krajcsi_202_24

KRAJCSI ATTILA

MENTAL REPRESENTATION FOR SIMPLE SYMBOLIC NUMBER PROCESSING

(AZ ELEMI SZIMBOLIKUS SZÁMFELDOLGOZÁS MENTÁLIS REPREZENTÁCIÓI)

AKADÉMIAI DOKTORI ÉRTEKEZÉS

BUDAPEST 2024

Table of Contents

Acknowledgments

First, I thank my wife, Réka, for her support in this challenge. She had to bear much more and much longer than I would have considered reasonable. I truly appreciate the assistance of Máté and Dávid, who did not cause too much trouble too often.

I appreciate the countless hours of discussion and work with many of my previous and current students and coauthors, who had many contributions to these works. I especially thank Petia Kojouharova and Gábor Lengyel for their continuous cooperation in these projects. Ákos Laczkó and Tamás Szűcs also made essential contributions to several of these works.

Many colleagues and friends influenced my work on a more general level. First, friends (who are also colleagues) from my first workplace at University of Szeged were essential: József Fiser, Ágnes Lukács, Zsuzsa Molnár, Dezső Németh, Mihály Racsmány, Tamás Sefcsik, Roland Tisljár (for those who wonder, the order is set alphabetically) (and for those who wonder, some of these friends were affiliated to the University of Szeged in an organizational level, while some of them in spirit). After the very active, turbulent, and enjoyable years at Szeged, a much calmer environment at ELTE provided the opportunity to focus on research even more strongly. I highlight the discussions with Ildikó Király and Zoltán Nádasdy there. I am also grateful for my previous mentors, Ilona Barkóczi, László Mérő, and Anikó Kónya, when I was a student at ELTE and later, Csaba Pléh at Szeged.

The work of this dissertation was influenced by a continuous discussion with many colleagues working in numerical cognition. I thank many of them for the scientific discussion and input: Daniel Algom, Daniel Ansari, Dana Chesney, Krzysztof Cipora, Stanislas Dehaene, Arianna Felisatti, Martin Fischer, Mateusz Hohol, Matthew Inglis, Veronique Izard, Craig Leth-Steensen, Melissa Libertus, Oliver Lindemann, Marie-Pascale Noël, Hans-Christoph Nuerk, Darko Odic, Manuela Piazza, Jérôme Prado, Bert Reynvoet, Francesco Sella, Emily Slusser, Tom Verguts, and Marco Zorzi.

This work is dedicated to my mom.

Chapter 1 Introduction

Introduction

Understanding numbers and, more generally, mathematics is fundamental not only in formal education but in everyday life (Butterworth, 1999; Dehaene, 1997; Garcia-Retamero et al., 2019). A cornerstone in cognitive science when studying mathematical thinking is finding the key representations supporting number understanding.

In the last decades, probably the most dominant research topic in numerical cognition has been the role and properties of an evolutionarily old and simple representation that may be a key to understanding numbers and high-level mathematical intuition. The present thesis discusses why this dominant model may be incorrect and provides an alternative account.

The approximate number system (ANS)

A representation that became one of the most widely studied mental components of numerical cognition was first described in 1967. In their seminal paper, Moyer and Landauer (1967) asked the participants to compare single-digit Arabic numbers. The participants indicated which of the two numbers was larger. A very simple effect, the **distance effect**, was described: The larger the numerical distance between the two values, the easier the task was in terms of error rate and reaction time. The analyses showed that the behavioral performance depended not simply on the numerical distance of the values but rather on the ratio of the two values. This **ratio effect** was interpreted as number comparisons obeying the psychophysical Weber's principle and thus revealing an evolutionarily old and simple representation that supports mathematical thinking even when working with symbolic numbers. While common sense assumes that mathematical thinking is a high-level, human-specific, cultural construct, the discovery of the comparison ratio effect suggests that mathematical thinking is rooted in an evolutionarily old system. No wonder that this important and striking finding was published in Nature.

After the initial description of the comparison ratio effect, a flood of empirical works confirmed the idea that this simple number representation exists not only in human adults but in human infants (Feigenson et al., 2004; Libertus & Brannon, 2010; Odic & Starr, 2018), and in non-human animals (Brannon & Merritt, 2011; Parrish & Beran, 2022).

A simple possible implementation of this representation – that latter was coined as the **approximate (or analog) number system (ANS)** – is a linear representation where the number representation is noisy, and the noise is proportional to the value (Figure 1) (Dehaene, 2007). In other words, the larger the number is, the noisier the representation is. According to the model, when comparing two numbers, the performance depends on the overlap of the two represented numbers (Dehaene, 2007).

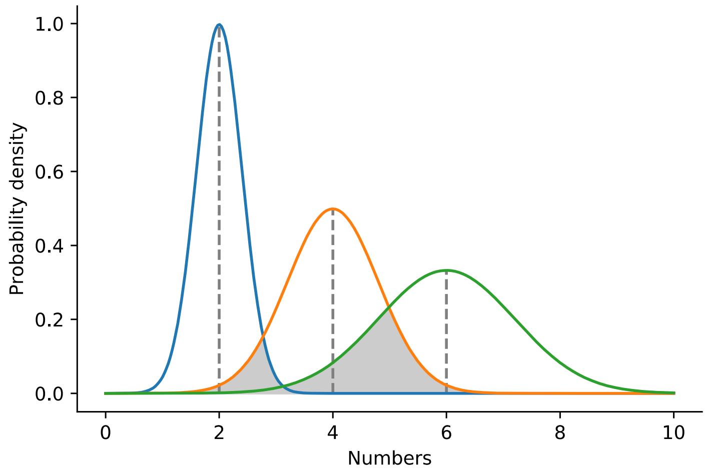


Figure 1 A possible formulation of the hypothesized number representation in the ANS account.

This simple representation later became the cornerstone of numerical cognition. First, it could account for a series of cognitive effects and phenomena. As explained above, it can explain the ratio effect. It can also explain the distance effect, where the distance effect is considered as a byproduct of the ratio effect. Similarly, the **size effect** (comparison of larger values is more difficult than the comparison of smaller values) can also be understood as a byproduct of the ratio effect. In other words, assuming a ratio effect, distance and size effects also must appear because distance and size properties correlate with the ratio property of the number pairs. Consequently, comparison ratio, distance, and size effects can be accounted for by the same mechanism.

As another example, the ANS may account for the interference between numerical and spatial information. A prominent example of numerical-spatial interference is the **SNARC effect** (Spatial-Numerical Association of Response Codes): When deciding about the parity of a number (i.e., is the number even or odd), participants respond faster to larger values with the right response button than with the left response button, and they respond faster to smaller values with the left response button than with the right response button (at least in left-to-right reading cultures) (Dehaene et al., 1993). The ANS account of the SNARC effect assumes that the ANS has a spatial orientation as if it was placed from left to right, and this spatial property interferes with the response location (Dehaene et al., 1993).

A next example of ANS accounting for a cognitive effect is the **priming distance effect**. For instance, when a participant decides if a single digit number is smaller or larger than five, the performance is influenced by the previous trial, and the strength of the influence depends on the numerical distance between the values of the previous and actual trials. The ANS account assumes that, similar to the comparison ratio effect, the size of the priming effect is related to the overlap between the prime and target values (Figure 1). (See additional effects the ANS is assumed to account for in Chapter 2.)

The ANS was outstanding not only because it could account for simple cognitive effects but also because it may be influential in higher-level mathematical capabilities. Like other psychophysical systems (i.e., systems obeying Weber's principle), the ANS has an acuity (or precision), the Weber fraction. There are **individual differences in the ANS acuity** (Halberda et al., 2008), and this acuity correlates with the mathematics grades in school and other more advanced mathematical abilities (Libertus et al., 2011; Schneider et al., 2017; Szkudlarek & Brannon, 2017). Moreover, training in ANS acuity may improve higher-level mathematical performance (Park & Brannon, 2013, 2014). Additionally, the impaired acuity may be the reason why some children live with developmental dyscalculia (a learning difficulty where children show specific impairment in mathematical tasks without having low intelligence) (Piazza et al., 2010). These latter examples characteristically demonstrate why the ANS is essential not only for basic science but for many applied areas, and its relevance can be fundamental in school and everyday life. (Find more examples in Chapter 2.)

However, the ANS model may be fundamentally flawed.

An alternative account, the discrete semantic system (DSS)

Issues with the ANS account

Like in the case of most models, there are empirical findings that are not entirely consistent with the ANS account. In general, such inconsistencies may exist not necessarily because the models are incorrect, but they may be rooted in invalid measurements, missing details in the models, or simply extreme data caused by random sampling (i.e., random noise in the measurement). However, in the case of the ANS, there may be more serious warning signs that the model is incorrect.

One fundamental issue in the case of the ANS is that the model cannot explain why **humans are able to discriminate numbers even with small ratios**. Weber's principle states that discrimination of two stimuli depends on the ratio of the two stimuli (Algom, 2021). Discrimination of the stimuli with a small ratio may be difficult, even close to random. In humans, the Weber fraction for symbolic numbers is measured to be around 0.1 (Krajcsi, Lengyel, & Kojouharova, 2018). With this Weber fraction, the discrimination of even the smallest-ratio single-digit number pair – that is, 8 vs. 9 – should be very difficult. Psychophysical models predict that the error rate for an 8 vs. 9 comparison with 0.1 Weber fraction should be around 20%. Even without appropriate empirical measurement, it seems to be unlikely that human adults cannot tell if 8 or 9 is larger in 20% of the trials, and empirical results confirm this anticipation (Krajcsi, Lengyel, & Kojouharova, 2018). This estimated Weber fraction predicts even larger error rates when generalized to multi-digit numbers (as ANS is assumed to back multi-digit comparison, e.g., Dehaene, 2007). This simple but characteristic and strong discrepancy is a fundamental sign that the ANS model may not necessarily be an appropriate account for the number comparison performance, although the psychophysics-based ratio effect is the most often cited phenomenon that is considered the defining feature of the ANS.

Another critical issue is that whenever a distance (or, equivalently, ratio or size) effect is observed in a numerical task, researchers assume that it is a sign of the ANS. However, **similar distance effects may be rooted in qualitatively different representations**. In a picture naming task, it was found that the reaction time depended on the meaning of the previous trial, and the strength of the effect depended on the semantic similarity of the meanings of actual and previous pictures

(Vigliocco et al., 2002). This priming distance effect is the exact parallel of the numerical priming distance effect (see above); still, considering the nature of language and psycholinguistic models, it is unlikely that the meaning is stored in simple psychophysical representations. This example can be considered as a warning sign that not all distance (or ratio or size) effects indicate an ANS-based processing.

The discrete semantic system model

Inspired by the latter linguistic priming distance effect, our lab started to explore the possibility of a representation similar to psycholinguistic or semantic networks. Can a network of nodes representing values and related concepts account for the phenomena that were attributed to the ANS?

We proposed a network of nodes that represent Arabic digits or other symbolic number notations, and related concepts, such as “smaller”, “larger”, “even”, “odd” (Figure 2) (see similar, but conceptually less far-reaching models, e.g., in Leth-Steensen & Marley, 2000; Verguts & Fias, 2004; Verguts & Van Opstal, 2014). We term this representation the **discrete semantic system (DSS)**. Various symbolic number processing effects that were formerly attributed to the ANS can be explained in this model as well.

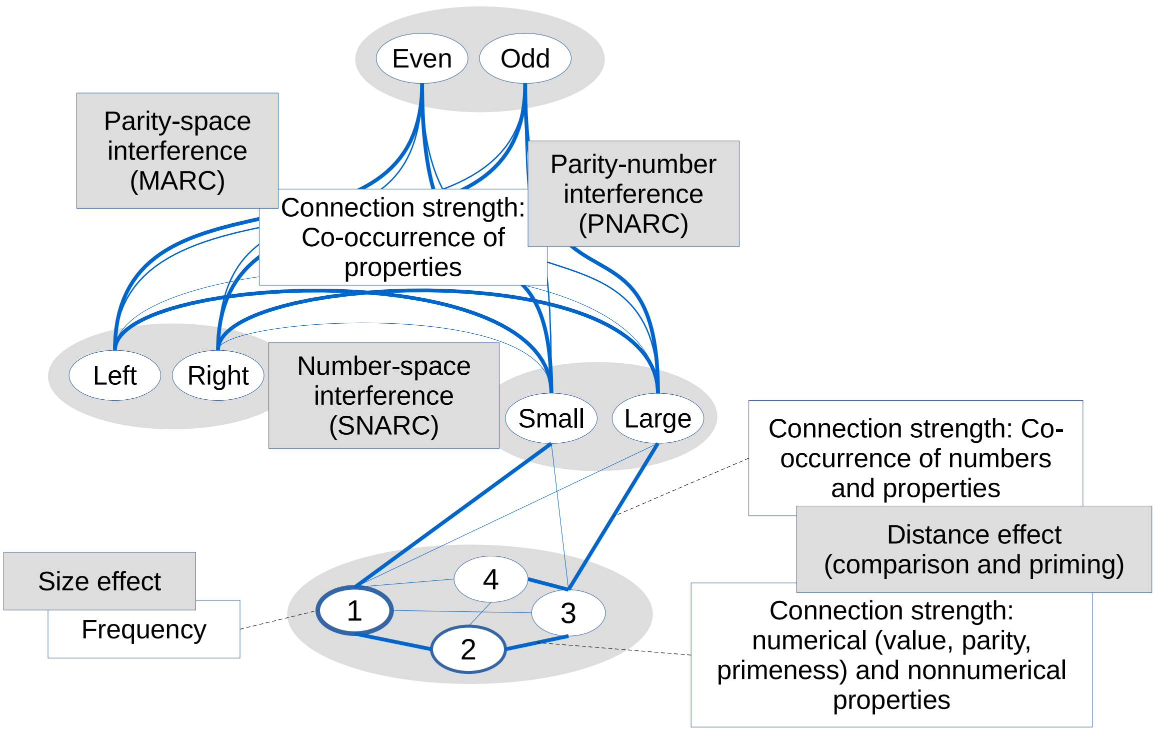


Figure 2 Hypothetical structure of the Discrete Semantic System. Nodes (white ellipses) of the network represent numbers and related concepts. Connections may have different weights, depicted as the width of the lines. Various properties of the network (white squares) may explain the symbolic numerical effects (gray squares).

