

On the magnitude representations of two-digit numbers

HANS-CHRISTOPH NUERK¹ & KLAUS WILLMES²

Abstract

In this review, we discuss how the magnitude of two-digit numbers is represented and argue that the magnitude of tens and units is represented separately at least in addition to an overall holistic magnitude representation. We start by introducing three model sketches: the holistic model, the decomposition model, and the hybrid model. We then provide an overview about the evidence against a purely holistic representation of two-digit numbers. Afterwards, we present four arguments why earlier findings favouring the holistic view may not be as conclusive as it seems and offer alternative explanations. We propose a more detailed model framework about how magnitude comparison of two-digit numbers may be performed. Finally, we suggest three hypotheses/questions which may guide future studies on two-digit number processing.

Key words: magnitude comparison, two-digit numbers, mental number line, holistic model, decomposed processing

¹ Dr. Hans-Christoph Nuerk, University Hospital RWTH Aachen, Section Neuropsychology, Department of Neurology, Pauwelsstr. 30, D-52057 Aachen, Germany; Phone: ++49 241-8089877 or -808 970; Fax: ++49 241-8082598 ; E-mail: nuerk@neuropsych.rwth-aachen.de

² University Hospital of the RWTH Aachen, Section Neuropsychology, Department of Neurology

Number magnitude representation

Number magnitude is activated automatically in various tasks when judgments about semantic number attributes (e.g. magnitude, parity) are required, however, even when no semantic attributes have to be processed, number magnitude gets activated nevertheless (e.g. Brysbaert 1995; Dehaene, Bossini, & Giroux, 1993; Dehaene, & Akhavein, 1995; Eger, Sterze, Russ, Giraud, & Kleinschmidt 2003; Fias, 2001; Fias, Brysbaert, Geypens, & d'Ydewalle, 1996; Fias, Lauwereyns, & Lammertyn 2001; Fischer, Castel, Dodd, & Pratt, 2003; Henik, & Tzelgov, 1982; Nuerk, Bauer, Krummenacher, Heller & Willmes, this issue; Nuerk, Iversen, & Willmes, 2004a; Nuerk, Wood, Willmes, in press a; see Moyer, & Landauer, 1967 and Restle, 1970, for early suggestions). Automatic number magnitude activation has been demonstrated in children as well (Berch, Foley, Hill, & McDonough-Ryan, 1999; Girelli, Lucangeli, & Butterworth, 2000; Rubinsten, Henik, Berger, & Shahar-Shalev, 2002) and recent data indicate that the automaticity of magnitude activation in normal children is different from the activation in children with developmental disorders like developmental dyscalculia (Rubinsten, & Henik, submitted), ADHD (Kaufmann, Delazer, Semenza, Willmes, & Nuerk, submitted) or children with visuo-spatial disorders (Bachot, Gevers, Fias, & Roeyers, this issue).

While it is virtually undisputed nowadays that magnitude is activated in various tasks, the nature and generality of magnitude activation is controversial in behavioral and brain imaging studies (for behavioral data, see Brysbaert, 1995; Dehaene, 1989; Dehaene, Dupoux, & Mehler, 1990, Greenwald, Abrams, Naccache, & Dehaene, 2003; Nuerk, Geppert, van Herten, & Willmes, 2002 a; Nuerk, Weger, & Willmes, 2001, 2002b, 2004 b, in press c; Ratinckx, Brysbaert, Fias, & Stevens, submitted, Stevens, Ratinckx, & Fias, 2003; Wood, Mahr, & Nuerk, this issue; for fMRI data Chochon, Cohen, van de Moortele, & Dehaene, 1999; Eger et al., 2003; Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003; see Dehaene, Piazza, Pinel, & Cohen, 2003 for a review). Here, we will be concerned mainly with the nature of magnitude representation, and in particular, the representation of two-digit numerals.

Three models of two-digit magnitude representation

In principal, one can discern three different models of two-digit magnitude representation which we would like to call the holistic model, the decomposition model, and the hybrid model (see Figure 1). We will introduce these three models first and then review the available evidence.

1. *The holistic model*: The holistic model assumes that we map two-digit numerals onto a holistic magnitude representation. Holistic in this context means that the base-10 structure of two-digit numbers in the Arabic number system (see Zhang, & Norman, 1995) is no longer retained when it comes to magnitude representation. Two different variants of the holistic model for two-digit numbers can be distinguished: linear coding with scalar variability and logarithmic coding with fixed variability (see the controversy of Brannon, Wusthoff, Gallistel, & Gibbon, 2001; Dehaene, 2001; Gallistel, Brannon, Gibbon, & Wusthoff, 2001). Both variants are designed to account for the problem size effect in

various tasks: larger numbers are slower and less accurately responded to than smaller numbers (see e.g., Brysbaert, 1995). The logarithmic coding fixed variability model assumes that the magnitude of larger numbers may be encoded more slowly, because magnitude is represented along the mental number line in a logarithmically compressed fashion. Logarithmic compression implies that larger numbers (or more specifically their logarithmic magnitude) are located closer together on the mental number line. Because the magnitude representations of all numbers are assumed to have equal variability, the overlap for larger numbers is relatively larger. This larger overlap leads to stronger interference with the magnitude representations of neighboring numbers. Therefore, magnitude comparisons of two larger numbers may be slower than magnitude comparisons of two smaller numbers with the same absolute distance. The linear coding – scalar variability notion of the holistic model assumes that numbers are mapped linearly onto the mental number line, but that their variability increases with size. Therefore, the overlap between the magnitude representations of two larger numbers may again be larger and, consequently, magnitude comparisons of larger numbers may be slower than those of smaller numbers.

2. *The decomposition model:* The decomposition model assumes that the magnitudes of the digits of two-digit numerals are represented separately. That is, two-digit numbers are no longer supposed to be mapped onto one single mental number line, but they may be mapped onto multiple mental number lines, namely one for each digit involved. Since participants can solve two-digit number comparisons for all kinds of numbers successfully from early school years on (Nuerk, Kaufmann, Zoppoth, & Willmes, 2004c), the mental number lines must then be labelled with regard to the place value of the represented digit. Different variants of the decomposition model may be discerned. One may assume that the representation of the single digits is digital, i.e. 37 is represented as $\{3\} * 10^1 + \{7\} * 10^0$ (cf. McCloskey, 1992, for similar suggestions or Ratinckx et al., submitted). Alternatively, one may assume that the representation of the individual digits is organized in much the same *analog* way as in the holistic model assumes for the overall two-digit number magnitude. Within the analog variant of the decomposition model, again, the linear coding - scalar variability hypothesis may be distinguished from the logarithmic coding - fixed variability hypothesis. Since even severely impaired patients are usually very good at approximately estimating a number of objects without counting them (for a review, see Delazer, & Bartha, 2001), there is no doubt that approximate magnitude representations of numbers in general and of two-digit numbers do exist. The decomposition model, however, questions that these approximate representations are used when participants have to deal with exact symbols of two-digit Arabic numbers.
3. *The hybrid model:* The hybrid model was shortly sketched by Nuerk et al. (2001). It assumes that two-digit numbers are represented in both ways, decomposed and holistically. I.e., besides the magnitude representations of the single digits, the approximate magnitude of the overall two-digit number is also activated. These representations may be activated in parallel and may activate and/or inhibit each other. Note, however, that following Occam's razor, the hybrid model is the most complex one and should only be favored when the two other models do not account for the data.

