code before the fact store can be accessed. Finally, the question of whether the internal code for the retrieval of multiplicative facts is the same as the one used in other numerical tasks, still remains open and deserves more experimental investigation.

Acknowledgments

We would like to thank Michel Désert and Annarita Bellato for helping us collect the data. Special thanks are also due to Jamie Campbell and Xavier Seron for their helpful comments. The first and the third authors were supported by the National Research Fund of Belgium. The second author benefitted from a grant to study 'psychological aspects of Dutch-French bilingualism' from the Flemish Government. This research project was also supported by a grant from the "Pole d'Attraction Inter-Universitaire, Pulia".

References

- Anderson, J.R., 1978. Arguments concerning representations for mental imagery. *Psychological Review* 85, 249–277.

- Ashcraft, M.H., 1992. Cognitive Arithmetic: A review of data and theory. *Cognition* 44, 75–106.
- A.D. Baddeley (1986), *Working memory*. Oxford University Press.
- S. Blankenberger and D. Vorberg (1997), The single-format assumption in arithmetic fact retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 23, in press.
- Brysbaert, M., 1995. Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *Journal of Experimental Psychology: General* 124, 434–452.
- Campbell, J.I.D., 1991. Conditions of error priming in number-fact retrieval. *Memory and Cognition* 19, 197–209.
- J.I.D. Campbell (1992), In defense of the encoding-complex approach: Reply to McCloskey, Macaruso Whetstone. In: J.I.D. Campbell (Ed.), *The nature and origin of mathematical skills*. Elsevier Science, North Holland, pp. 539–556.
- Campbell, J.I.D., 1994. Architectures for numerical cognition. *Cognition* 53, 1–44.
- 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,(2) 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.
- J.I.D. Campbell and J.M. Clark (1992), Cognitive number processing: An encoding-complex perspective. In: J.I.D. Campbell (Ed.), *The nature and origin of mathematical skills*. Amsterdam, Elsevier, pp. 457–491.
- J.I.D. Campbell and D.P.M. Tarling (1996), Retrieval process in arithmetic production and verification. *Memory and Cognition*, 24, 156–172.
- Caramazza, A., Hillis, A.E., 1990. Where do semantic errors come from?. *Cortex* 26, 95–122.
- Caramazza, A., Hillis, A.E., Rapp, B.C., Romani, C., 1990. The multiple semantics hypothesis: Multiple confusions?. *Cognitive Neuropsychology* 7, 161–189.
- Cohen, L., Dehaene, S., 1994. Amnesia for arithmetic facts: A single-case study. *Brain and Language* 47, 214–232.
- M. Coltheart (1978), Lexical access in simple reading tasks. In: G. Underwood (Ed.), *Strategies of information processing*. London, Academic Press.
- Cornet, J.-A., Seron, X., Deloche, G., Lories, G., 1988. Cognitive models of simple mental arithmetic: A critical review. *Cahiers de Psychologie Cognitive* 8, 551–571.
- Clark, J.M., Campbell, J.I.D., 1991. Integrated versus modular theories of number skills and acalculia. *Brain and Cognition* 17, 204–239.
- Dehaene, S., 1992. Varieties of numerical abilities. *Cognition* 44, 1–42.
- 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., 1991. Two mental calculation systems: A case-study of severe acalculia with preserved approximation. *Neuropsychologia* 29, 1045–1074.
- Dehaene, S., Cohen, L., 1995. Towards an anatomical and functional model of number processing. *Mathematical Cognition* 1, 83–120.
- Dehaene, S., Dupoux, E., Mehler, J., 1990. Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance* 16, 626–641.
- Frenck-Mestre, C., Vaid, J., 1993. Activation of number facts in bilinguals. *Memory and Cognition* 21, 809–818.
- Fias, W., Brysbaert, M., Geypens, F., d'Ydewalle, G., 1996. The importance of magnitude information in numerical processing: Evidence from the SNARC effect. *Mathematical Cognition* 2, 95–110.
- W. Fias, M. Brysbaert and G. d'Ydewalle, *Multiple routes in number processing? Yes for verbal numbers, no for Arabic numbers*, submitted.
- Gielen, I., Brysbaert, M., Dhondt, A., 1991. The syllable-length effect in number processing is task-dependent. *Perception and Psychophysics* 50, 449–458.
- Gonzalez, E.G., Kolers, P.A., 1982. Mental manipulation of arithmetic symbols. *Journal of Experimental Psychology: Learning, Memory and Cognition* 4, 308–319.