In the DSS model, the ratio effect is not a single effect, but it is, in fact, comprised of two **independent effects, the size and distance effects**, that add up to form a seeming ratio effect. In the DSS model, **the size effect is a frequency effect**. It is well known in the literature that the frequency of symbolic numbers is related to their value in everyday life: Smaller numbers are more frequent, and larger numbers are rarer (Dehaene & Mehler, 1992). It is also well known that more frequent stimuli are easier to process. Therefore, smaller numbers can be processed faster. This latter rule is a possible version of the previously observed numerical comparison size effect (Figure 2). Regarding the **distance effect**, in the DSS model, the digit nodes may form associations with other nodes, such as “small”, or “large”. When processing numbers, small numbers will be associated more strongly with the “small” node, and large numbers will be associated with the “large” node. When two numbers are compared, numbers of a pair with a small numerical distance have more similar associations with “small” and “large” nodes than numbers of a pair with a large numerical distance. Dissimilar association (i.e., numbers with large numerical distance) may make the decision easier, creating the distance effect (see a more specific similar account in Verguts & Fias, 2004). Alternatively, to account for a distance effect, one may consider the strength of the connections between the number nodes, and spreading activation may influence number processing, where the strength of the connection and, consequently, the influence of the spreading activation causes a distance effect (Figure 2).

The DSS model can account for various **interferences** as well, such as the SNARC or other effects. Connections may be formed between property pairs (for similar models, see also, e.g., Hines, 1990; Leth-Steensen & Marley, 2000; Proctor & Cho, 2006), which connections may help the processing of congruent items and hamper the processing of incongruent items (Figure 2). (See more details about the DSS model in Chapters 2 and 3.)

It is important to highlight that the DSS model applies only to symbolic numbers (such as Arabic numbers, number words, Roman numbers) but not to nonsymbolic numbers (such as arrays of dots, series of events, series of sounds). According to our alternative model (similar to the classic model), nonsymbolic numbers are processed by the ANS. To include both symbolic and nonsymbolic number processing in the model and to contrast the classic and the alternative view, it is important to highlight that according to the classic model, both symbolic and nonsymbolic numbers are processed by the ANS, which we term the **pure ANS framework**. Contrarily, in our alternative model, while nonsymbolic numbers are processed by the ANS, symbolic numbers are processed by the DSS, which we term the **hybrid ANS-DSS framework** (Figure 3).

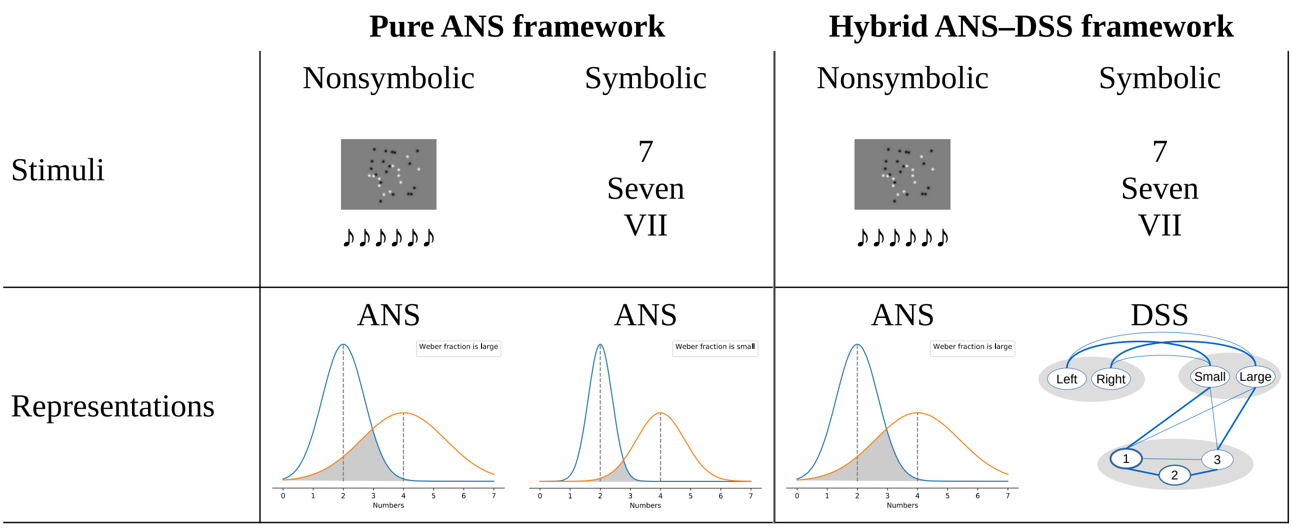


Figure 3 Hypothesized representations that handle simple numerical operations according to the pure ANS framework (left) and the hybrid ANS–DSS framework (right).

These considerations show that a network of discrete nodes, such as the DSS, may account for the same symbolic numerical phenomena that were attributed to the ANS. In other words, the DSS is a viable alternative model for symbolic number processing.

When someone wants to see whether the DSS model is not only viable but whether it is better than the ANS model, there are two limitations that must be handled. One limitation is that the two models have similar predictions for many phenomena. This is not surprising in the sense that the DSS model was created to account for the same phenomena that were attributed to the ANS. However, this also means that the two models cannot be contrasted based on the overlapping predictions, but differing predictions should be found. A second limitation to be handled is that while there may be phenomena for which the two models have different predictions, to our knowledge, with a few exceptions, the literature mostly reports phenomena that can be accounted for by both the ANS and the DSS models. To investigate whether the DSS is not only a viable but a better account than the ANS model, new tests had to be designed where the two models have characteristically differing predictions.

Contrasting the models

In a series of studies, new test cases had been designed to contrast the ANS and the DSS models for symbolic number processing, or more generally, to contrast the pure ANS framework and the hybrid ANS-DSS framework. Here, our main findings are summarized (see also Figure 4 as a visual summary of the findings along with the Chapters of the present dissertation; see the main findings summarized for all the relevant papers here in Table 2).

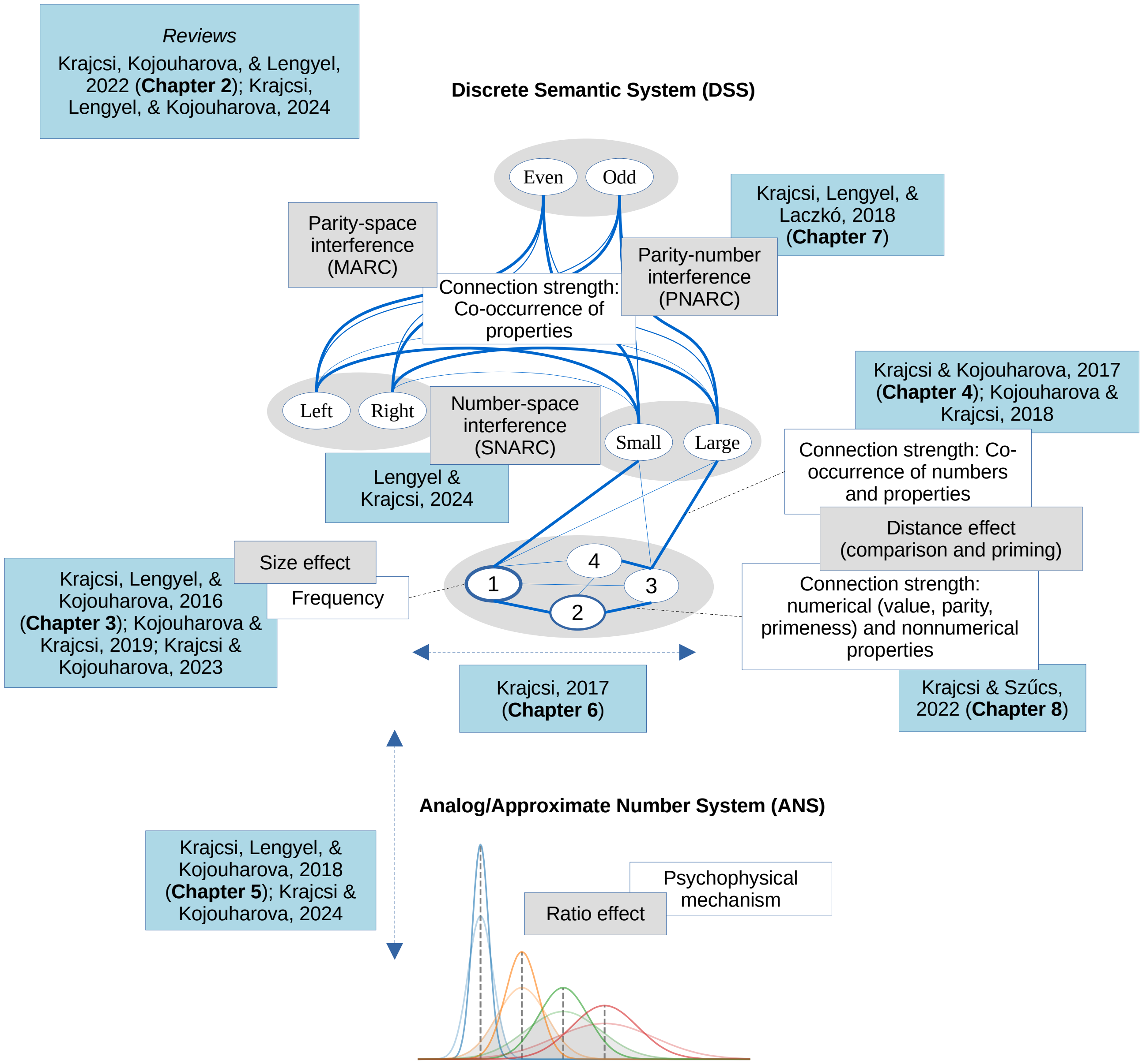


Figure 4 An outline of the DSS and ANS models with the relevant effects (gray squares), properties of the models to account for these effects (white squares), and the papers and chapters of the dissertation that test these effects (blue squares).

Contrasting the ANS and the DSS models in comparison task performance. The defining signature of an assumed ANS activation is the comparison ratio effect that obeys Weber's principle. For this effect, the ANS and the DSS have characteristically different accounts: in the DSS model, instead of a psychophysics-based ratio effect, the sum of a frequency-based size effect and an association-based distance effect is assumed. However, the comparison performance predictions of the two models are highly similar: depending on the specific versions or implementations of the models, their performance predictions of a single-digit comparison task may correlate stronger than 0.9 (Krajcsi et al., 2016). (See also Chapter 3 for more details.) While the predictions are similar, they also have small differences, and theoretically, it is possible to investigate which model predicts empirical data more precisely. Unfortunately, measured data are mostly too noisy to contrast the two models, and one cannot tell which model provide better prediction (Krajcsi et al., 2016). (See also Chapter 3 for more details.) Therefore, other tests should be designed to investigate which model describes symbolic comparison better.

Comparison size effect. Since the DSS model assumes that the size effect is a frequency effect, it also predicts that the size effect should change in line with the stimulus frequency. To avoid the lifelong experience with Arabic numbers, in a first test, new artificial digits were introduced, where their frequency was manipulated experimentally (Krajcsi et al., 2016). The participants learned new symbols that meant the values between 1 and 9; then, they compared number pairs in that new notation. In one group of participants, the comparison task included the digits that mimicked the frequencies of Arabic numbers in everyday life: Smaller numbers were more frequent, and the frequency was proportional to the power of the value (Dehaene & Mehler, 1992). In another group, the digit frequency was uniform; that is, each digit appeared to have the same frequency. While the size effect was observed in the first group, it was missing in the second group. This means that the size effect followed the frequency of the stimuli (Krajcsi et al., 2016). (See also Chapter 3 for more details.)

One may argue that this size effect is not the same thing as observed in Arabic numbers; the stimuli were new digits, after all. However, the same additional effects were observed with the new artificial notation as observed with Arabic numbers in the literature: The comparison distance effect was present, too; also, the decision time depended on the previous stimulus, and the effect was proportional with the numerical distance of the previous and the actual values, which is the priming distance effect (Krajcsi et al., 2016). If these effects are considered to be the sign of numerical processing in Arabic numbers, then these effects should be the sign of numerical processing in new symbols as well.

Importantly, the effect of the stimulus frequency on the size effect was observed in Arabic numbers, too. In another comparison study with Arabic numbers, not only everyday digit statistics were mimicked in one group, and uniform distribution was used in another group, but in a third group, reversed-everyday statistics were also used (i.e., large values were more frequent than small values). Although with Arabic numbers, small values were always compared faster, the difference between comparing small values vs. comparing large values (i.e., the size effect) was influenced by the stimulus frequency: The size effect was largest with everyday frequency, smaller with uniform frequency, and smallest with reversed-everyday frequency (Kojouharova & Krajcsi, 2019). In other words, the Arabic number comparison size effect seemed to show a combination of previous experience for stimulus frequency and the session’s experience for stimulus frequency. In sum, the frequency-based size effect can be observed with Arabic numbers, too, where we interpret our results as the combined experience of previous and actual stimulus frequency.

One may argue that the latter result of the combined size effects is not the combination of previous and present experience of the frequencies but the combination of the ANS-based ratio effect and the session’s frequency-based size effect. To investigate this possibility, in a follow-up study, participants compared new artificial digits, and the frequency of the stimuli was changed in the middle of the session (Krajcsi & Kojouharova, 2023). If the performance reflects the combination of the two frequencies, then frequency alone may be responsible for the combination of various statistics, and it is possible that in the previous Arabic number study, the stimulus frequency alone can cause the size effect. Indeed, the results showed that responses reflect the two frequencies seen in a single session (Krajcsi & Kojouharova, 2023). Furthermore, no additional sign of the ANS was observed. For example, if the size effect changes observed in our previous studies are, in fact, psychophysics-based ratio effect changes, then the distance effect should be changed, too (remember that, in the ANS model, the distance and size effects are simply two different ways to measure the ratio effect). However, no such changes in the distance effect were observed (Krajcsi & Kojouharova, 2023).

All of these results demonstrate that the symbolic comparison size effect depends on the stimulus frequency, and no other influence has been observed yet (Kojouharova & Krajcsi, 2019; Krajcsi et al., 2016; Krajcsi & Kojouharova, 2023). These results seriously challenge the ANS account and are in line with the DSS model.

Comparison distance effect. According to the DSS model, similar to the symbolic comparison size effect, the symbolic comparison distance effect may also depend on the statistics of the stimuli. However, for the distance effect, the statistics are not simply the frequency of the stimulus but the frequency at which a number and the “larger”–“smaller” categories are observed together. In a single-digit comparison, if the stimuli have uniform frequency distribution (i.e., each digit is shown with the same frequency), then 1 is never larger, 9 is always larger, and the other digits are larger proportionally with their values (see Table 1 with a modified series).