Reviewing the evidence for and against decomposed processing of two-digit numbers

In this section, we first review the evidence for holistic processing of two-digit numbers. However, we do this shortly, because such reviews can also be found elsewhere. In the second part, we review the evidence for decomposed processing in greater detail, because such a review of relatively new data from our and other labs has not been published so far. In the third section, we will give a couple of reasons why we think that previous evidence for two-digit numerical processing is not fully conclusive.

Evidence for holistic processing of two-digit numbers

In 1981, Hinrichs, Yurko and Hu rejected the simple place-value model when they investigated two-digit number processing. They found that the magnitude comparison response times decreased “as a logarithmic function of the absolute difference between the two numbers” and suggested that two-digit numbers are compared as integrated quantities. Dehaene and colleagues (1990; see also Dehaene, 1989) confirmed and extended these results and conclusions. They also found that RT diminished logarithmically with increasing distance from the standard.³ However, Dehaene and colleagues were aware that besides a holistic model based on an analog mental number line the so called interference model based on separate comparisons for tens and units captured “the essential features of [their] comparison data” (Dehaene et al., 1990, p. 634; cf. also Hinrichs, Yurko, & Hu, 1981, for early suggestions of the interference model). To distinguish between holistic and decomposed processing in their Experiments 1-3, Dehaene et al. conducted a fourth experiment in which they manipulated the SOA (-50ms, 0ms, +50ms) between presentation of the tens and units digit (see also Wood et al., this issue, for SOA data regarding two-digit magnitude comparison). They argued that the influence of units - as indexed by unit magnitude regression slopes - might be larger when the units appeared earlier. They found, however, that the regression coefficients capturing the influence of units did not differ significantly between the three SOA conditions. In the magnitude comparison with the standard 55, the units were coded by just deleting the decade value from the two-digit number for numbers smaller than 55. For numbers larger than 55, the units “were also included for the standard 1-9 by pairing the ones digits symmetrically with respect to 5 [4 with 6, 3 with 7 etc. ...]” (p. 629). These regression coefficients were negative in all three conditions, but did not differ significantly between the critical decade-first and the unit-first conditions.

Based on the assumption that “interference can play a role only when the units comparison finishes before the decades comparison; as long as the result of the units comparison is not available, it cannot bias the subject in any direction” (p. 633), Dehaene and colleagues concluded that these results are inconsistent with the interference model: They assumed that

³ Note, however, that a simple logarithmic function only explained the data best for the comparison with the standard 65. For standards 55 and 66, there were strong differences between within (e.g. 61_66) and between-decade trials (e.g. 59_66) that could not fully be accounted for by a simple logarithmic function. Dehaene and colleagues argued that the standards 55 and 66 were not representative because they were tie numbers while the only standard with successful fitting of the logarithmic curve, 65, was the representative one.

if the units of two-digit numbers are presented a little earlier than the decades, they are available earlier (for data questioning this seemingly trivial assumption, see Wood et al., in press a). When the units are available earlier, they should lead to more interference; i.e. the unit regression slopes should differ between unit-first and decade-first conditions. Since this was not the case, Dehaene and colleagues suggested that the magnitude of two-digit numbers is processed holistically rather than decomposed.

Evidence for decomposed processing of decade and unit magnitude

In this review, we present evidence for decomposed processing of tens and units from four different sources. First, we outline in detail, how the so-called unit-decade compatibility effect is inconsistent with the assumption of holistic processing for two-digit numbers. Since it was found in our lab, we address some questions about the compatibility effect that we have investigated in the last years and outline for the first time the conditions under which the effect can be found or not found. Second, we present priming data of Ratinckx and colleagues (submitted) that suggest that contrary to earlier claims, priming of two-digit numbers is – indeed – specific to the position of the decade and units and not only to overall magnitude. Third, we shortly describe decade crossing effects in the number bisection task that are also not consistent with a holistic point of view. Finally, we give a short review on carry-over effects in simple calculation which also follow the base-10-structure of Arabic numbers.

1. The unit-decade compatibility effect

Based on the studies mentioned above, it was widely accepted in the last years that two-digit number comparisons are based on an analog (or holistic) representation of magnitude (see, however, McCloskey, 1992, for a diverging view). In 2001, a first study questioned the holistictness assumption. Nuerk and colleagues (2001) used the *unit-decade-compatibility* (further also called compatibility) manipulation to investigate magnitude representation of two-digit numbers. In a two-digit number comparison task (e.g., 42 vs. 57, short 42_57) a given number comparison was defined as *unit-decade-compatible* if both comparisons between tens and units led to the same decision (e.g., for 42_57, both $4 < 5$ and $2 < 7$). In contrast, a number comparison was defined as *unit-decade incompatible*, if the two comparisons for units and tens led to different decisions (e.g., for 47_62; $4 < 6$, but $7 > 2$). In the above two examples, absolute overall distance is 15 in both cases. In addition, the different stimulus groups were constructed such that logarithmic distance, distance of the logarithms and mean absolute and logarithmic problem size did not differ between compatible and incompatible stimulus groups. If a single analog (holistic) magnitude representation would be sufficient, then no compatibility effect should be obtained, because distance and other measures did not differ between compatible and incompatible stimuli. However, Nuerk and colleagues consistently found compatibility effects for the Arabic notation: Incompatible number pairs were responded to more slowly than compatible number pairs in both, a participant-based and an item-based ANOVA (Nuerk et al., 2001, see also Nuerk et al., 2002b, 2004b, c, in press b). A number of objections could be brought forward against this initial finding. We outline the most important ones.

- ∄ *The compatibility effect may just reflect competition at a response or output level which is similar to common attentional congruity effects. It has nothing to do with the magnitude representation of two-digit numerals.*

Incompatible trials lead to faster responses and more errors than compatible trials in adults (Nuerk et al., 2001, in press b) and even in children from grade 2 on (Nuerk et al., 2004c). It is true that these main effects could be explained by a pure response competition account. However, the compatibility effect interacts with magnitude related information. With regard to RT, large unit distances (e.g. 38_51 vs. 41_58) usually lead to much greater compatibility effects than small unit distances (e.g. 42_53 vs. 43_52) even when overall distance is held constant between all respective stimulus groups. With regard to errors, small decade distances usually produce larger compatibility effects than large decade distances. In sum, both the magnitudes of the decade and the unit distances influence the compatibility effect in a specific way and in a reverse direction. Small decade distances and large unit distances tend to lead to the largest compatibility effects (Nuerk et al., 2001; 2004b, c, in press b). It may be true that part of the interference we observe in the compatibility effect is located at the response level, but it is specifically influenced by the separate magnitude (and distance) representations of the tens and the units in our study. In our view, this is hard to reconcile with a view that magnitude is holistically processed and that we are just looking at a pure response conflict. Rather, these interactions suggest that there is a specific influence of the (irrelevant but interfering) unit comparison on the relevant decade comparison. When the distance between the irrelevant units in incompatible trials is large the activation of a wrong response may be particularly pronounced and thus causes the strongest interference. When the decade distance is large and the decade comparison fast and easy, the units have less of a chance to interfere. Altogether, one could presume that stronger activations of the irrelevant response and weaker activations of the relevant response both lead to more interference in incompatible trials.