- E.G. Gonzalez and P.A. Kolers (1987), Notational constraints on mental operations. In: G. Deloche and X. Seron (Eds.), *Mathematical disabilities: A cognitive neuropsychological perspective*. Hillsdale, NJ, Lawrence Erlbaum, pp. 27–42.
- Humphreys, G.W., Riddoch, M.J., Quinlan, P.T., 1988. Cascade processes in picture identification. *Cognitive Neuropsychology* 5, 67–103.
- S.M. Kosslyn (1983), *Ghosts in the mind's machine: Creating and using images in the brain*. New York, Norton.
- LeFevre, J., Bisanz, J., Mrkonjic, L., 1988. Cognitive arithmetic: Evidence for obligatory activation of arithmetical-facts. *Memory and Cognition* 16, 45–53.
- LeFevre, J., 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., 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.
- Marsh, L.G., Maki, R.H., 1976. Efficiency of arithmetical operations in bilinguals as a function of language. *Memory and Cognition* 4, 459–464.
- Mandler, G., Shebo, B.J., 1982. Subitizing: An analysis of component processes. *Journal of Experimental Psychology* 111, 1–22.
- McCarthy, R.A., Warrington, E.K., 1988. Evidence of modality-specific meaning systems in the brain. *Nature* 334, 428–430.
- McClain, L., Shih Huang, J.Y., 1982. Speed of simple arithmetic in bilinguals. *Memory and Cognition* 10, 591–596.
- McCloskey, M., 1992. Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition* 44, 107–157.
- McCloskey, M., Alimonsa, D., Sokol, S.M., 1991. Facts, rule, and procedures in normal calculation: Evidence from multiple single-patient studies of impaired arithmetical fact retrieval. *Brain and Cognition* 17, 154–203.
- McCloskey, M., Caramazza, A., Basili, A., 1985. Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition* 4, 171–196.
- McCloskey, M., Harley, W., Sokol, S.M., 1991. Models of arithmetic fact retrieval: An evolution in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning Memory and Cognition* 17, 377–397.
- M. McCloskey, P. Macaruso and T. Whetstone (1992), The functional architecture of numerical processing mechanisms: Defending the modular model. In: J.I.D. Campbell (Ed.), *The nature and origin of mathematical skills*, Elsevier Science, North Holland.
- Noël, M.-P., 1991. Influence des systèmes de notation des nombres sur les mécanismes d'encodage et de traitements numériques. [Influence of national systems for numbers on the encoding mechanisms and numerical processes]. *L'Année Psychologique* 91, 581–607.
- Noël, M.-P., Seron, X., 1992. Notational constraints and number processing: A reappraisal of the Gonzalez and Kolers (1982) study. *The Quarterly Journal of Experimental Psychology* 45A, 451–478.
- A. Paivio (1986), *Mental representations: A dual-coding approach*. Oxford, Oxford University Press.
- Pynte, J., 1974. Readiness for pronunciation during the reading process. *Perception and Psychophysics* 16, 110–112.
- T. Shallice (1987), Impairments of semantic processing: Multiple dissociations, in: M. Coltheart, G. Sartori and R. Job (Eds.), *The cognitive neuropsychology of language*. London, Lawrence Erlbaum, pp. 111–127.
- Sokol, S.M., McCloskey, M., Cohen, N.J., Alimonsa, D., 1991. Cognitive performance and processes in arithmetic: Inferences from the performance of brain-damaged subjects. *Journal of Experimental Psychology: Learning Memory and Cognition* 17, 355–376.
- Vorberg, D., Blankenberger, S., 1993. Mentale repräsentation von zahlen [Mental representation for numbers]. *Sprache Kognition* 12, 98–114.