Numbers	1	2	3		7	8	9
Chance of being smaller in a comparison	100%	80%	60%		40%	20%	0%
Chance of being larger in a comparison	0%	20%	40%		60%	80%	100%

Table 1. The chance of being smaller or larger in a comparison task when the symbols are presented with equal probability.

In a series of tests, we utilized the fact that if some numbers are omitted from the 1-9 series, then the association statistics will be changed, while obviously, the values themselves stay the same (see Table 1). This way, the association statistics and the values can be dissociated, which can be applied for appropriate tests: The ANS model assumes that the performance depends on the values (more specifically, their ratios), while the DSS model assumes that the performance depends on the stimulus statistics. Investigating participants’ performance when comparing new artificial digits, the performance was predicted better by the association statistics than by the values of the digits (Krajcsi & Kojouharova, 2017). (See Chapter 4 for more details.) The same result was replicated with Arabic numbers as well (Kojouharova & Krajcsi, 2018). These results, again, cannot be accounted for by the ANS model but are in line with the DSS model.

In the presented tests so far, usually the statistical properties of the symbolic stimuli were manipulated, and it was found that comparison size and distance effects depended mainly on these statistics. However, it is possible that these statistical properties can influence not only symbolic but also nonsymbolic comparisons. To test this possibility, while participants compared arrays of dots numerically, the statistics of the stimuli were manipulated similarly to the manipulations applied in the previously described studies (Krajcsi & Kojouharova, 2024). It was found that the nonsymbolic comparison distance effect depends mainly on the values of the stimuli and not the statistics of them. This clearly shows again that symbolic and nonsymbolic comparisons work qualitatively differently. It was also found that the nonsymbolic comparison size effect was somewhat influenced by the frequency of the stimuli; the influence was qualitatively different in terms of the drift-diffusion model (see more details below) than in symbolic comparison. Again, the result confirms that symbolic and nonsymbolic comparisons are processed qualitatively differently, confirming the hybrid ANS-DSS framework.

Test of the psychophysical model. One of the first tests above investigated whether the overall ANS model prediction or the DSS model prediction accounts better for the symbolic number comparison. In an alternative test, the psychophysics-based ANS model prediction was used to contrast symbolic (Arabic) and nonsymbolic (arrays of dots) number comparisons. A psychophysics-based performance for both symbolic and nonsymbolic number comparison would be in line with the pure ANS framework (Figure 3). On the other hand, if psychophysics-based model can account only for the nonsymbolic comparison but fails to account for the symbolic comparison, that would support the hybrid ANS-DSS framework. Similar tests can be found in the literature (e.g., see the summary of Dehaene, 2007), but our test investigated more detailed parts of the predictions. Overall, it was found that while the psychophysical model described the nonsymbolic number comparison without observable biases, it missed several details of the symbolic comparison (Krajcsi, Lengyel, & Kojouharova, 2018). One key issue is that the model predicts a much larger error rate for small ratio comparisons (e.g., 8 vs 9) than what can be observed in empirical data. This is basically the empirical and more formal confirmation of our previous note that while the ANS (and Weber's principle) supposes that small ratios cannot be discriminated, humans, in general, can tell which number is large; for example, they can tell that 9 is larger than 8, or 90 is larger than 89. Another key issue is rooted in drift-diffusion model-based results (Ratcliff et al., 2004, 2016; Ratcliff & McKoon, 2008; Shinn et al., 2020; Voss et al., 2013). The drift-diffusion model of simple decisions assumes that a decision is based on the accumulation of relevant evidence, where the decision depends, for example, on the decision threshold (set by, for example, the accuracy-speed trade-off intention), starting point of the information accumulation (any bias towards a response before the trial), or the drift rate (quality of the accumulated information, that is, the difficulty of the task). Importantly, in a psychophysical task, when the stimuli are equivalent in the relevant dimension, the task is impossible to solve, and the related drift rate is zero. In line with this assumption, in nonsymbolic comparison, as the ratio of the two stimuli gets smaller (i.e., the task gets harder), the drift rate tends to zero. However, in symbolic comparison, as the ratio of the stimuli gets smaller, the drift rate tends to a non-zero value (Kojouharova & Krajcsi, 2018). (See also Chapter 5 for more details.) Again, this result is in line with the everyday observation and simple intuition that humans can successfully perform numerical comparisons for any number pairs. Overall, while the psychophysical model successfully describes the nonsymbolic number comparison, it fails to account for symbolic number comparison, which is again a critical challenge for the pure ANS framework and is in line with the hybrid ANS-DSS framework.

Independence of the comparison distance and size effects. Previously described results have already demonstrated that manipulating the stimulus frequency changes the size effect but not the distance effect (Kojouharova & Krajcsi, 2019; Krajcsi et al., 2016; Krajcsi & Kojouharova, 2023). Conversely, manipulating the association strength between the values and the "large" – "small" properties changes the distance effect but not the size effect (Kojouharova & Krajcsi, 2018; Krajcsi & Kojouharova, 2017). These results show that the distance and size effects can be changed independently. This is essential in the contrast of the ANS and DSS models since the ANS proposes that because the distance and size effects come from the same psychophysical ratio effect, they cannot dissociate. On the other hand, in the DSS model, the symbolic distance and size effects are two independent effects. In another study, the same independence was demonstrated with the correlational method (Krajcsi, 2017). When comparison distance and size effect slopes are measured in a nonsymbolic comparison (here, arrays of dots), after disattenuating the role of the imperfect reliability (Spearman, 1904, 1910), the two effects correlate around 1, which is exactly the prediction of a psychophysical model, since the distance and size effects are simply two ways to measure the single ratio effect; therefore, the distance and size effect slopes should correlate perfectly. Conversely, in the symbolic number comparison (here, Arabic numbers), the distance and size effect slopes show a much lower correlation, demonstrating that symbolic number comparison is not a psychophysical effect. (See Chapter 6 for more details.) This result pattern is again a strong counter-argument against the pure ANS framework and confirms the hybrid ANS-DSS framework.

Interference effects. The ANS model assumes that it has a spatial orientation that may interfere with other spatial components of a task, such as the response locations (Dehaene et al., 1993). A more general account assumes that continuous stimulus properties may be represented either similarly or in a single, more general representation, and these properties (such as numerical value and space) may interfere with each other (Cantlon et al., 2009; Walsh, 2003). Contrarily, the DSS model assumes that any property pair may interfere with each other, similar to other linguistic or conceptual interference models (Hines, 1990; Leth-Steensen & Marley, 2000; Proctor & Cho, 2006).

One important contrast of these opposing models is whether numerical values may interfere with discrete properties that are unlikely to be represented on a continuous representation. Such a discrete property may be the parity (whether a number is even or odd). In a relevant contrast, it was investigated whether the parity and the value of a number interfere (which is possible in the DSS model) or not (in line with the ANS model). In a parity decision task, it was found that parity interferes with the values of the numbers, supporting the DSS model (Krajcsi, Lengyel, & Laczkó, 2018) again (See also Chapter 7 for more details.) .

The DSS also assumes that the interference comes from the relevant labels stored in the system. It may explain the interference not only for symbolic numbers but also for nonsymbolic numbers. In many nonsymbolic number-space interference tasks, the participants make decisions where the response buttons denote specific verbal labels. Many previous studies have demonstrated that this numerical-spatial interference is observable even if the task is nonsymbolic, but the numerical information is relevant in the task (Cleland et al., 2020; Cleland & Bull, 2019). According to the DSS model, in those cases, the interference may come not from the nonsymbolic representation of the value, but the related labels, such as “small” – “large”, “even” – “odd”, or “left” – “right”. The DSS model predicts that these interferences exist only if those labels are activated. To test this assumption, in a color decision task, where numbers are not relevant, the stimuli were either symbolic numbers (Arabic numbers) or nonsymbolic numbers (arrays of dots) in either red or blue. It was found that numerical-spatial interference is observable only with symbolic numbers but not with nonsymbolic numbers (Lengyel & Krajcsi, 2024). The pure ANS framework assumes that evolutionarily old number perception was evolved for nonsymbolic stimuli, hence this interference should be present for nonsymbolic stimuli, too. This result again supports the DSS account but contradicts the ANS account.

Priming distance effect. Another potential test case is the relation between the priming distance effect (PDE) and the comparison distance effect (CDE; see the ANS section above). According to the ANS model, both effects depend on the overlapping noisy number representations, which in turn depend on the ratio of the stimulus values and the participants' Weber fraction. This means that the PDE and the CDE slopes should correlate (similar to the comparison distance and size effect slopes). Contrarily, in the DSS model, the two effects may be rooted in different mechanisms (Figure 2). One serious challenge of such a correlational contrast is that the reliability of the PDE may be relatively small. In the statistical and methodological literature, it is well known, but in practical research, it is often ignored that low reliability attenuates the measured correlation (Spearman, 1910). Consequently, when any variable's reliability is low in a correlation, the observed correlation cannot be high, even if the true correlation is high. The only way to ensure that correlation is measured correctly is to ensure that the reliability is satisfyingly high (Spearman, 1910). Often, the only available method to increase reliability is to increase the number of trials. In a multi-session study, participants performed comparison tasks for approximately 6 hours (Krajcsi & Szűcs, 2022). This long multi-session data collection ensured that the reliability was high enough that the true correlation between the PDE and CDE slopes can be estimated, and a potentially perfect correlation can be observed. However, it was found that the correlation of the two slopes is low, much lower than predicted by the ANS model (Krajcsi & Szűcs, 2022). (See also Chapter 8 for more details.) This result, again, confirmed the DSS model.

To conclude, all of our works (Figure 4, Table 2) revealed phenomena that cannot be accounted for by the ANS model; therefore, a revision of the current mainstream model is needed. We propose that these findings are in line with the DSS model of symbolic number processing and the more general hybrid ANS-DSS framework. Future works will reveal whether the DSS model can be an appropriate account of simple symbolic number processing or if another alternative model to the ANS should be found.

Reference	Main findings
(Krajcsi et al., 2016)	• The predictions of the ANS and DSS models for symbolic number comparison performance are so similar that usual contrasting methods do not work. • Symbolic number comparison size effect depends entirely on stimulus frequency in new symbols.
(Kojouharova & Krajcsi, 2019)	• Arabic number comparison size effect depends partly on stimulus frequency.
(Krajcsi & Kojouharova, 2023)	• Symbolic number comparison size effect with new symbols can combine various stimulus frequencies. No additional sign of the ANS in the symbolic number comparison size effect.

(Krajcsi & Kojouharova, 2017)	• Symbolic number comparison distance effect depends dominantly on stimulus statistics in new symbols.
(Kojouharova & Krajcsi, 2018)	• Arabic number comparison distance effect depends dominantly on stimulus statistics.
(Krajcsi & Kojouharova, 2024)	• Nonsymbolic number comparison is dominantly not influenced by the stimulus statistics. However, stimulus frequency modifies the size effect. This size effect is characteristically different from the symbolic comparison size effect.
(Krajcsi, Lengyel, & Kojouharova, 2018)	• Psychophysical models describe nonsymbolic number comparison without biases, while they describe symbolic number comparison with systematic inaccuracies.
(Krajcsi, 2017)	• Nonsymbolic number comparison distance and size effect slopes correlate perfectly, while symbolic distance and size effect slopes correlate weakly.
(Krajcsi, Lengyel, & Laczkó, 2018)	• Parity interferes with the number value in the parity decision (PNARC effect).
(Lengyel & Krajcsi, 2024)	• The number-response side interference (SNARC) effect can be observed with symbolic numerical stimuli but not with nonsymbolic stimuli.
(Krajcsi & Szűcs, 2022)	• The symbolic comparison distance effect and the priming distance effect slopes do not correlate.

Table 2 Summary of the main findings in the order of the discussion here.

Manuscripts that are directly related to the topic of the dissertation

In this section, all of our published or submitted papers that are directly relevant to the topic of the present dissertation are listed. First, the papers that are also the chapters of the dissertation are listed (see also Figure 4). They are sorted by the order of the chapters. Second, additional papers that contrast the ANS and DSS models are mentioned, sorted by publication date (these are also included in Figure 4). Finally, papers that critically investigate the role of the ANS in other numerical cognition phenomena are summarized, sorted by publication date (these are not mentioned in Figure 4).

Papers that are part of the dissertation

Chapter 2. Krajcsi, A., Lengyel, G., & Kojouharova, P. (2024). A new framework for elementary number processing: The hybrid ANS–DSS account [Submitted].

- A comprehensive review of many effects attributed to the ANS, alternative accounts provided by the DSS, and contrasting tests. It also discusses related models and the big picture of what representations may support number understanding.

Chapter 3. Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.01795

- The chapter discusses and contrasts the quantitative predictions of the ANS and DSS models for a comparison task. It finds that the models cannot be contrasted with fit on empirical data applying the usual measurement precision. In an experimental manipulation, the dominant role of the stimulus frequency on the comparison size effect is demonstrated.

Chapter 4. Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.02013

- The work demonstrates that the symbolic comparison distance effect reflects dominantly the conditional probability of the stimuli.

Chapter 5. Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analog number system: different symbolic and nonsymbolic numerical distance and size effects. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.00124

- The work investigates the psychophysical model in symbolic and nonsymbolic comparison tasks and finds that while the model describes nonsymbolic tasks appropriately, it is biased for symbolic tasks.

Chapter 6. Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. Psychonomic Bulletin & Review. 24(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

- The work demonstrates that while the nonsymbolic comparison distance and size effects correlate, the symbolic distance and size effects do not.

Chapter 7. Krajcsi, A., Lengyel, G., & Laczkó, Á. (2018). Interference between number magnitude and parity: Discrete representation in number processing. Experimental Psychology. 65(2), 71–83. https://doi.org/10.1027/1618-3169/a000394

- The work explores the number-parity interference.

Chapter 8. Krajcsi, A., & Szűcs, T. (2022). Symbolic number comparison and number priming do not rely on the same mechanism. Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-022-02108-x

- The study finds that the comparison distance effect and the priming distance effect do not rely on the same mechanism.

Additional papers that contrast the ANS and the DSS models

- Kojouharova, P., & Krajcsi, A. (2018). The Indo-Arabic distance effect originates in the response statistics of the task. Psychological Research. https://doi.org/10.1007/s00426-018-1052-1

    - The paper finds that stimulus association statistics influence the comparison distance effect not only with new symbols but also with Arabic numbers.