- ∄ *The compatibility effect is just a perceptual effect due to the perceptual organization of the original stimuli in a column-based (unit digit above unit digit; decade digit above decade digit) organization.*

This objection was raised immediately when we published the compatibility effect in 2001. However, the compatibility effect prevails when the stimuli are presented in a diagonal fashion rather than with a columnar perceptual alignment (Nuerk et al., 2004b; Ratinckx, Nuerk, van Dijk, & Willmes, in press). Additionally, compatibility even interacts with unit distance for German number words (which have an inverted order of units and tens information, e.g. 21 is “einundzwanzig”, i.e. one-and-twenty) where there is no column-based organization of single digits (Nuerk et al., 2002b).

In several yet unpublished follow-up experiments, we rather have observed that the opposite of the above objection is true. The compatibility effect becomes weaker as the predictability of the decades and the units is increased by the perceptual organization. For instance, in the original experiment, we had not used any mask and presented Arabic digits and number words in a randomized order. We were lucky, because this probably prevented participants from attending only to the relevant decade digit. In a later follow-up experiment for the above-mentioned SOA experiment in Dehaene’s study (Knops, Nuerk, & Willmes, 2003; Nuerk, Knops, & Willmes, submitted), we used an “XX” type mask, so that the mask perfectly determined the location of the decade and the unit digit

(the decade digit appeared where the left symbol of the mask appeared and the unit digit on the right). We observed no compatibility effect for an SOA of 0 ms. Because all stimuli were between-decade stimuli, we hypothesized that it is sufficient to concentrate on the 100% relevant decade digit. This may have diminished the influence of the irrelevant unit distance.

Behavioral data of two follow-up experiments confirmed this interpretation. When we repeated the same experiment with 50% within-decade trials and 50% between-decade trials, we observed a huge compatibility effect again. When the decade digits are not 100% valid for the magnitude comparison, but only valid in 50% of the trials, decade and unit magnitude are again both processed, because it does not make sense to direct attention only to the decade digit. Note that the stimuli and perceptual organization of the experiment were the same as before, the only change was the additional inclusion of the 50% within-decade trial stimuli so that the strategy to attend only to the decade digits did not make sense anymore.

Similarly, in a control condition for a TMS study about the compatibility effect (Nuerk, Knops, Dambeck, Foltys, & Willmes, 2004), we used a random dot pattern mask extending over a large part of the presentation screen. Again, we observed a large compatibility effect.

In sum, it is not the appearance of a mask, but the type of mask and how this mask can strategically be used which may conceal the compatibility effect. The compatibility effect is most likely to be masked under artificial and highly predictable perceptual conditions, namely perfect prediction of the location of the decade digit and 100 % relevance in case of many trials (e.g. more than 2000 in Knops et al., 2003). When the stimuli themselves or the validity of the decade digit is not predictable as in an ecological encounter with any given number, the compatibility effect is reliable. So, a focused attentional spotlight on highly predictable perceptual features can destroy the compatibility effect rather than produce it.

Hence, the compatibility effect has now been replicated in different labs (Nuerk et al., 2004c; Ratinckx et al., in press; Meeuwissen, personal communication), with different stimuli (Ratinckx et al., in press) and with different populations (Nuerk et al., 2004c; Wood, Nuerk, Freitas, Freitas, & Willmes, in press a) and it seems to be fairly stable. It is influenced by perceptual organization and attentional mechanisms, but it is not restricted to one certain perceptual organization. Part of its locus may be a response conflict but the opposite interactions with decade distance and unit distance do – in our view - show that the magnitudes of the single digits are processed. However, not only the digits, but also their place value is processed at a certain level because of the still accurate performance in incompatible trials: If the place value had not been computed and instead the comparison of the single (unit or decade) digits had been randomly computed for response, 50 % of the digit comparisons lead to the wrong decision in incompatible trials. However, even children in grade 2 can solve the task very accurately.

2. The place-value digit priming effect in two-digit numbers

Originally, priming data seemed to be in favor of an analog magnitude representation of two-digit numbers. Semantic distance effects for primes have been observed regularly (Brysbaert, 1995; Reynvoet, & Brysbaert, 1999). Reynvoet and Brysbaert (1999) specifically examined the effect of the first decade crossing with priming. They compared the effects of primes of equal distance to the target within and between decades (e.g. 7_9 vs. 11_9) and found no difference between these two types of trials. They concluded that unit-decade structures did not play any role in the representation of two-digit numbers. However, as Nuerk and coworkers (2001) have argued teen numbers are specific and very high-frequent numbers and their corresponding number words have – in most Western languages – not even a transparent ten-unit structure; therefore the results may not be representative.

Ratinckx and colleagues (submitted) conducted just such a priming study over the whole range of two-digit numbers in a naming task. They found a facilitating priming effect when prime and target shared a digit at the same position (e.g., primes 18 and 21 for the target 28). In contrast, an interference effect was observed when prime and target shared one digit at different positions of the place-value system (e.g. primes 82, 86, or 72 for target 28). Most intriguing, however, was the finding that even when primes and targets shared the same digits, but at different positions (e.g. prime 18 for target 81) there was an interference effect. So, clearly, the place-value of the constituent digits is processed and not only the digits themselves as reported in previous studies (Greenwald et al., 2003).

In sum, these results suggest that not only the overall magnitude of two-digit numbers and not only the magnitude of the constituent digits get activated, but that the magnitude of digits in relation to their position and value in the place-value Arabic number system is activated and represented.

3. Decade crossings in the number bisection task

A third piece of evidence is provided by a task which is usually used in the assessment of number magnitude representation in patients (e.g., Dehaene & Cohen, 1997). In the number bisection task, participant is asked to name the number which is exactly the numerical middle of two outer numbers (e.g., what is the numerical middle between 1 and 9?). In the verification version of this task, the participant is presented with three numerals in ascending order of magnitude and is asked whether the number displayed in the middle is really the numerical middle of the two outer numbers (e.g. 35_38_41) or not (35_39_41). The bisection task, originally designed to assess pure magnitude representation, is sensitive to numerical variables like the length of the bisection interval defined by the two outer numbers or the distance of the middle number to the numerical middle in “No” trials (34_35_46 is rejected faster than 34_39_46; Nuerk et al., 2002a). In addition, it is sensitive to other non-quantity related attributes of two-digit numbers like parity or whether or not the three numbers are part of a multiplication table (e.g. 21_24_27 is faster than 22_25_28).

Most important, however, for the current review is that this task is highly sensitive to decade crossings. In a regression analysis, we introduced a variable called decade crossing which was positive (+1) when the triplet crossed a decade (e.g. 35_38_41) and negative (-1) when all three numbers were within a decade, i.e. had the same decade digit (e.g. 32_35_38).