- Kojouharova, P., & Krajcsi, A. (2019). Two components of the Indo-Arabic numerical size effect. Acta Psychologica, 192, 163–171. https://doi.org/10.1016/j.actpsy.2018.11.009
    - The paper observes that stimulus frequency influences the comparison size effect not only with new symbols but also with Arabic numbers.
- Krajcsi, A., Kojouharova, P., & Lengyel, G. (2022). Processing Symbolic Numbers: The Example of Distance and Size Effects. In J. Gervain, G. Csibra, & K. Kovács (Eds.), A Life in Cognition: Studies in Cognitive Science in Honor of Csaba Pléh (pp. 379–394). Springer International Publishing. https://doi.org/10.1007/978-3-030-66175-5_27
    - The paper reviews evidence supporting the DSS model, mostly in the example of the comparison distance effect.
- Krajcsi, A., & Kojouharova, P. (2023). Stimulus frequency alone can account for the size effect in number comparison. Acta Psychologica, 232, 103817. https://doi.org/10.1016/j.actpsy.2022.103817
    - The work shows that various stimulus frequency patterns throughout a session can be summarized by the DSS.
- Krajcsi, A., & Kojouharova, P. (2024). Different sources of the numerical comparison size effect [Submitted].
    - The paper investigates whether stimulus statistics can influence the nonsymbolic comparison and finds that, in some cases, stimulus statistics does not influence the nonsymbolic comparison and, in some other cases, they influence the nonsymbolic comparison but qualitatively differently than it influences symbolic comparison.
- Lengyel, G., & Krajcsi, A. (2024). SNARC effect emerges only with symbolic numbers [Submitted].
    - The work demonstrates that the SNARC effect is related to symbolic numerical tasks, and it is not observable in nonsymbolic tasks where the symbolic components of the tasks are avoided.

Papers that investigate the role of the ANS in other numerical cognition phenomena

- Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. Experimental Psychology, 60(4), 227–234. https://doi.org/10.1027/1618-3169/a000191
    - The study shows that subitizing (fast and precise enumeration of small sets) depends on the spatial arrangement of the set; therefore, subitizing cannot be accounted for by the ANS (which is not sensitive to the spatial arrangement of the stimuli) but more probably by a pattern recognition mechanism.

- Krajcsi, A. (2020). Ratio effect slope can sometimes be an appropriate metric of the approximate number system sensitivity. Attention, Perception, & Psychophysics, 82(4), 2165–2176. https://doi.org/10.3758/s13414-019-01939-6
  - The simulation presents that the numerical distance effect slope as a proxy of the ANS Weber fraction can be used only in special circumstances, and generally, it is better to calculate Weber fraction with sigmoid fit.
- Krajcsi, A., Fedele, M., & Reynvoet, B. (2023). The approximate number system cannot be the leading factor in the acquisition of the first symbolic numbers. Cognitive Development, 65, 101285. https://doi.org/10.1016/j.cogdev.2022.101285
  - Our extended mathematical model shows that the ANS cannot be the main driving mechanism behind the initial understanding of the first number words in preschoolers.
- Krajcsi, A., Chesney, D., Cipora, K., Coolen, I., Gilmore, C., Inglis, M., Libertus, M., Nuerk, H.-C., Simms, V., & Reynvoet, B. (2024). Measuring the acuity of the approximate number system in young children. Developmental Review, 72, 101131. https://doi.org/10.1016/j.dr.2024.101131
  - The review summarizes unresolved methodological issues of measuring the ANS Weber fraction. It suggests that majority of the previous works measured the ANS Weber fraction invalidly and/or unreliably in children, which questions the validity of those studies' conclusions.

The structure and topics of the present dissertation

The present dissertation discusses the detailed theoretical and empirical considerations of whether the ANS or the DSS models can be appropriate to account for essential numerical cognition phenomena.

The present introduction provided a bird's eye view of the classic ANS model, our alternative DSS model, and some works providing contrast of the two models.

The second chapter provides a longer and more extensive review of the ANS and DSS models, also discussing other alternative models that are or may be strongly related to the DSS model, and discussing further representations that may be essential in understanding and processing numbers.

The next four chapters (Chapters 3-6) describe empirical works that contrast the ANS and DSS accounts or the pure ANS framework and the hybrid ANS-DSS framework by investigating the comparison distance, size, and ratio effects.

The final two empirical chapters (Chapters 7-8) describe additional contrasts of the models by investigating the nature of number-related interferences and the relation of priming and comparison distance effects.

The final chapter (Chapter 9) discusses follow-up questions about what the potential role of the nonsymbolic ANS and the DSS representations may have in number understanding, and what other representations may be relevant in building up an understanding of numbers.

All chapters except the present and the last one are manuscripts that either have been published (Chapters 3-8) or have been submitted (Chapter 2). These chapters represent the preprint or postprint versions of those manuscripts and not the final published versions. All chapters have their separate figure and table numbering and reference list, but the whole dissertation has common page numbering.

Acknowledgments

I thank Kimberly Brosche, Petia Kojouharova, and Gábor Lengyel for their comments on an earlier version of the present chapter.

References

Algom, D. (2021). The Weber–Fechner law: A misnomer that persists but that should go away. *Psychological Review, 128*(4), 757–765. https://doi.org/10.1037/rev0000278

Brannon, E. M., & Merritt, D. J. (2011). Chapter 14—Evolutionary Foundations of the Approximate Number System. In S. Dehaene & E. M. Brannon (Eds.), *Space, Time and Number in the Brain* (pp. 207–224). Academic Press. https://doi.org/10.1016/B978-0-12-385948-8.00014-1

Butterworth, B. (1999). *The mathematical brain*. Macmillan.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91. https://doi.org/10.1016/j.tics.2008.11.007

Cleland, A. A., & Bull, R. (2019). Automaticity of access to numerical magnitude and its spatial associations: The role of task and number representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(2), 333–348. https://doi.org/10.1037/xlm0000590

Cleland, A. A., Corsico, K., White, K., & Bull, R. (2020). Non-symbolic numerosities do not automatically activate spatial–numerical associations: Evidence from the SNARC effect. *Quarterly Journal of Experimental Psychology, 73*(2), 295–308. https://doi.org/10.1177/1747021819875021

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Vol. XXII* (pp. 527–574). Harvard University Press.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and mental number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Garcia-Retamero, R., Sobkow, A., Petrova, D., Garrido, D., & Traczyk, J. (2019). Numeracy and Risk Literacy: What Have We Learned so Far? *The Spanish Journal of Psychology*, *22*. https://doi.org/10.1017/sjp.2019.16

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Hines, T. M. (1990). An odd effect: Lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*(1), 40–46. https://doi.org/10.3758/BF03202644

Kojouharova, P., & Krajcsi, A. (2018). The Indo-Arabic distance effect originates in the response statistics of the task. *Psychological Research*. https://doi.org/10.1007/s00426-018-1052-1

Kojouharova, P., & Krajcsi, A. (2019). Two components of the Indo-Arabic numerical size effect. *Acta Psychologica*, *192*, 163–171. https://doi.org/10.1016/j.actpsy.2018.11.009

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02013

Krajcsi, A., & Kojouharova, P. (2023). Stimulus frequency alone can account for the size effect in number comparison. *Acta Psychologica*, *232*, 103817. https://doi.org/10.1016/j.actpsy.2022.103817

Krajcsi, A., & Kojouharova, P. (2024). *Different sources of the numerical comparison size effect* [Submitted].

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analogue number system: Different symbolic and nonsymbolic numerical distance and size effects. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00124

Krajcsi, A., Lengyel, G., & Laczkó, Á. (2018). Interference between number magnitude and parity: Discrete representation in number processing. *Experimental Psychology*, *65*(2), 71–83. https://doi.org/10.1027/1618-3169/a000394

Krajcsi, A., & Szűcs, T. (2022). Symbolic number comparison and number priming do not rely on the same mechanism. *Psychonomic Bulletin & Review*, *29*(5), 1969–1977. https://doi.org/10.3758/s13423-022-02108-x

Lengyel, G., & Krajcsi, A. (2024). *SNARC effect emerges only with symbolic numbers* [Submitted].

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*(1), 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292–1300. https://doi.org/10.1111/j.1467-7687.2011.01080.x

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Odic, D., & Starr, A. (2018). An Introduction to the Approximate Number System. *Child Development Perspectives*, *12*(4), 223–229. https://doi.org/10.1111/cdep.12288

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition*, *133*(1), 188–200. https://doi.org/10.1016/j.cognition.2014.06.011

Parrish, A. E., & Beran, M. J. (2022). Approximate Number System (ANS). In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 381–386). Springer International Publishing. https://doi.org/10.1007/978-3-319-55065-7_1068

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A Diffusion Model Account of the Lexical Decision Task. *Psychological Review, 111*(1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, *20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12372

Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *eLife*, *9*, e56938. https://doi.org/10.7554/eLife.56938

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. https://doi.org/10.2307/1412159

Spearman, C. (1910). Correlation Calculated from Faulty Data. *British Journal of Psychology, 1904-1920*, *3*(3), 271–295. https://doi.org/10.1111/j.2044-8295.1910.tb00206.x

Szkudlarek, E., & Brannon, E. M. (2017). Does the Approximate Number System Serve as a Foundation for Symbolic Mathematics? *Language Learning and Development*, *13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion Models in Experimental Psychology. *Experimental Psychology*, *60*(6), 385–402. https://doi.org/10.1027/1618-3169/a000218

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Chapter 2: The hybrid ANS–DSS account

The present chapter has been submitted:

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2024). A new framework for elementary number processing: The hybrid ANS–DSS account [Submitted].

A new framework for elementary number processing: The hybrid ANS–DSS account

Attila Krajcsi[1], Petia Kojouharova[2], Gábor Lengyel[3]

[1]Cognitive Psychology Department, Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

[2]Research Centre for Natural Sciences, Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary

[3]Department of Cognitive Science, Central European University, Budapest, Hungary

Abstract. *In the last few decades, an evolutionarily old, simple representation, the Approximate Number System (ANS), was assumed to account for a series of simple number processing phenomena and also for providing a primary basis for number understanding. In this paper, we provide an alternative framework, in which—although nonsymbolic numerosities are still processed with the ANS—symbolic numbers are processed by a network of nodes, which we term the Discrete Semantic System (DSS); this is similar to linguistic or conceptual models and partly similar to some of the former alternative numerical models. We argue that this hybrid ANS–DSS framework not only provides an appropriate description of already known phenomena, but it also accounts for several additional phenomena that the classic ANS account is unable to handle. Finally, in contrast with the classic model, we argue that a more diverse set of modules could be the basis for number understanding and mathematical thinking.*

Keywords: Approximate Number System; Discrete Semantic System; number understanding;

Introduction

In 1967, a simple representation was proposed for number processing, which was later termed the Approximate Number System (ANS). The ANS is an evolutionarily old representation that works according to Weber's law in a similar way to how representations process perceptual physical properties. In recent decades, it has been the widely held belief that this system is the very base of number understanding. The present review proposes an alternative framework, in which nonsymbolic (e.g., arrays of dots or series of sound events) and symbolic values (e.g., Indo–Arabic numbers, number words, and Roman numbers) are processed differently: While nonsymbolic numerosities are still processed by the evolutionarily old and simple ANS representation, symbolic numbers are handled by a network of nodes similar to the mental lexicon or conceptual networks. This framework also supposes that the ANS has a more limited role in forming numerical abilities.

This review starts with a brief summary of the ANS account. Based on the fact that the ANS model has been supported by a considerable amount of evidence and that, in the last few decades, hundreds of works have presented their results in the ANS framework, it can be said that this classic account is deeply rooted in scholarly thinking about numerical cognition. When presenting an alternative account to such a convincing and persuasive model, it is important to first recall and re-examine why the model in question has been so attractive, what exactly the main pieces of evidence are, and what the main arguments that support this comprehensive account are. In light of that background, the following section explains the main features of the alternative account, presenting how an entirely different architecture with different mechanisms could explain the very same symbolic phenomena that have been previously explained by the ANS and also how the ANS may be limited to nonsymbolic numerosity processing. The goal of that section is to show that our alternative proposal is viable and may explain key phenomena that have been believed to support the ANS account. However, because the two accounts make very similar predictions, the second section does not have the scope to assess which framework produces a more appropriate description of the available data. To assess the competing frameworks, the following section contrasts the two models by presenting several phenomena for which the two models make different predictions. The anticipated result of those contrasts is that all of the phenomena described in this paper are in line with our alternative proposal and simultaneously introduce crucial challenges for the ANS account. In the final section, we examine where the meaning of numerical information and mathematical concepts may come from and what role the discussed representations could have in broader mathematical cognition.

The Approximate Number System account

While the ANS model is well-known among scholars in the field of numerical cognition, it is important to summarize not only the model and its main variations but also the main findings and arguments that support the model. In this way, this summary makes it explicit why the ANS model is firmly believed to be comprehensive and appropriate and also lists phenomena that an alternative model needs to account for. In this summary, we also recount some phenomena that were initially believed to support the ANS model, but later findings produced counter evidence or at least conclusions that are open to interpretation and debate. Considering these phenomena is essential to understanding why the ANS is believed to be a model that accounts for an extremely wide range of phenomena and also making it explicit what the ANS model can actually account for.

The ANS and the ratio effect. In the 1960s, a very simple experiment profoundly changed how cognitive scientists perceived mathematical thinking. In the seminal work of Moyer and Landauer (1967), participants compared two single-digit Indo–Arabic numbers, and it was found that their behavioral performance depended on the ratio of the two values: Number pairs with larger ratios are easier to process. They argued that the ratio-dependent performance was a sign of a psychophysical representation working according to Weber’s law as in perceptual comparisons. This representation was later termed, among other names, the Approximate Number System. The most important general consequence of this finding was that mathematical understanding could be framed in an entirely new way: While it had formerly been believed that mathematics is an abstract, complex, human-specific and culture-dependent quality, the existence of the ratio-based comparison performance and the assumed psychophysical representation demonstrated that even symbolic mathematics is partly rooted in a simple, evolutionarily old mechanism.

The ratio effect is considered to be a key signature of the ANS (Table 3), and this effect has been found in various conditions. For example, the ratio effect was observed in many other tasks, such as matching tasks, where participants have to decide whether two symbolic numbers are the same or different (Dehaene & Akhavein, 1995), also priming tasks, where the priming effect of a previous stimuli depends on the ratio of the priming and target stimuli (Koechlin et al., 1999; Reynvoet & Brysbaert, 1999), as well as approximation tasks, where only the approximate result (and not the exact result) of a calculation task needs be found (Dehaene et al., 1999). The ratio effect is observable in numerical tasks, independent of the modality or the notation of the stimuli, e.g., regardless of whether it is visual or auditory, an object or an event, and Indo–Arabic digits or arrays of dots (Dehaene et al., 1998; Eger et al., 2003; Wynn, 1996). The ratio effect can be observed in not only human adults but also children and infants (Feigenson et al., 2004; Izard et al., 2009) and also nonhuman animals (Hauser & Spelke, 2004). The ratio effect is present in not only behavioral measurements but also neuroscientific studies (Dehaene, 1996; Dehaene et al., 1999; Eger et al., 2003; Nieder, 2005; Piazza & Dehaene, 2004; Piazza & Izard, 2009). In technical terms, the ratio effect is often measured as the distance effect (number pairs with larger numerical distance are easier to process than number pairs with smaller distance) or occasionally as the size effect (improving performance by decreasing the numerical magnitude of the stimuli). Even so, the distance and size effects are considered to be technically simpler ways of measuring the ratio effect. The widely observed ratio effect is one of the most important pieces of evidence for why a single representation, the ANS may be behind a wide range of phenomena for various species, ages, and tasks.

Representational implementation. On the representational level, the most frequently cited description of the ANS supposes a noisy representation of values with a Gaussian distribution, where the dispersion of the distribution is proportional to the values, i.e., larger numbers are noisier (Figure 5). This simple description is capable of explaining the comparison ratio effect: The difficulty of the task is proportional to the overlap of the two noisy representations of the two to-be-compared numbers (e.g., see the gray areas in Figure 5). (Find a detailed mathematical description how this mechanism is able to generate the ratio effect in Dehaene (2007).) In an alternative formulation, the scale of the representation is logarithmic, and the dispersion of the distribution is a constant value. While, in some cases, this alternative formulation gives slightly different quantitative predictions compared to the previously described linear version, the difference is usually so small in proportion to the precision of the data that, most of the time, the two models make practically the same predictions (Dehaene, 2007).

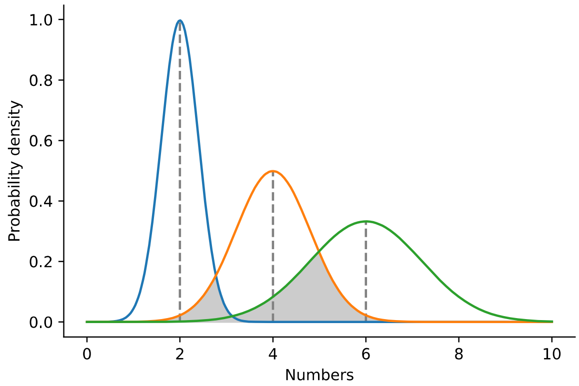


Figure 5. A possible formulation of the hypothesized number representation in the ANS account. Numbers are stored in a linear scale, and number representations are noisy Gaussian distributions, where the standard deviation of the distribution is proportional to the value that it represents. Number processing performance depends on the overlapping of relevant number representations.

The sensitivity of the ANS. Another essential feature of the ANS that bolsters its explanatory power for further phenomena is its sensitivity. The dispersion of the Gaussian representation depends on the sensitivity of the system expressed as the Weber fraction (Figure 6), in addition to the to-be-represented value. In mathematical terms, the standard deviation of the distribution is the product of the to-be-represented value and the Weber fraction. Consequently, an ANS with a smaller Weber fraction has more narrow distributions, which, therefore, makes the ANS more sensitive, e.g., resulting in a smaller overlap of two to-be-compared number representations, which in turn leads to a better comparison performance. The sensitivity of the ANS (the dispersion of the Gaussian representation) may account for a series of phenomena (Table 3, Figure 7). It has been demonstrated that there are individual differences in how sensitive the ANS is (Halberda et al., 2012), and this sensitivity improves as people age (Halberda et al., 2012; Halberda & Feigenson, 2008; Piazza et al., 2010): In infancy, one's ANS is capable of discriminating only several-fold differences (Cordes & Brannon, 2008; Feigenson et al., 2004), while, in adulthood, approximately a 10% difference can be discriminated (Halberda et al., 2012; Halberda & Feigenson, 2008). These individual differences may be essential in higher-level mathematical performance and abilities: The ANS sensitivity has been found to correlate with various math performance indexes (Halberda et al., 2008; Schneider et al., 2017). Moreover, improving the sensitivity of one's ANS via training tasks may lead to improved general math ability (e.g., Park & Brannon, 2013). Furthermore, it has been proposed that extremely low sensitivity may be the source of developmental dyscalculia (DD), which is a condition where one's numerical abilities are selectively impaired (Molko et al., 2003; Piazza et al., 2010; Price et al., 2007).

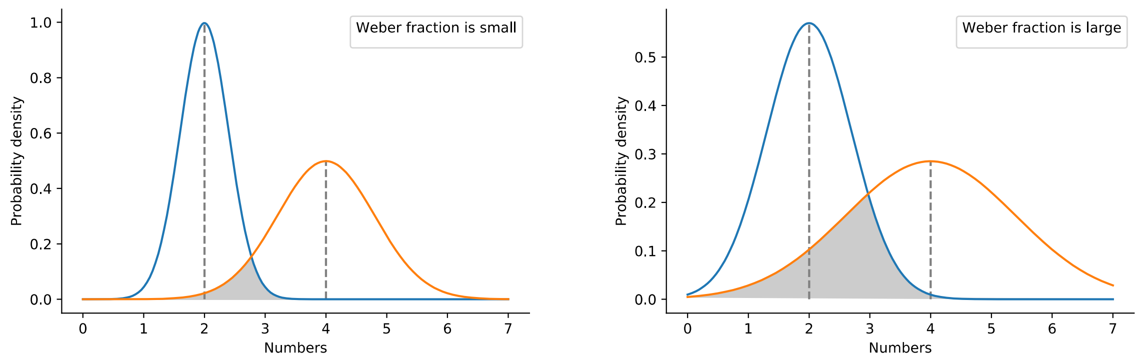


Figure 6. Illustration of the effect of the sensitivity of the ANS expressed with the Weber fraction. The smaller the Weber fraction is, the more sensitive the representation, which results in a smaller overlap between relevant number representations.

Differing Weber fractions are also important when the different precision of symbolic and nonsymbolic operations are accounted for. In a comparison task, symbolic number processing is more precise than nonsymbolic numerosity processing. (Note that the possible differences between Weber fractions in approximate symbolic and nonsymbolic calculations are not discussed in the literature, because approximate symbolic calculations do not require precise representation. This may be an inconsistency in the model which we discuss in the following subsection.) To explain this difference, the model proposes that there may be an ANS for nonsymbolic numerosity processing with a higher Weber fraction, and a separate ANS for symbolic number processing with a lower Weber fraction (Dehaene, 1997, 2007). Note that this means that there are at least two separate ANSs with different Weber fractions, but these separate systems otherwise work similarly.

The limited precision of the ANS with specific Weber fractions can explain additional phenomena (Table 3, Figure 7). When one knows the Weber fraction and the specific representational overlap (Figure 5) it is possible to specify which neighboring numbers can be reliably discriminated. Larger values are less likely to be discriminated because of their larger noise. The limit of reliably discriminable number pairs may account for additional phenomena, such as subitizing or the development of subset-knowers. Subitizing is a relatively fast and relatively precise set enumeration process working up to 4 as opposed to the fast but imprecise estimation process and also to the precise but slow counting process (Kaufman et al., 1949). It was proposed that sets can be subitized up to 4, because only numbers up to 4 can reliably be discriminated by adults' ANS, and this ANS limitation is in line with the limitation of subitizing (Dehaene & Cohen, 1994; Gallistel & Gelman, 2000). A related proposal is that subitizing may be impaired in dyscalculia (Desoete et al., 2009), which is another sign of an ANS deficit in mathematical disabilities.

The limitation of ANS-based discrimination may also explain the development of subset-knowers. In the development of symbolic numbers, preschool children first learn the meaning of 1 (measured with the give-a-number task, where children are asked to give a specific amount of objects from a larger set) without knowing other larger numbers (Wynn, 1990, 1992). Typically, a few months later, they learn the meaning of 2, then after some time the meaning of 3, and then 4. Interestingly, while children who already know a counting list up to a relatively high value are not able to apply all of those number words in the give-a-number task. These children are termed subset-knowers, which highlights the fact that they can only use a portion of their counting list properly. After being able to reliably give 4 in the give-a-number task, preschoolers are able to give any other numbers available in their counting list, and it is supposed that they understand the main principle how counting is used to specify the cardinality of a set (Wynn, 1990, 1992). It was proposed that the slow learning process of the subset-knowers is rooted in the slow development of the ANS's Weber fraction, and the cardinality principle is understood after the point when the ANS cannot discriminate larger values (Piazza, 2010).

Interference effects. The ANS is also applied to account for numerical interference effects (Table 3, Figure 7). For example, the spatial–numerical association of response codes (SNARC) effect in which the location of the response interferes with the value of the stimulus (e.g., in a parity decision task, participants responded faster on the left side for small numbers compared to large numbers and responded faster on the right side for large numbers compared to small numbers) is explained by the ANS (Dehaene et al., 1993). Similarly, the ANS is used to account for the size congruency effect in which numerical values interfere with the physical size of symbols (Henik & Tzelgov, 1982) and for the interference of number magnitude and duration (Oliveri et al., 2008). For the SNARC effect, it was proposed that the ANS may have a spatial property that interferes with the response location of the paradigm (Dehaene et al., 1993). In a more general account, it was supposed that the continuous numerical representation may interfere with similar continuous representations and potentially lead to various interference effects (Cantlon et al., 2009; Cohen Kadosh et al., 2008; Henik et al., 2012; Walsh, 2003).

Neural localization. Based on various studies in neuroscience, a converging picture was drawn regarding the brain areas that are involved in the simple numerical processing: Many fMRI and neuropsychological studies have found that the bilateral intraparietal sulcus (IPS) is a necessary area of the brain for solving ANS-related tasks (Dehaene et al., 1999; Eger et al., 2003; Nieder, 2005; Piazza & Dehaene, 2004; Piazza & Izard, 2009).

ANS property	**Related phenomena and selected references**
Ratio effect in operations	In various tasks: Comparison task, approximation task, priming effect, matching task (Dehaene et al., 1998, 1999; Dehaene & Akhavein, 1995, 1995; Eger et al., 2003; Koechlin et al., 1999; Moyer & Landauer, 1967) In various notations and modalities (Dehaene et al., 1998; Eger et al., 2003) In various ages and species (Halberda et al., 2008; Hauser & Spelke, 2004; Izard et al., 2009)

ANS property	**Related phenomena and selected references**
Sensitivity (Weber fraction)	Development with age (Feigenson et al., 2004; Piazza, 2010) Developmental dyscalculia (DD) (Piazza et al., 2010) Predictor of general math ability (Halberda et al., 2008; Schneider et al., 2017) Developing low level ANS sensitivity to improve general math ability (Park & Brannon, 2013) Symbolic–nonsymbolic difference (Dehaene, 1997)
Reliable discrimination of values	Subitizing (Dehaene & Cohen, 1994) Development of subset-knowers (Piazza, 2010)
Spatial/continuous property	SNARC effect, size congruency effect and other interference effects (Dehaene et al., 1993)
Neural localization	Intraparietal sulcus activation (Nieder, 2005; Piazza & Dehaene, 2004)

Table 3. Foundational properties of the ANS (left) with the appropriate phenomena (right) supporting those properties.

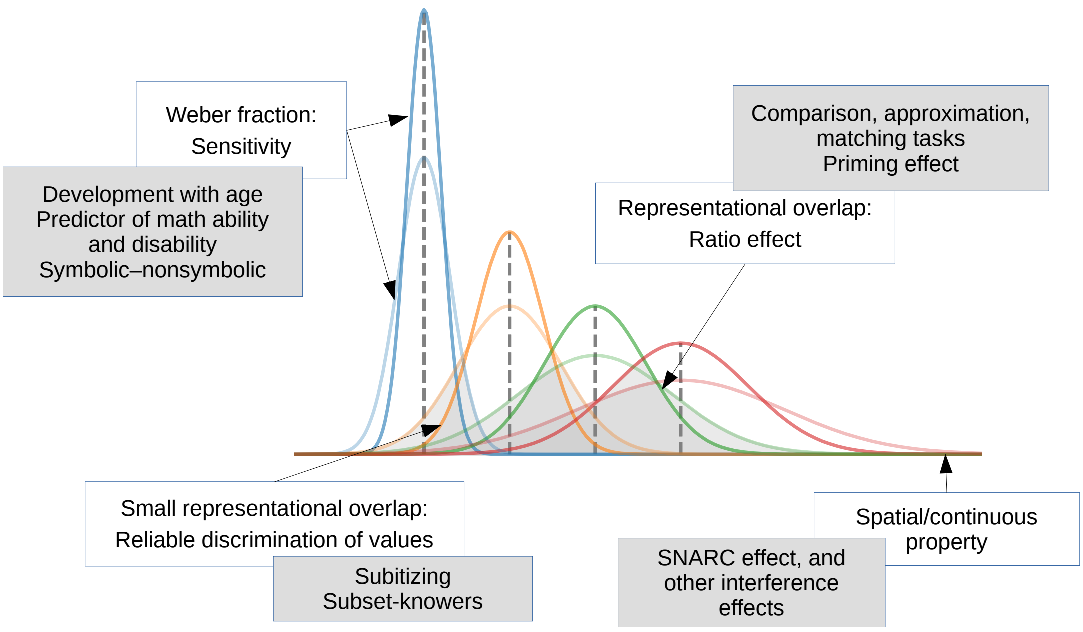


Figure 7. Illustration of the key features of the ANS (white boxes) with the related phenomena (gray boxes) on the representational model of the ANS.

Even this brief summary may show why the ANS model has become popular: With only a few assumptions of a noisy representation lead by the Weber fraction it can account for a series of phenomena, and, by adding a few other details (such as the supposed spatial property or the localization of the system in the brain), the model may explain additional phenomena. Table 3 and Figure 7 summarize the main properties of the ANS representation, and the main phenomena those properties may account for.

In the context of presenting an alternative account, it is important to highlight that, although it may seem that there is a near endless list of phenomena supporting the ANS account, in fact, the range of such phenomena is quite limited. For example, in his seminal paper, Dehaene (1992) convincingly argues that the ANS may be the root of several numerical phenomena; he lists only four main related phenomena: subitizing, estimation (quantity of sets), comparison and approximate calculation. Similarly, in a comprehensive and technical summary of the model, Dehaene (2007) lists the phenomena the ANS model accounts for: symbolic and nonsymbolic number discrimination, comparison and identification, simple calculations, neural coding, and development.