This variable became the strongest predictor in a regression analysis for non-bisectable trials (in which the middle number was not the numerical middle) and did explain more variance than quantity related variables like bisection range and distance to the middle (see above). For non-bisectable trials, decade crossing remained a strong predictor in the regression analysis, even when the covarying influence of the length of the bisection interval had already been partialled out. In sum, the decision whether or not another number is in the middle between two outer numbers is not based on one analog magnitude representation only, but is highly sensitive to the decade-unit structure of the digits involved. A strong influence of the decade crossing variable is not consistent with the idea of one simple holistic magnitude representation driving performance in the bisection task.

4. Carry-over effects in calculation

Another source of evidence for decomposed processing of two-digit numbers comes from studies on calculation. It has repeatedly been shown (e.g. Deschuyteneer, De Ramme-laere, & Fias, this issue) that additions or subtractions with carry-overs over the decade break are much harder to perform than without carry-overs. If they were performed only on the basis of an analog magnitude representation, such unit-decade-related effects should not influence performance when overall problem size effects are matched. Even for simple calculations the decomposed decade-unit structure of the Arabic place-value system thus influences performance.

In the study of Deschuyteneer and colleagues one can find another source of evidence against purely holistic processing. For trials without carry-overs, the problem size effect of overall magnitude disappeared when the magnitudes of the decade digits and unit digits were controlled separately. Trials with large decade digits and small unit digits (i.e., with a larger problem size) were not faster than trials with small decade digits and large unit digits (i.e., with a smaller problem size). Thus, in this analysis, not the overall problem size was important, but problem size of the individual digits determined performance.

Why earlier evidence for holistic processing may not be conclusive

There are at least four reasons why the good fit of logarithmic distance curves to the RT data as performed in previous studies may not be fully conclusive for the issue whether two-digit numerals are processed in a decomposed or in a holistic fashion: (i) stimulus selection, (ii) confounds between compatibility, overall distance, and decade distance, (iii) the issue of fitting conclusiveness, and (iv) automatic comparison of the available digits in a comparison task with a fixed standard.

(i) Stimulus selection

One simple, but already very important reason why Dehaene and colleagues (1990; see also Hinrichs et al., 1981; Dehaene, 1989) did not find significant differences in unit magnitude (or distance) regression slopes may have been the standard number they had used and

the relations of the standard to other numerals. When standards like 55 or 65 are used, unit distance between standard and target is restricted to a maximum of 4 (when decade numbers themselves which are high-frequent and special are hence not considered here). However, as outlined above, compatibility effects are usually much harder to find for small unit distances (e.g., 34_52, absolute unit distance $4 - 2 = 2$) and much easier to detect for large unit distances (e.g., 38_51, absolute unit distance $8 - 1 = 7$). In some of our experiments, the compatibility effects for small unit distances did not reach significance. The choice of a standard 65 (for which Dehaene found good logarithmic distance fitting) does – in principal – restrict the investigation of specific influences of unit magnitude representations to those stimulus conditions which are now known to most likely produce null effects when unit magnitude representation is examined. It would thus be better to choose standards like 52 and 58 in order to allow large compatible or incompatible unit distances (e.g. for numbers like 71, 79, 31, 39) for investigating the influence of unit magnitude. The restriction to unit distances which most likely produce null effects makes it hard to interpret these null effects conclusively.

(ii) Confusions between compatibility, decade distance and overall distance

In an experimental design with a fixed standard of 65, compatibility, unit distance, decade distance and overall distance are confounded unavoidably. Consider, for instance, the comparisons 41_65 and 49_65: The pair 41_65 is compatible, but also has a larger overall distance than 49_65. For fitting the responses within one-decade, therefore compatibility and overall distance are confounded. When the RT data are fitted with a (logarithmic) distance measure within that decade, overall distance may explain part of the variance which might actually be attributed to compatibility. Another possibility would be to compare (compatible) trials like 41_65 with the mean RT of incompatible trials like 36_65 and 46_65. In such a case, mean overall distance is matched. However, the average unit distance is 4 ($= 5 - 1$) for compatible trials and 1 for incompatible trials. As outlined in the above section, such a small unit distance is the most unfavorable condition to produce incompatibility interference. So, comparing such trials is not a good solution either. A final possibility would be to compare compatible trials like 41_65 with incompatible trials like 39_65. Here the overall distance is a little larger in the incompatible condition working against the compatibility effect. What is worse, in addition to overall distance, the decade distance of the incompatible condition is always larger than in compatible trials. So the compatibility effect would have to work against two distance effects, a slightly higher overall distance effect and a larger decade digit distance effect. Because the unit distance of 4 (which is maximal when a standard like 65 is used) is rather small, the compatibility effect will likely not be sufficient to overcome the influence of two distance effects working against it.

One can easily construct more examples like this, but one will always find that with one fixed standard like 65 such confusions are unavoidable. They can only – in part – be controlled when groups of stimuli are selected specifically and constructed such that the relevant variables do not differ on average between these groups of stimuli (which is a standard

method in word recognition research).⁴ For single stimuli, such confusions cannot be avoided. Because of all these confusions one cannot – in our opinion – conclusively resolve which numerical variable drives performance in a two-digit number comparison task and which does not.

(iii) The issue of fitting conclusiveness

The problem of fitting conclusiveness can be elaborated with the following example. Let us suppose, a participant (or an algorithm) uses perfectly decomposed processing of two-digit numbers in a number comparison task. RT performance is influenced by both, decade and unit distance. When the unit distance comparison is in the same direction as the decade distance comparison, its influence may be facilitating, when it leads to a different response, its influence may be interfering. Let us assume, for convenience, that this influence is linear – the bigger the distance, the stronger facilitation or interference. Let us further assume that the influence of the relevant decade digit distance on performance is much larger than the influence of the irrelevant unit distance on performance, let's say 10 times as large. For convenience, let us assume that this influence is linear, too.

What would happen, if we computed a linear regression analysis over RT data with the variables overall distance, decade distance and unit distance? The result would be very clear: If there would not be any noise, overall distance would explain 100 % of the variance; the variables decade distance and unit distance would not add any variance. Following the line of argument that is commonly used with such fitting algorithms, we would argue that the RT in our magnitude comparison task are linearly determined by absolute overall distance and that these results are consistent with a holistic (however, in this example linear) mapping of number magnitude onto one analog mental number line, although in the above example there was perfectly decomposed processing. We would conclude the opposite of what had happened in the production of the data.