Debates about phenomena supporting the ANS account

While the ANS model is extremely efficient at accounting for a series of phenomena (Table 3, Figure 7), the above listing was intentionally far-fetched, including a few phenomena that, based on later findings, are no longer attributed to the ANS, and also a few phenomena over which there is debate whether the ANS can account for them. When forming alternative explanations, it is essential to consider these phenomena, either because they are not relevant in the original account (and, consequently, it is not necessary to include them in an alternative model) or it remains inconclusive as to which model those phenomena support.

Subitizing. While the ANS was hypothesized to be the root of subitizing, the fast and precise enumeration of small sets (Dehaene & Cohen, 1994), later findings suggest another generator for the effect. It was found that, because the subitizing effect cannot be observed for multiples of ten, it may not be the ratio of the values and consequently the ANS that produces the effect (Revkin et al., 2008). In addition, it was found that subitizing is sensitive to the spatial pattern formed by the items; hence, subitizing could be rooted in a pattern detection mechanism (Krajcsi et al., 2013; Mandler & Shebo, 1982). These findings highlight that subitizing is not supported by the ANS and is more probably supported by a pattern recognition mechanism.

Symbolic matching task. Distance effect was also observed in a matching task (i.e., participants have to decide whether two symbols are the same or different) of single digit numbers (Dehaene & Akhavein, 1995). Initially, it was interpreted that the effect reflects automatic ANS activation even if the semantic processing of the values is not necessary and visual matching could be sufficient to solve the task. However, later works have shown that the visual similarity and the numerical distance of the digit pairs correlate in some notations, such as Indo–Arabic digits, and that the seeming numerical distance effect is, in fact, the artifact of the visual similarity (Cohen, 2009; Defever et al., 2012; García-Orza et al., 2012; Wong & Szűcs, 2013).

Subset-knower children. It was also proposed that subset-knowers' symbolic number understanding is led by the ANS, and they learn their first symbolic numbers slowly because of the specific sensitivity improvement rate of the ANS (Piazza, 2010). However, the presented data of this proposal supposes that one-year-old infants may differentiate 1 vs. 2 and 2 vs. 3 (see Figure 1a(ii) in Piazza, 2010). This means that even the youngest children who are able to provide relevant responses to the give-a-number task should be at least two-knowers; however, one-knowers are usually reported in the literature (Wynn, 1990, 1992). In addition, in a longitudinal study, Shusterman and her colleagues (2016) demonstrated that the ANS sensitivity did not improve when children were subset-knowers and only did so when they became cardinality-principle-knowers. This result suggests that it is most probably number knowledge that modifies the measured ANS sensitivity and not the other way around, and that the relationship between number knowledge and the ANS sensitivity is stronger for cardinality-principle-knowers than for subset-knowers. A recent analysis also demonstrated that, even if some parameters of the development are handled more flexibly than in the original proposal, the ANS sensitivity improvement would predict a much slower development of subset-knowers than what has been observed; therefore, the ANS cannot be the main cause of subset-knowers' development (Krajcsi et al., in preparation).

Multi-digit symbolic comparison. The ANS has also been believed to play an essential role in multi-digit symbolic number comparison tasks, because the distance effect can be seen in not only single-digit comparison but also multi-digit comparison (Dehaene, 2007; Dehaene et al., 1990). However, the same effect can also be accounted for by a serial decomposition model in which multi-digit numbers are compared power by power (Hinrichs et al., 1982; Poltrock & Schwartz, 1984). In the latter explanation, the largest powers are first compared (e.g., for three-digit numbers, it is the hundreds), and, if there is a difference, a decision is made; if there is no difference between the largest powers, the second largest power (e.g., the tens) is considered, and so on. In this account, the larger the difference between the numbers is (in terms of the largest power where the difference is), the faster a decision is made, which is a distance effect. In addition, another single-digit distance effect is added to the power-based distance effect, which results in a more gradual effect. Huber et al. (2016) presents a more comprehensive description of the related effects and a computational model suggesting that the multi-power comparison may indeed rely on a decompositional process.

Nonnumerical features in nonsymbolic stimuli. When presenting nonsymbolic stimuli (e.g., arrays of dots), the numerical property of the stimuli may correlate with nonnumerical perceptual properties (e.g., the number of the dots correlates with the luminance of the stimulus). Because of this correlation, there is persistent concern as to whether perceptual properties are appropriately controlled for in empirical studies and whether the measured ANS sensitivity validly measures the ANS instead of measuring perceptual features. While most of works aim to control for perceptual features, recent consensus in the literature concludes that, in most of the common paradigms, perceptual features influence the measured numerical task performance (Gebuis & Reynvoet, 2012a, 2012b), and perceptual features are impossible to remove (DeWind et al., 2015). These issues with the nonsymbolic stimuli lead some authors to conclude that a number-specific representation does not exist (Leibovich et al., 2017). However, there are some special paradigms and analysis methods that could overcome this issue. For example, in infant studies, habituation paradigms allow the researchers to control for orthogonal features across different stimuli (Xu & Spelke, 2000); after simultaneously measuring the perceptual and numerical properties, analysis can separate the numerical and nonnumerical components (DeWind et al., 2015). (See more examples in the commentaries on Leibovich et al., 2017.) These latter works demonstrate that numerical information on nonsymbolic stimuli is indeed processed. While careful design of the paradigms and innovative analyses may provide results that convincingly demonstrate that the numerical information of the nonsymbolic stimuli is processed, it is important to highlight that many other methods in the literature cannot validly measure numerical processing. For example, in most nonsymbolic comparison tasks used in studies the effect of correlated perceptual properties cannot be avoided (DeWind et al., 2015; Gebuis & Reynvoet, 2012a, 2012b). In addition, when both the stimuli's numerical and perceptual properties vary in a paradigm, participants have to focus on the numerical information instead of the parallel nonnumerical information; consequently, these paradigms may measure inhibitory processes as well (Leibovich & Ansari, 2016). Critical reviews of animal studies suggest that previous works have not provided convincing evidence that animals process numerical information, rather than only correlating perceptual information (Gebuis et al., 2016). Therefore, many former methods that intended to measure numerical processing, in fact, simultaneously measure numerical processing, perceptual processing, and inhibitory processes. This ambiguity of the measured generators may be a critical issue in many works.

The cause of dyscalculia. DD has been proposed as related to the ANS in several ways. First, impaired ANS may cause DD (Molko et al., 2003; Piazza et al., 2010; Price et al., 2007). Second, impaired connections between the ANS and other relevant representations may be another cause of DD (Wilson et al., 2006). Third, impaired subitizing, which is supposed to be an ANS-based function, may also contribute to DD symptoms (Desoete et al., 2009). However, cumulative results of recent decades suggest that several cognitive deficits can also be observed in DD, such as working memory deficit, executive function problems, and symbolic processing issues; furthermore, ANS-related deficits are neither the only nor the main problem (Kucian & von Aster, 2015; Rubinsten & Henik, 2009). The problems mentioned above should also be taken into account when re-evaluating the results about the cause of dyscalculia: (a) subitizing is not related to the ANS; therefore, a potential subitizing deficit is not an ANS deficit, (b) paradigms that utilize nonsymbolic sets that can be subitized are not valid measurements of the ANS, and (c) perceptual control is usually suboptimal in DD studies, which leads to invalid results where inhibition and other perceptual abilities may be measured instead of the ANS. Overall, it remains unclear as to how strongly the ANS contributes to DD symptoms.

The ANS is responsible for symbolic comparison even though it cannot solve it. Finally, it is worthwhile to note that the ANS explanation of the symbolic comparison task contains a critical paradox, which the literature usually ignores. According to the ANS account, behavioral performance in the symbolic (e.g., Indo–Arabic) comparison task is caused purely by the ANS: Error rate and reaction time can be entirely explained with the psychophysical model, and no other mechanism should be supposed to account for the behavioral data. However, based on the properties of the ANS, the ANS representation is not able to solve the comparison task by itself, because the Weber fraction observed for typical adults would not allow the ANS to compare even some single-digit number pairs with low ratios. For example, according to a calculation method (Pica et al., 2004), for a 0.14 Weber fraction value (which is typical for adults measured with nonsymbolic stimuli), a person would produce a 22% error rate for the 6 vs. 7 comparison task, 25% for 7 vs. 8, and 28% for 8 vs. 9. These error rates are clearly larger than the typical human performance. We may suppose that symbolic number processing utilizes a smaller Weber fraction; even so, a realistic 0.07 Weber fraction predicts a 12% error rate for the 8 vs. 9 comparison task, which is still too high compared to human performance (Krajcsi, Lengyel, & Kojouharova, 2018). This issue becomes even more serious for larger neighboring numbers if one supposes that multi-digit numbers are also processed by the ANS (Dehaene, 2007; Dehaene et al., 1990). The limitation of the ANS in solving precise numerical tasks has been acknowledged and other representations have been proposed to account for symbolic precise numerical tasks, such as the verbal word frame in the triple code model (Dehaene, 1992, 1997), but the model still supposes that behavioral performance in symbolic number comparison is led purely by the ANS, i.e., a system which is unable to solve the task. To summarize, one key proposal of the ANS framework is that behavioral performance in precise symbolic comparison is exclusively lead by an imprecise system that is unable to solve the task. While additional mechanisms may help to solve the task, these additional mechanisms, which are critical in solving the task, are invisible in terms of observable behavior. Normally, the scientific community would not accept such an argument. This partial contradiction of the model is also reflected in another incoherent detail of the ANS account: While the ANS is generally assumed to support imprecise and approximate calculations, precise comparison is an exception. In defense of the original ANS proposal, one may question the ANS's inability to predict comparison performance, as the ANS model has been repeatedly demonstrated to be able to correctly predict the symbolic comparison performance (Dehaene, 2007; Moyer & Landauer, 1967). In fact, while the overall fit of the ANS prediction to the symbolic comparison data is relatively good, there are specific details of the predictions that fail to capture important parts of the empirical data, such as the performance with small-ratio number pairs (Krajcsi, Lengyel, & Kojouharova, 2018). One may also raise the issue that an additional precise mechanism may not be invisible in terms of observable behavior, but it contributes to the behavioral performance, and the ANS is responsible for only a part of behavioral performance. This argument supposes that behind the observed performance there is an unspecified component (generated by the precise mechanism), while the rest of the performance is the ANS-predicted behavioral pattern. The main problem with this idea is that the unspecified component is entirely speculative without any empirical or theoretical support; as a consequent, the ANS performance pattern becomes entirely speculative, which would mean that this reasoning undermines the original and critical argument for the ANS, i.e., the observable ratio effect in the comparison task. In other words, if the ANS predicted behavioral pattern cannot be observed directly in the symbolic comparison task, it raises the question of why an ANS would be there in the first place. To summarize, the ANS's account of symbolic precise number comparison is questionable, since typical human performance is better than the performance that the imprecise ANS predicts and since it is supposed that behavioral performance is led purely by a representation that cannot solve the problem. This issue is critical because it undermines the central reasoning that supports the ANS model.

Phenomena	**ANS-based explanation**	**Alternative explanations with selected references**
Subitizing	Small representational overlap for small numbers	Pattern detection of sets (Krajcsi et al., 2013; Mandler & Shebo, 1982)
Symbolic matching task	Representational overlap based ratio effect	Visual similarity (Cohen, 2009)
Development of subset-knower children	Improving sensitivity	ANS cannot explain the symbolic number learning trajectory (Krajcsi et al., in preparation)

Phenomena	**ANS-based explanation**	**Alternative explanations with selected references**
Multi-digit comparison	Representational overlap based ratio effect	Multi-step decompositional processing (Hinrichs et al., 1982; Poltrock & Schwartz, 1984)
Nonnumerical features in nonsymbolic tasks	Nonnumerical features can be controlled for	Nonnumerical features can be controlled for but they rarely are; many studies are thus invalid (DeWind et al., 2015; Gebuis & Reynvoet, 2012a)
Dyscalculia	Impaired ANS sensitivity, disconnected ANS, impaired subitizing	Working memory problems, executive function and inhibition issues, symbol processing problems, possibly with ANS-related deficit (Kucian & von Aster, 2015; Rubinsten & Henik, 2009)
Symbolic comparison	Representational overlap-based ratio effect	Noisy ANS is unable to solve a precise task by itself

Table 4. Alternative explanations of formerly ANS-accounted phenomena.

After summarizing several important debates about the supporting evidence for the ANS model, it is worthwhile to highlight that, while one may have the impression that the ANS is supported by a great number of phenomena, which provides a diverse set of evidence for the existence of the ANS, in fact, the list of phenomena can be grouped into a few key features (Table 3), and as summarized in this section, several phenomena have been found to be driven by other representations, and, for many phenomena, it remains inconclusive whether they support the ANS account or not (Table 4).

The discrete Semantic System and the hybrid ANS–DSS account

The core of our alternative account is that symbolic and nonsymbolic numerical information is processed by different types of representations and that symbolic phenomena previously attributed to the ANS can be accounted for by an entirely different architecture. This section first outlines the main properties of such an alternative model and explains how the appropriate symbolic phenomena can be explained in this new architecture (i.e., the DSS model), and how symbolic and nonsymbolic number processings are related (the hybrid ANS–DSS account). Subsequently, we review some other numerical cognition models, which can be considered to be similar or related models to account for different symbolic processing. Note, however, that this section only intends to demonstrate that the alternative model is viable and that it may make the same predictions as the ANS model for most effects; this section will not handle the question of which competing account is more appropriate when it comes to explaining a wider range of phenomena. This question is discussed in the following section.

The Discrete Semantic System

Our alternative model questions the source of the key signature of the ANS model, i.e., the ratio effect: In the ANS model, whenever the ratio effect (or its functional equivalents, the distance or size effects) is observed, it has been assumed that this is a sign that the representation is working according to the psychophysical rules (Dehaene, 2007; Moyer & Landauer, 1967). However, similar effects can also be observed in other cases where the generator is more than likely a different mechanism. In a psycholinguistics study, it was demonstrated that, in a picture naming task, a previous item not only influences the reaction time of the response for the current item, but this priming effect is proportional to the semantic distance between the previous and actual items (Vigliocco et al., 2002). Most probably, in a picture naming task, it is neither a single nor few continuous representations similar to the ANS that generate this semantic distance effect, but rather some other architecture may be responsible for the effect, for example, a network of nodes for the words where spreading activation may cause the observed distance effect. By extending the possibility of a network of nodes responsible for the distance effect, we propose a model that may comprehensively account for other related symbolic numerical phenomena as well (Krajcsi et al., 2016).