The above example illustrates a general problem with fitting procedures that is not restricted to numerical processing. The “foremost criterion of model selection” (Myung, & Pitt, 1997, p. 80) is descriptive adequacy which implies that a model can be considered adequate only if it provides a good fit to the data. However, the best fit does not naturally imply that the respective model is the best or the true one (Myung, & Pitt, 1997; Schmidt-Weigand, 1999). The true model may not be included in the set tested. However, even if it is included, it may not provide the best fit because it may provide the best fit for data minus noise, but not for data plus noise. Myung and Pitt (1997, p. 81) illustrate the importance of this point and discuss differences between the Fuzzy Logic Model of Perception (FLMP) of Massaro and colleagues (Massaro, & Cohen, 1993; 1994; Oden, & Massaro, 1978) and the Linear

⁴ Note, however, that when overall distance is matched between compatible and incompatible stimulus groups, decade distance is necessarily a bit larger in incompatible stimulus groups. As overall distance is composed of 10 times decade distance *plus* unit distance for compatible trials and 10 times decade distance *minus* unit distance for incompatible trials, the decade distance for incompatible trials must be larger (+1 for 42_57 and +2 for 47_62, for a mathematical elaboration, see Nuerk et al., 2002). So, if participants only compared the decades they should have been faster in incompatible trials. As it is sufficient to concentrate on the decades when having to compare between-decade two-digit number comparisons, it is even more remarkable that the compatibility effect usually overrides this decade distance effect.

Integration Model (LIM) of Anderson (1981). Myung and Pitt (1997) claim that the FLMP even fitted data patterns produced by the LIM better than the LIM itself (see also Cutting, Bruno, Brady, & Moore, 1992; Massaro, & Cohen, 1993). Because one can always fit any data set perfectly, any fitting model is by itself nothing more than statistical mimicking (Ratcliff, 1988; Van Zandt, & Ratcliff, 1995).

What does this mean for our case of fitting two-digit number comparison data? In the above example, a comparison process based on a linear holistic magnitude representation and a comparison process based on decomposed magnitude would produce exactly the same data pattern because overall distance = $10 * \text{decade distance} + 1 * \text{unit distance}$ (when unit distance is defined to be negative for incompatible trials with the smaller two-digit number having the larger unit digit). So, when both decade distance and unit distance would be represented and would explain some part of the variance in the data, overall distance measures would always be the strongest predictors, because they include variance of both, the decade distance and the unit distance predictors as well as possibly the variance of a holistic magnitude representation. This inconclusiveness of such fitting patterns does in our opinion not change substantially when logarithmic rather than linear overall distance measures become predictors. In almost all cases investigated, we would have explained more variance if we had forced logarithmic (or linear) decade distance and unit distance as predictors in a regression analysis than if we had only used logarithmic overall distance. That separate predictors for decade distance and unit distance explain more variance than one single predictor for (logarithmic) overall distance, however, does not imply that the decomposed model is the true model. Models with more free parameters tend to explain more variance than models with only one free parameter.

In sum, the results of such regression analyses or other fitting procedures to the data are in our opinion not conclusive for deciding between the holistic, the decomposition, and the hybrid model of two digit number processing. Mostly, very similar amounts of variance can be explained by the holistic and the decomposed processing assumptions. Other methods like the isolated variation of the compatibility attribute must be developed to decide which of the two models is true.

(iv) Automatic comparison of available digits

We know that the magnitude of single digits is automatically activated (e.g. Pavese & Umiltà, 1999). However, in this special issue, we have two sources of evidence that not only the magnitude of the constituent digits are automatically activated, but that such magnitudes are automatically compared to each other even when these magnitude comparisons are totally irrelevant and useless for the task at hand.

Wood and colleagues (this issue) investigated how the compatibility effects are altered when the single digits of the two two-digit numbers are presented at different SOAs. The task in both experiments of their study was the same as in the other studies described: to compare the two two-digit numbers and to indicate by button press which is larger. In Experiment 1, SOA was manipulated such that the unit digit of one number and the decade digit of the other number were presented at the same SOA (e.g., in 37_52 the digits 3 and 2 and the digits 7 and 5 were presented at the same SOA). They found that the compatibility effect decreased. Further analyses showed that the magnitude relation between the digits appearing

together strongly determined performance. When the digit comparison would have led to the same response as the overall comparison of the two-digit numbers, responses were fast and otherwise slow. In the above example, this is not the case: In the comparison of 3 and 2, 3 is the larger digit, but it is part of the smaller number. The same holds for the comparison of 5 and 7. In the SOA study of Wood and colleagues, these digit comparisons had stronger influence on performance than the compatibility effect.

Of particular interest with regard to two-digit magnitude comparisons with fixed standards is Experiment 2 of Wood and colleagues: There, they made the same SOA manipulation as in Experiment 1, but now the digits of one two-digit number were presented at the same SOA. So in the above example the digits 3 and 7 and the digits 5 and 2 were presented together at the same SOA. The compatibility effect in this study was reversed: Again, subsequent analyses showed that this could be explained by a digit comparison of the digits appearing together. In the above example, the results of this digit comparison process are congruent with the response. The decade digit 3 is smaller than the decade digit 7 and it is also the smaller decade digit in the relevant (decade) digit comparison process in the two-digit number magnitude comparison task. The same holds for the digit comparison process of 5 and 2. In sum, in the incompatible trial 37_52 both digit comparisons led to congruent responses.

Consider now a compatible trial like 42_57 in the same experimental setting. Here both digit comparisons would lead to responses incongruent with the correct response. The decade digit 4 is *larger* than the unit digit 2, but it is the *smaller* decade digit in the relevant decade digit comparison process, while the decade digit 5 is *smaller* than the unit digit 7, but the *larger* decade digit. When the relationship between unit-decade compatibility and the congruity is analyzed over all items, one can find a negative correlation between compatibility and the above congruity effect. This implies that if congruity of the magnitude comparison of digits appearing together at the same SOA influences responses as indicated in the Wood et al. study, it works against the compatibility effect. Indeed, the unit-decade compatibility effect was inverted in Experiment 2 of the Wood et al. study, but became regular after the digit congruity effects were partialled out in an ANCOVA.

This finding has clear implications for a fixed comparison with a number like 65. First, consider the number 49. 49_65 is an incompatible two-digit number comparison because $4 < 6$ but $9 > 5$. With regard to the above mentioned congruity effects, the relation between its constituent digits (appearing together like in Experiment 2 of Wood et al.) is congruent. The decade digit 4 is *smaller* than the unit digit 9 and also *smaller* than the relevant other decade digit 6. Secondly, now consider the comparison with number 51. 51_65 is a compatible two-digit number comparison because $5 < 6$ and $1 < 5$. With regard to the above mentioned congruity effects, the relation between its constituent digits (appearing together like in Experiment 2 of Wood et al.) is incongruent. While the decade digit 5 is *larger* than the unit digit 1, it is *smaller* than the relevant other decade digit 6. When the standard 65 is chosen, unit-decade compatibility and congruity of the digits appearing together are inversely related in almost all trials. Their effects work against each other. It is therefore well conceivable, that they cancel each other out as indicated by the Wood et al. data. Thus, a compatibility effect in number comparison studies with a fixed standard could not be observed under such circumstances.

In sum, when participants see digits in a magnitude comparison task, they sometimes cannot help but comparing these digits, even when this comparison is task-irrelevant. The congruity of this comparison is inversely related to unit-decade compatibility. When only two digits of one two-digit number are presented in a fixed comparison study, this comparison may therefore interfere with the unit-decade compatibility effect. When two two-digit numbers are presented at the same time in a variable comparison study, six digits comparisons are possible. The influence of the two decade-unit digit comparisons within the two numbers is much smaller unless it is enhanced by special SOA presentation modulations as in Wood et al.'s study.

A new model framework for two-digit number processing

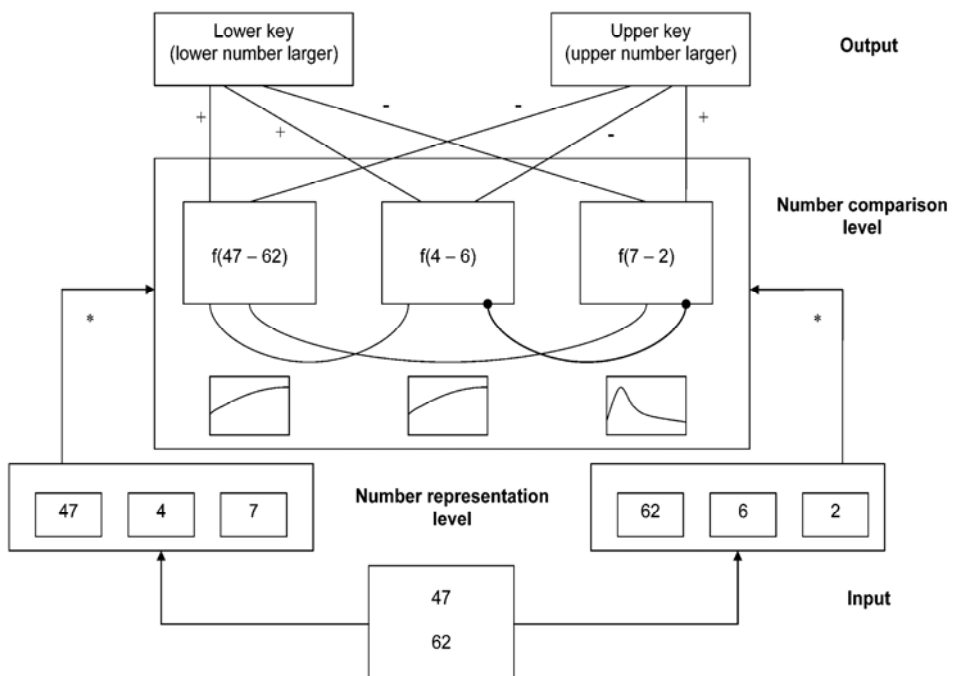
In this last section, we would like to suggest that the magnitude comparison of two-digit numbers might be accounted for by separate but interactive comparison processes for decade magnitude, unit magnitude and eventually overall magnitude (cf. McClelland, & Rumelhart, 1981, for an interactive model). The idea is that although the comparison of decade magnitude and of overall magnitude is sufficient to make a decision, unit comparisons nevertheless also influence responses until their activation decays. So, a compatible unit comparison would add additional activation to the correct response key and thus facilitate responses while an incompatible unit comparison would inhibit the correct response key and thus slow down responses. Of course, the decade comparison has to have stronger connection weights to the final output representation, because incompatible trials are nevertheless responded to correctly.

Activation of both comparisons (decade and unit comparison) may accumulate faster when the distance between the two numbers is large and slower when the distance is small. With regard to decade comparisons, this would lead to faster decisions for large decade distances and to slower decisions for small decade distances (i.e. to the ordinary distance effect). With regard to the unit comparisons, this leads to greater incompatibility interference for large unit distances and to smaller incompatibility interference for small unit distances. Such a model would also produce an overall distance effect because overall distance is comprised of 10 times decade distance and unit distance and both play a role in this model framework. Eventually, an overall comparison may additionally be needed to account for the exact parametric fitting of a logarithmic distance effect.

As an SOA study (Knops et al., 2003) shows, there is a certain time window in which unit comparisons may play a significant role. When the decade digits are presented long enough before the unit digits, the decade comparison (and the overall comparison) may be largely finished before the unit comparison really starts to inhibit or facilitate correct responses. However, when the unit comparison starts before the decade comparison, the response is first determined by the result of the unit comparison before the decade comparison (and eventually the overall comparison) starts to activate the correct response. However, after a while (at larger SOAs) the unit comparison may be identified as irrelevant and its activation may decay or be inhibited. Therefore, the compatibility effect disappears again at larger SOAs because the facilitating or interfering activation of the unit comparison is no longer influential.

Figure 1:

Model framework for modeling the compatibility effect and its SOA modulations by way of an example. The model framework depicts 4 levels: an input level, a number representation level, a number comparison level and an output level. The input of two digit numbers is organized into three different representations: overall magnitude, decade magnitude, and unit magnitude, for both presented numbers. All three representations are compared automatically and in parallel with the respective representations of the other number at the magnitude comparison level. The asterisk (*) at the arrows from the representation to the comparison level indicates that three arrows - one from each of the three different representations to each of the three magnitude comparisons - should be in the graph, however, for visibility only one arrow is depicted. The activation of the magnitude comparison is a monotonically increasing function of the actual (not the absolute) distance between the two numbers which may be linear, linear with scalar variability depending on the size of the two numbers, logarithmic, or the distance of the logarithms with some constant added such that a positive distance activates responses to the upper key and inhibits responses to the lower key and vice versa for a negative distance. The model framework postulates reciprocal, but asymmetric inhibition (incompatible) or facilitation (compatible) between the decade comparison and the unit comparison which is dominated by the decade comparison activation. If the unit comparison is activated first (because units appear first in the task), the unit comparison starts, but decays after some time: This would lead to higher interference when the unit comparison starts shortly before the decade comparison, but less interference when units are presented long before decades. It is not clear whether an overall analog representation and comparison also play a role, but there is no reason to deny their influence.



How do the above number-related attentional processes fit into such a model framework? Attention might be assumed to facilitate the selection of the relevant comparison (i.e., the decade comparison) and to inhibit the irrelevant comparison (i.e., the unit comparison). This view could be incorporated in such a model by attentional modulation of the decade and unit comparison activation. The stronger the attentional cues help to select the relevant decade comparison and to inhibit the irrelevant unit comparison, the stronger is the pre-activation of the relevant decades and the weaker is the pre-activation of the irrelevant unit comparison. Such pre-activation might then be changed by the informativeness of the attentional cue much like in a standard Posner task. When the cue is fully (100%) informative as in this study, participants may strongly attend to the decade digit and thus produce a null or even a reverse compatibility effect. When the cue is informative in a much smaller percentage of the trials, pre-activation may be less pronounced. Similar arguments can be made when the visual salience of the attentional cue is varied.