Building blocks of the DSS. Our account proposes that while nonsymbolic numerosities may be processed by the ANS, symbolic numbers are processed by an entirely different representation, i.e., a network of nodes, which we term the Discrete Semantic System (DSS). The main features of this alternative representation are rooted in psycholinguistics models and models of conceptual networks. The word “semantic” in the name of the system mainly refers to the systems that inspired the present model, and not to the possibility that this representation could be the main source of numerical meaning. (See additional considerations about the source of meaning in The source of meaning and essential representations for numerical cognition.) In this alternative model, numbers, such as single Indo–Arabic digits, number words, and special multi-digit numbers are represented in nodes, and these nodes are connected (Figure 8). The word “discrete” in the name of the model refers to the fact that the nodes are independent units as opposed to continuous (i.e., analog as in the alternative name of the ANS) representation. The nodes may be notation-specific, i.e., there may be separate nodes for Indo–Arabic numbers, number words, and so on; consequently, related effects may be notation-dependent. The strength of the connection may depend on various features (see a nonnumerical example of a similar semantic network in McClelland & Rogers, 2003). The values of the numbers may influence the connections of the nodes: Numerically closer numbers may have stronger connections with each other (see the width of the connecting lines between the numbers in Figure 8). However, other properties may also have an effect on those connections, such as parity (e.g., even numbers have relatively stronger connections with other even numbers than with odd numbers), primeness, and also nonnumerical properties, such as lucky numbers. Nodes may also reflect the frequency of the appropriate values (see the width of the lines around the specific numbers in Figure 8). The network may include not only the numbers, but also related numerical concepts, such as small, large, even, odd, and related nonnumerical concepts, such as left and right (see some examples of those concepts in Figure 8). These concepts may be connected with the values according to the relevant properties of those values, e.g., larger numbers have stronger connections with the concept “large” than smaller numbers.

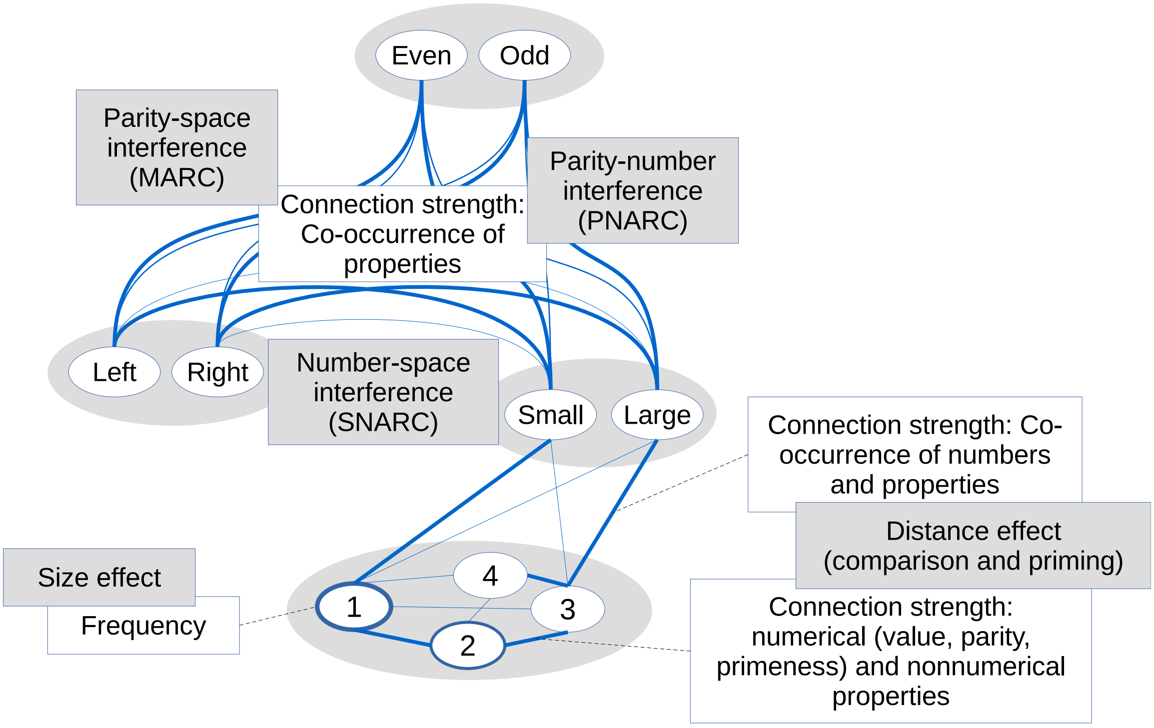


Figure 8. Hypothetical structure of the Discrete Semantic System. Nodes (white ellipses) of the network represent numbers and related concepts. Connections may have different weights, depicted as the width of the lines. Various properties of the network (white squares) may explain the symbolic numerical effects (gray squares).

Accounting for symbolic phenomena in the DSS model

The distance and size effects. This simple architecture may account for the same symbolic phenomena that the ANS accounts for. In the DSS representation, the ratio effect is not a single effect, but is, in fact, two separate effects (i.e., the distance and the size effects) which combine in a way that resembles a ratio effect. The size effect may be a frequency effect: Smaller numbers are more frequent than larger numbers (Dehaene & Mehler, 1992; Rinaldi & Marelli, 2020), and more frequent stimuli may be easier to process, which leads to a size effect. The distance effect may come from at least two sources in this model. One possibility is that the nodes of the values are connected, and the strength of the connection may partly rely on the numerical distance of the nodes (i.e., numerically closer number pairs are connected more strongly), where a spreading activation may cause interference between close numbers, which leads to a distance effect (see the connection strengths between the numbers in Figure 8). Another possibility is that numbers are connected with the nodes of “small” and “large”, and relatively small numbers may have stronger connections with the “small” node, while relatively large numbers may have stronger connections with the “large” node (see the connection strengths between the numbers and the “small”–“large” nodes in Figure 8). Numbers with smaller distances have more similar connection strengths with the “small” or “large” nodes, which could make a decision more difficult, for example, in a comparison task, which leads to a distance effect. This explanation of the distance and size effects provides not only an appropriate qualitative but also a quantitative description. The upper part of Figure 9 displays the quantitative ANS prediction for the performance in a comparison task, where rows and columns are the to-be-compared numbers, and cells represent the difficulty of the task (i.e., larger values and respective darker shades denote more difficult number pairs). In the lower part of the same figure, two possible quantitative descriptions of the distance and size effects can be seen according to the DSS model (left and middle tables, respectively), where combining the two effects produces a prediction for the difficulty in a comparison task (table on the right) that is very similar to the prediction of the ANS. (For a more detailed and technical discussion of various possible quantifications of the model predictions see Krajcsi et al. (2016). See additional relevant properties of the comparison task in the Comparison distance and size effects subsection.)

In symbolic approximate calculations, the distance and size effects can arise from similar mechanisms as in comparison tasks. Similarly, the addition of these effects may form a seeming ratio effect that has been interpreted as a sign of psychophysical representation.

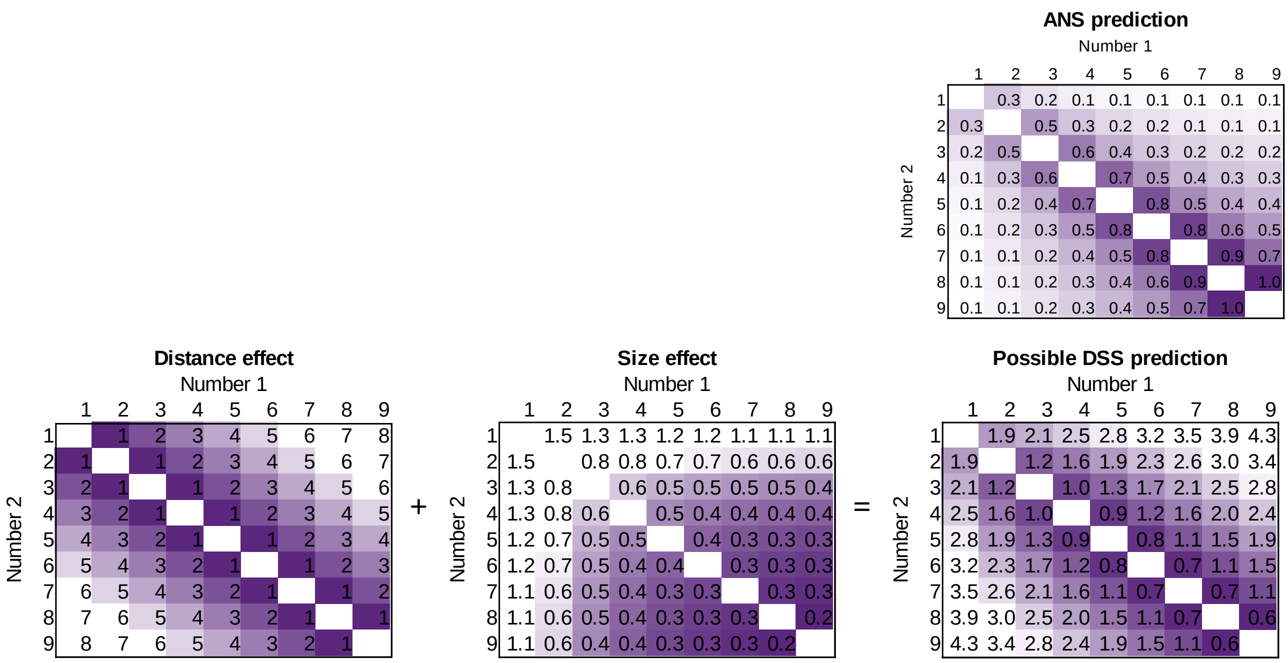


Figure 9. Predicted performance in the symbolic number comparison task according to the ANS model (top) and the DSS model (bottom). The DSS model supposes that the overall performance (right) is the sum of the distance effect (left) and the size effect (middle). The heatmaps display the predicted performance on an arbitrary scale (cells of the heatmap) as a function of the two to-be-compared numbers (rows and columns). Darker shades denote worse performance. The ANS performance values are calculated as $a \times \log(large/distance) + b$, while the DSS performance values are calculated as $a_1 \times distance + a_2 \times (x_1^{-1} + x_2^{-1}) + b$, where *large* is the larger number, *distance* is the distance between the two numbers, x_1 and x_2 are the two numbers, and a, a_1, a_2 and b are free parameters (the parameters a and a_2 are set to 1, a_1 to 0.4, and b to 0). (See more technical details in Krajcsi et al. (2016).)

The priming distance effect (PDE). This effect can be explained by the same connections: either by the connections between the numbers or, if activation spreads bidirectionally between numbers and “small”–“large” nodes, by similar spreading activation which could cause a numerical distance-based priming effect.

Interference effects. These effects, such as the SNARC effect, can be readily explained in the present framework. Numbers may activate relevant numerical properties (such as small, large, even, odd, etc.). These numerical properties and other nonnumerical properties (such as left and right) may be connected, and the interference may come from the connections of the property pairs (see the connected “small”–“large” nodes and “left”–“right” nodes in Figure 8). This model has an advantage over the ANS account, which is that it can include additional interference effects beyond the numerical magnitude interference effects, such as the MARC effect where the parity interferes with the response location (Nuerk et al., 2004) (see the related connection in Figure 8). (See some other relevant models in the Related models subsection.)

Handling composite numbers. Although the DSS may handle some simple numerical processing, numbers beyond integers (such as the comparison of multi-digit numbers or negative numbers) may require additional mechanisms, which could rely on the DSS. For example, multi-digit comparison can be performed via a power-by-power comparison (Hinrichs et al., 1982; Huber et al., 2016; Poltrock & Schwartz, 1984), where within-power operations may be supported by the DSS, while higher-level between-power organizing steps could be handled by other processes. Similarly, in a comparison task, negative numbers can be processed as positive numbers with the additional step of reversing the smaller-larger responses (Huber et al., 2016; Krajcsi & Igács, 2010). (See further examples of the role of the cognitive control processes in number processing in Huber et al. (2016).)

<table>
<tr><th>Symbolic numerical effect</th><th>ANS explanation</th><th>DSS explanation</th></tr>
<tr><td>Numerical distance effect (easier comparison in pairs with a large distance)</td><td rowspan="3">Size of the overlap of noisy signals depending on the ratio of the values</td><td>Strength of the connections between (1) number nodes depending on the semantic relation and/or (2) between number nodes and smaller–larger nodes depending on their conditional frequencies</td></tr>
<tr><td>Size effect (easier comparison in pairs with small values)</td><td>Frequency of the numbers</td></tr>
<tr><td>Priming distance effect (stronger priming effect when the prime is close to the target)</td><td>Strength of the connections between nodes depending on semantic relation</td></tr>
<tr><td>SNARC effect (easier parity decision for small numbers with left response position and large numbers with right response position)</td><td>Interference between the automatic spatial property of the ANS and the response location</td><td>Interference between small–large and left–right nodes</td></tr>
</table>

Table 5. Some symbolic numerical effects and their ANS and possible DSS explanations.

The present list of DSS-based accounts (Table 5) offers explanations for the main symbolic phenomena that has been attributed to the ANS model (Table 3), which have not been invalidated by former accounts (Table 4).

Note that the proposed DSS model not only embraces the symbolic phenomena that has been attributed to the ANS, but it also accounts for additional phenomena that were irrelevant in the ANS model, such as the MARC effect or the semantic congruency effect (SCE). In the SCE, large numbers are chosen faster than small numbers when the task is to choose the larger value, and small numbers are chosen faster when the task is to choose the smaller value (Leth-Steensen & Marley, 2000).

General considerations on the DSS model

The reason why the discovery of the ANS is considered to be astonishing is partly because it contradicted the commonsense idea that mathematical thinking is a high-level, complex, and human- and culture-specific ability. According to the ANS model, mathematical thinking is based on an extremely simple, evolutionarily old mechanism. In that context, the DSS model is a return to common sense: Simple symbolic mathematical operations are based on a conceptual or linguistic-like mental system. In light of common sense, the DSS model is less exciting; it is intuitive to think that abstract mathematical concepts like numbers can be handled by a system that is similar to a conceptual network or the mental lexicon, i.e., by a system that partially processes symbolic and abstract concepts.