Such a model would represent a hybrid approach in between the models of McCloskey and colleagues (McCloskey, 1992) and the model of Dehaene (Dehaene, 1992; Dehaene, & Cohen, 1995; 1997). As outlined in Nuerk et al. (2001), tens and units have separate access to magnitude representation or separate magnitude representations themselves thus confirming the ideas of McCloskey (1992) that 37 would be represented as {3} 10EXP1, {7} 10EXP0. However, the comparison at each level is analog thus confirming the idea of Dehaene and colleagues (Dehaene, 1989; Dehaene et al., 1990): Moreover, we have yet no reason to deny an important role of an overall magnitude representation in two-digit number comparison as suggested by Dehaene. An analog approximate magnitude representation can be used for magnitude estimation: The functional activation in magnitude estimation is located in similar brain regions as functional activation in magnitude comparison (Pinel, Le Clec'h, van de Moortele, Naccache, Le Bihan, & Dehaene, 1999; Simon, Mangin, Cohen, Le Bihan, & Dehaene, 2002) and there is no reason yet to suggest that it is not used in two-digit number comparison. At the moment we therefore suppose that this overall approximate magnitude representation is used in addition to separate representations of tens and units as suggested by the compatibility effect. Finally, the suggested model framework does not yet contain reciprocal inhibition between decade and unit digits at the representation level. However, recent data from Stevens and colleagues (2003) suggest that a smaller distance between the constituting digits of two-digit numbers (i.e., when they are closer together on the mental number line) may have deteriorating effects on performance. For instance, on average, a number like 23 is named slower than the number 26 because the distance between 2 and 6 in 26 is larger thereby evoking less interference between the constituting digits. If this result holds, even reciprocal inhibition from the separate representation of decade and unit digits on the representation level would have to be assumed.

Three hypotheses about two-digit number representation

(i) Pure decomposed processing of two-digit numbers

For numerical estimation and approximation tasks, there is strong evidence that an approximate magnitude representation is needed. Therefore, in our model, we assume that it also plays a role for two-digit Arabic number processing. However, Arabic numbers are

exact representations. As we have outlined above, the data indicating an overall logarithmic magnitude representation could – in principle – be explained also by the appropriate composition of decade and unit digits and the representations of their magnitude. A good example is the missing overall problem size effect in the study of Deschuyteneer and colleagues (this issue). Without a carry-over effect, it is not the overall magnitude that determines performance in their study, but the magnitudes of the constituent digits. Trials with large decade digits and small unit digits (i.e. with a large overall problem size) are not different from trials with small decade digits and large unit digit (i.e. with a small overall problem size).

Thus, overall problem size and distance effects could well be explained by problem size effects of the individual decade and unit digits. As there is now evidence for decomposed processing of tens and units, one might use Occam's razor to question whether an analog magnitude representation is needed additionally to account for the studies examining the representations of two-digit numbers. The simplest pure decomposition hypothesis might therefore be that *for Arabic numbers only the magnitudes of the constituent digits are processed*. They might have different activation weights for different place-values (because after all, we can successfully solve two-digit comparisons even in incompatible trials), but there might be no need for an approximate analog representation in this study. Future studies should – in our opinion – examine, if an impact of overall magnitude activation beyond the digit level can really be demonstrated in two-digit Arabic number comparison studies.

(ii) Is there a substantial difference between two-digit and other multi-digit magnitude representations?

Marc Brysbaert (personal communication) postulated that nobody would assume that there is an overall magnitude representation for all existing numbers, and that namely the question whether or not multi-digit numbers are processed in an analog fashion is – in principle – restricted to two-digit numbers. If this is the case, we should find substantial differences, for instance in the compatibility effect or in digit priming effects between two- and three-digit numbers. In line with hypothesis (i), one might postulate that there is *no substantial difference between the representations of two- and three-digit numbers*. Of course, three-digit number processing is more complex; it requires at least three digit magnitude representations, the issue of spatial resolution of the constituent digits is more complex, and for sure, more working memory resources are needed. However, these questions are questions about the impact of task demands or work loads, but not questions of a strict qualitative difference between two- and three-digit number processing. Such differences have yet to be shown.

(iii) The parietal lobes as a digit integration system?

Originally, it has been assumed that the parietal lobes are the locus of the magnitude representation system. In a seminal paper, Dehaene and colleagues (2003) have recently integrated different findings in the last years and separated the location of an analog magnitude system (the intraparietal sulcus IPS) from the location of another system (superior parietal regions). If two-digit numbers are represented in a decomposed fashion and the different digits are differentially evaluated with respect to their position in a place-value system, there

must be a structure where this is accomplished. The value of an Arabic digit depends critically on its spatial position in the Arabic place-value system. Because the parietal lobes have been shown to subserve both magnitude and spatial representation, they seem to be particularly suited for the integration of two-digit numbers into a two-digit magnitude representation. Thus, while the identification of the single digits in a multi-digit display is probably performed by the visual number form area (BA 19/37), their integration into a differential place-dependent value system is presumably performed by parietal regions.

References

- Anderson, N. H. (1981). *Foundations in information integration theory*. New York: Academic Press.
- Bachot, J., Gevers, W., Fias, W., & Roeyers, H. (this issue). Number sense in children with visuospatial disabilities: Orientation of the mental number line. *Psychology Science, Special Issue "Brain and Number"*.
- Berch, D. B., Foley, E.J., Hill, R. J., & McDonough-Ryan, P. (1999). Extracting parity and magnitude from Arabic numerals: Developmental changes in number processing and mental representation. *Journal of Experimental Child Psychology*, 74, 286-308.
- Brannon, E.M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number. *Psychological Science*, 12, 238-243.
- 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.
- Chochon, F., Cohen, L., van de Moortele, P. F., & Dehaene, S. (1999). Differential contributions of the left and right inferior parietal lobules to number processing. *Journal of Cognitive Neuroscience*, 11, 607-630.
- Cutting, J. E., Bruno, N., Brady, N. P., & Moore, C. (1992). Selectivity, scope, and simplicity of models: A lesson from fitting judgments of perceived depth. *Journal of Experimental Psychology: General*, 121, 364-381.
- Dehaene, S. (1989). The psychophysics of numerical comparison: a reexamination of apparently incompatible data. *Perception and Psychophysics*, 45, 557-566.
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44, 1-42.
- Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, 16, 16-36.
- Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 314-326.
- Dehaene, S., & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, 1, 83-120.
- Dehaene, S., & Cohen, L. (1997). Cerebral pathways for calculation: double dissociation between rote verbal and quantitative knowledge of arithmetic. *Cortex*, 33, 219-250.
- 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., 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.
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20, 487-506.

- Delazer, M., & Bartha, L. (2001). Transcoding and calculation in aphasia. *Aphasiology*, 15, 649-679.
- Deschuyteneer, M., De Rammelaere, S., & Fias, W. (this issue) The addition of two-digit numbers: Exploring carry versus no-carry problems. *Psychology Science, Special Issue "Brain, & Number"*.
- Eger, E., Sterze, P., Russ, M. O., Giraud, A.L., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, 37, 1-20.
- Fias, W. (2001). Two routes for the processing of verbal numbers: Evidence from the SNARC effect. *Psychological Research/Psychologische Forschung*, 65, 250-259.
- 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.
- Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal Representation of Symbolic and Nonsymbolic Magnitude. *Journal of Cognitive Neuroscience*, 15, 47-56.
- Fias, W., Lauwereyns, J., & Lammertyn, J. (2001). Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Cognitive Brain Research*, 12, 415-423.
- Fischer, M., Castel, A., Dodd, M., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, 6, 555-556.
- Gallistel, C. R., Brannon, E.M., Gibbon, J., & Wusthoff, C.J. (2001). "Subtracting pigeons: Logarithmic or linear?": Reply. *Psychological Science*, 12, 247.
- Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology*, 76, 104-122.
- Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003) Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 235-247.
- Henik, A., & Tzelgov, J. (1982) Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory and Cognition*, 10, 389-395
- Hinrichs, J. V., Yurko, D. S., & Hu, J. M. (1981). Two-digit number comparison: Use of place information. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 890-901.
- Kaufmann, L., Delazer, M., Semenza, C., Willmes, K., & Nuerk, H.-C., (submitted). A broad examination of elementary and complex number processing skills in ADHD children.
- Knops, A., Nuerk, H. C., & Willmes, K. (2003). Die semantische Größenrepräsentation: Existiert mehr als nur ein numerischer Zahlenstrahl? In: J. Golz, F. Faul, R. Mausfeld (Hrsg.). *Experimentelle Psychologie. Abstracts der 45. Tagung experimentell arbeitender Psychologen* (p. 105), Regensburg: Roderer.
- Massaro, D. W., & Cohen, M. M. (1993). Perceiving asynchronous bimodal speech in consonant-vowel and vowel syllables. *Speech Communication*, 13, 127-134.
- Massaro, D. W., & Cohen, M. M. (1994). Visual, orthographic, phonological, and lexical influences in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1107-1128.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88, 375-407.
- McCloskey, M. (1992). Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition*, 44, 107-157.
- Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, 215, 1519-1520.

- Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin and Review*, 4, 79-95.
- Nuerk, H.-C., Bauer, F., Krummenacher, J., Heller, D., & Willmes, K. (in press a) The Power of the Mental Number Line: How the Magnitude of Unattended Numbers affects Performance in an Eriksen Task. *Psychology Science*, Special Issue "Brain and Number".
- Nuerk, H.-C., Geppert, B. E., van Herten, M., & Willmes, K. (2002a). On the impact of different number Representations in the number bisection task. *Cortex*, 38, 691-715.
- Nuerk, H.-C., Iversen, W. & Willmes, K. (2004a). Notational modulation of the SNARC and the MARC (Linguistic Markedness Association of Response Codes) Effect. *Quarterly Journal of Experimental Psychology: A*, 57, 835-863.
- Nuerk, H.-C., Kaufmann, L., Zopoth, S., & Willmes, K. (2004c). On the development of the mental number line. More or less or never holistic with increasing age. *Developmental Psychology*. 40, 1199-1211.
- Nuerk, H.-C., Knops A., & Willmes, K. (submitted). Is numerical processing analog? How decomposed two-digit number processing can be affected by strategy. Manuscript submitted for publication.
- Nuerk, H.-C., Knops, A., Dambeck, N., Foltys, H., & Willmes, K. (2004b). Is magnitude representation in the HIPS? A TMS study of two-digit number processing. 22nd European Workshop of Cognitive Neuropsychology, Bressanone, 25.-30. 1. 2004.
- Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition*, 82, B25-33.
- Nuerk, H.-C., Weger, U., & Willmes, K. (2002b). A unit-decade compatibility effect in German number words. *Current Psychology Letters: Behaviour, Brain, & Cognition*, 2, 19-38.
- Nuerk, H.-C., Weger, U., & Willmes, K. (2004b). On the perceptual generality of the unit-decade compatibility effect. *Experimental Psychology*, 51, 72-79.
- Nuerk, H.-C., Weger, U., & Willmes, K. (in press b). Language effects in magnitude comparison: Small, but not irrelevant. *Brain and Language*.
- Nuerk, H.-C., Wood, G. & Willmes, K. (in press a). The universal SNARC effect: The association between magnitude and space is amodal. *Experimental Psychology*.
- Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, 85, 172-191.
- Pavese, A. & Umiltà, C. (1999). Symbolic distance between numerosity and identity modulates Stroop interference. *Journal of Experimental Psychology: Human Perception and Performance* 24, 1535-1545.
- Pinel, P., Le Clec'h, G., van de Moortele, P. F., Naccache, L., Le Bihan, D., & Dehaene, S. (1999). Event-related fMRI analysis of the cerebral circuit for number comparison. *NeuroReport*, 10, 1473-1479.
- Ratcliff, R. (1988). A note on mimicking additive reaction time models. *Journal of Mathematical Psychology*, 32, 192-204.
- Ratinckx, E., Brysbaert, M., & Fias, W., & Stevens, M. (submitted). The mental representation of two-digit Arabic numerals examined with masked priming: 28 rapidly activates $\{2\} \times \{10\} + \{8\}$ in number naming. Manuscript submitted for publication.
- Ratinckx, E., Nuerk, H.-C., van Dijk, J.-P., & Willmes, K. (in press). Effects of interhemispheric communication on two-digit Arabic number processing. *Cortex*
- Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, 83, 274-278.
- Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, 72, 191-201.

- Rubinsten, O., & Henik, A. (submitted). Automatic activation of internal magnitudes: A study of developmental dyscalculia.
- Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology*, 81, 74-92.
- Schmidt-Weigand, F. (1999). Analyses of Single Words' Response Time Distributions: Applying Alternative Methods to Common Questions of Visual Word Recognition. Unpublished Diploma-Thesis: Philipps-University Marburg, Marburg, Germany.
- Simon O., Mangin J. F., Cohen L., Le Bihan D., & Dehaene S. (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. *Neuron*, 33, 475-87.
- Stevens, M., Ratinecx, E., & Fias, W. (2003). Naming of two-digit numbers: Position-based priming effects of constituting digits. 3rd Aachen-Ghent Brain, & Number Workshop, Aachen, 10.-11.6.2003.
- Van Zandt, T., & Ratcliff, R. (1995). Statistical mimicking of reaction time data: Single-process models, parameter variability, and mixtures. *Psychonomic Bulletin and Review*, 2, 20-54.
- Wood, G., Mahr, M., & Nuerk, H.-C. (this issue). Deconstructing and reconstructing the base-10 structure of Arabic numbers. *Psychology Science*, Special Issue "Brain and Number".
- Wood, G., Nuerk, H.-C., Freitas, G., Freitas, P. & Willmes, K. (in press a). What do semi-illiterate adults know about two-digit Arabic numbers? *Cortex*
- Zhang, J., & Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition*, 57, 271-295.

Authors' note

This research was supported by funding from the Interdisciplinary Group for Clinical Research, IZKF, of the RWTH Aachen and from the DFG grant KFO 112/TP2 to Klaus Willmes and Dieter Heller supporting Hans-Christoph Nuerk and by the EU funded Marie Curie Research Training Network NUMBRA. We would like to thank Tobias Versin and Sabrina Clemens for their assistance in preparing this article. Correspondence should be addressed to Dr. Hans-Christoph Nuerk, University Hospital RWTH Aachen, Department of Neurology – Section Neuropsychology, Pauwelsstr. 30, D-52057 Aachen, Germany (nuerk@neuropsych.rwth-aachen.de).