The DSS model took inspiration from conceptual and linguistic models. In that context, the DSS does not include any new details at the moment. However, what is new is offering a coherent and comprehensive model that can account for the symbolic phenomena that were formerly attributed to the ANS. While these models that provided inspiration offer alternative explanations and predictions for simple number processing phenomena, they also have limitations. (1) Since relevant conceptual or linguistic models are less quantitative than psychophysical models, the DSS is also less quantitative than the ANS. For example, for the DSS, it is harder to provide a mathematical description of number comparison performance that relies on a consensual description of linguistic or conceptual models, even if those models offer some constraints (Krajcsi et al., 2016). Note, however, that the psychophysical models also have their own limitations, which make them imperfect when applied to number processing (Krajcsi et al., 2016; Krajcsi, Lengyel, & Kojouharova, 2018). (2) The DSS is more flexible than the ANS because a network of nodes is more flexible, and the relevant source models offer various functional descriptions,, e.g., there may be more free parameters or there may be more flexible rules in the DSS than in the ANS. Even so, this does not mean that the DSS model is ad hoc or entirely arbitrary, because it is limited by many properties and suppositions of its source models.

While the DSS model is proposed in this paper as a representation that builds upon an entirely different architecture than the ANS, there is still the possibility of functional similarity between the two models. In its extreme form, it is possible that the DSS is just another formulation or implementation of the ANS model in a sense that one main organizing force of the network is the meaning of the values; while the DSS includes discrete units, overall the nodes will be arranged according to a virtual number line that mimics the ANS. In fact, there are critical computational differences and constraints between the ANS and DSS models that ensure that neither model is an implementation of the other. (1) One defining difference is that while the ANS is a continuous system (as reflected in its alternative name, the Analog Number System, where analog means continuous), the DSS includes discrete nodes, i.e., independent units. Since the ANS is a continuous representation, relevant modification of the representation (e.g., adaptation) is either applied across the whole continuum or across a specific range of the continuum; relatedly, the modification of specific number representations is generalized to neighboring values or to the whole number range.

In contrast, in the DSS model, it is possible that the modifications of some values are not generalized to neighboring numbers, e.g., making the number 13 unlucky will not make either 12 or 14 unlucky. Similarly, in the ANS model, generalization may occur through the neighboring values. However, in the DSS model, generalization may occur through any relevant properties, e.g., parity. (2) Other critical constraints come from the models that inspired the ANS and DSS models. The ANS model is rooted in the psychophysical models; therefore, it inherently follows Weber’s law and related formulas. On the contrary, the DSS model was inspired by the psycholinguistic and conceptual network models, and at least in its initial form, the DSS works according to rules that have been established in those cognitive areas. (3) Following these differences, there are several specific cases in which the two models have different predictions, as discussed in more detail below. To conclude, the DSS is not only a possible implementation of the ANS, resulting in functionally equivalent models, but it is also an entirely different architecture with different functional rules, leading to different predictions in some cases.

Because the DSS model took inspiration from the linguistic and conceptual network models and because the DSS network may include nodes that are not strictly numerical in nature (e.g., lucky numbers and birthday dates), it is a question whether the DSS is a separate numerical network or is the part of the linguistic or conceptual system. At the moment, it is yet unknown whether the DSS is a functionally separate system (e.g., whether it can dissociate from other similar linguistic or conceptual systems or whether individual differences of the DSS correlate with individual differences of similar linguistic or conceptual systems) or is a part of a more general cognitive system. (See more details about the relation of the verbal system in the triple code model and the DSS in the subsection Related models.)

Finally, it should be noted that, while the DSS in its current form includes several specific details, it is by no means a final or complete description of simple symbolic number processing. The DSS model highlights that simple symbolic number processing relies on a network of nodes instead of a continuous imprecise number representation, and the functional description of the DSS builds on similar conceptual and linguistic models. Beyond those main starting points, several details are incomplete (e.g., see the possible various sources of the comparison and priming distance effect), further details may be refined based on new empirical findings (e.g., the notation specificity of the comparison distance effect), or additional details may be clarified based on similar alternative models (e.g., see the Related models section).

To summarize, the DSS is an alternative model that comprises a network of nodes that can account for the elementary symbolic numerical effects that the ANS accounts for.

The hybrid ANS–DSS account

In our alternative framework, the DSS processes only symbolic numbers, and nonsymbolic numerosities are still processed by the ANS. In the classic view, which we term the pure ANS account in this paper, both symbolic and nonsymbolic numbers are processed by the ANS, although the Weber fractions may differ when symbolic and nonsymbolic stimuli are handled (left side of Figure 10). In contrast, in our general framework, which we term the hybrid ANS–DSS account, symbolic numbers are processed by the DSS, and nonsymbolic numerosities are still handled by the ANS (right side of Figure 10). Recently, in line with our proposal, several extensive and critical reviews of the literature have suggested that the relation of symbolic and nonsymbolic processing may not be as strong as previously proposed (Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016).

In the hybrid account, if symbolic input and output are strongly related to the DSS, then it is reasonable that the two representations can communicate, as evidenced by some simple phenomena, e.g., the quantity of a set can be verbalized (i.e., the ANS may send the information to the DSS) or an imprecise number of knocking can be generated based on verbal instruction (i.e., the DSS may send the information to the ANS).

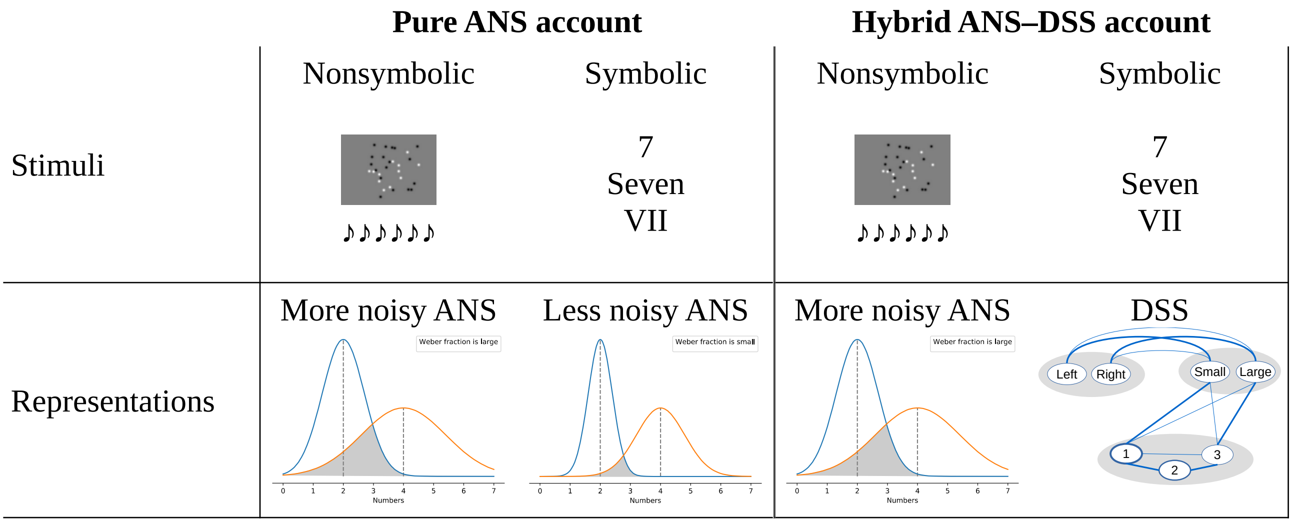


Figure 10. Hypothesized representations that are handling simple numerical operations according to the pure ANS account (left) and the hybrid ANS–DSS account (right).

While, in the pure ANS account, the ANS is responsible for the basis of the semantic understanding of numbers, it is yet to be determined how this role is shared in the hybrid account. Can the ANS play a similarly essential role in the hybrid account as in the pure ANS account? We discuss this question later in the subsection The source of meaning and essential representations for numerical cognition.

Related models

The framework presented in this paper may be similar to some other models formerly introduced in the literature, although clear and essential differences can be described. There are several models for various aspects of basic numerical processing. Some of those models consider simple psychophysical processing from a new perspective (Burr & Ross, 2008; Dakin et al., 2011; Gebuis et al., 2016; Leibovich et al., 2017; Stoianov & Zorzi, 2012) or extend the original ANS model (Walsh, 2003). From our perspective, it is important to note that those works mostly concentrate on nonsymbolic numerosity processing; therefore, they are mostly orthogonal to the DSS model. On the other hand, some other models offer non-psychophysics-based solutions (Banks et al., 1976; Humphreys, 1975; Vigliocco et al., 2002); these models could be in line with the DSS model. In the present subsection we discuss some of the relatively frequently cited models that may be related to our hybrid ANS–DSS account or to the DSS representation itself.

Verbal word frame of the triple code model. The DSS may show some similarity with the verbal word frame of Stanislas Dehaene's triple code model (Dehaene, 1992, 1997). In that model, while the ANS is responsible for the meaningful understanding of numbers, its precision is limited. To offer a mechanism that can handle precise numerical operations, a verbal word frame is proposed. This verbal component is a part of the verbal system and stores precise (i.e., not approximate) numerical knowledge in a linguistic form. It is notable that the verbal component of the triple code model is similar to the DSS, because both of them rely on mechanisms that have been proposed for linguistic models. The main difference between these two models is the range of phenomena that it is proposed they account for: In our framework, the DSS accounts for symbolic phenomena that are accounted for by the ANS in the triple code model and not by the verbal word frame.

Number primitives. Another model in which similarity with the DSS can be observed is the model of number primitives. Joseph Tzelgov and his colleagues investigated whether different forms of numbers are stored in long-term memory (i.e., whether those numbers are primitives—in their terminology). They assumed that, if a number automatically interferes with the physical size of its symbol (automatic numerical size congruency effect), then that number is a primitive. They found that positive single digits and zero show interference with size; thus, they are primitives (Henik & Tzelgov, 1982; Pinhas & Tzelgov, 2012), while negative numbers and ratios are not primitives (Kallai & Tzelgov, 2009; Tzelgov et al., 2009). (See the related issue of composite numbers in the section Accounting for symbolic phenomena in the DSS model.) The model of primitives and the DSS model may overlap in that the primitives in that model and the nodes in the DSS model can be the same constructs. In addition, although the relation of the primitives and the ANS model is mostly not discussed in the aforementioned works, the primitives cannot be reconciled with the ANS model. Beyond these similarities, we note that, while the model of number primitives discusses mostly the automaticity of those numbers, the DSS model and the hybrid ANS–DSS framework offer a more comprehensive explanation for basic number processing phenomena.

Polarity and markedness model. The DSS-based explanation of interference effects is not unique in that similar accounts have been applied for the numerical interference effects. As one of these similar accounts, the polarity model suggests that antonyms have "positive" and "negative" polarity, and this related polarity could cause the interference (Leth-Steensen et al., 2011; Proctor & Cho, 2006). Similarly, according to the markedness model, one of the words in antonym pairs is marked, which leads to not only slower processing of these words, or slower learning (Hines, 1990), but also interference effects (Nuerk et al., 2004; Patro et al., 2014). (See a more detailed comparison of the related interference models in Krajcsi, Lengyel, & Laczkó, 2018). All of those models, together with the DSS model, suggest that discrete units of features (small, large, left, right, etc.) interfere because of common properties or because of specific connections. Interference-specific models, such as the polarity or the markedness model, may offer additional specific mechanisms as to how and why the interference effects emerge. On the other hand, in the DSS-based account, a more comprehensive explanation of numerical effects is provided.

Connectionist model of simple number handling. The DSS model may show strong similarities with the connectionist model, which was built by Tom Verguts and his colleagues to account for various numerical effects (Verguts et al., 2005; Verguts & Fias, 2004, 2008). The model proved to be successful in accounting for a series of effects in various tasks, such as the distance and size effects, the semantic congruity effect, the end effect, and the SNARC effect in comparison, naming and parity tasks, not only in symbolic numbers but also in symbolic nonnumerical order processing (see a more detailed description of those effects and the connectionist models in Gevers et al., 2006; Verguts et al., 2005; Verguts & Fias, 2004, 2008; Verguts & Van Opstal, 2014). In their initial model, there could be two input layers that represent nonsymbolic and symbolic input, and the numbers are processed in a hidden number layer. In that hidden number layer, the nodes are functionally arranged according to their values, and the representation is noisy. When only nonsymbolic input is used, the noise is wider for large values, similar to the functional description of the ANS (Figure 5) and also similar to the tuning curves of some cells observed in neurophysiological studies (Nieder et al., 2002). (For a similar model, see Leth-Steensen & Marley, 2000; and Zorzi & Butterworth, 1999.) However, when introducing the symbolic input, the number layer represents values on a linear scale with fixed width of noise. Note that this latter feature is critical; because of it the network cannot produce the ratio-based performance of the Weber's law by itself (Verguts & Fias, 2004), so the everyday frequency of numbers was introduced to the model to simulate the size effect (Verguts et al., 2005). While the works describing the model are mostly neutral as to whether this model is an ANS model, it has been interpreted as an ANS implementation (Dehaene, 2007). Although it was also highlighted that the model does not rely on the overlap of representations but instead on the connections of the nodes, even so, this difference was interpreted in terms of additional input–output connections and task-relevant decision components and not in terms of number representational differences (Verguts & Fias, 2008). In addition, it was stated that the number representation of that model is not discrete (Verguts & Fias, 2008). Nevertheless, it is notable that the version accounting for symbolic processes cannot generate the ratio effect by itself (in the ANS model the frequency of the numbers is not relevant), and this effect is the defining feature of the ANS model. Therefore, our interpretation is that this connectionist model cannot be considered an ANS implementation. Alternatively, we propose that this connectionist model can be considered a model of discrete nodes, overlapping with the DSS model (and also overlapping with the model of number primitives). In line with this proposed overlap, the nodes of the connectionist model are similar to the nodes of the DSS model; the ordinal arrangement of the nodes in the number layer is consistent with the semantic relations of the DSS nodes; and, the fixed-width noise of the connectionist model is consistent with a spreading activation between the nodes in the DSS model. Furthermore, the introduction of the frequency of the numbers to account for the size effect in the connectionist model is consistent with the frequency-based explanation of size effects that we outlined above (Figure 8). As a final example, the explanation of the SNARC effect is also similar in the two models: In the connectionist model, the effect comes from the automatic connection of the response layer and the categorical magnitude layer (Gevers et al., 2006), whereas the DSS model supposes the existence of similar connections between the relevant property pairs. In terms of the differences between the two models, on one hand, the connectionist model has a more specific description of the exact processing and learning mechanisms, implemented in a working connectionist model, which detailed description and specificity is lacking in the DSS model for some phenomena. On the other hand, the DSS model and the hybrid ANS–DSS framework offer a broader perspective and a more comprehensive interpretation of the phenomena that has been discussed intensively in the literature over the last few decades. Moreover, while the connectionist model is mostly formulated on a technical level, the DSS model is described on a more psychological and cognitive level.

Connectionist model for more complex number handling. Another connectionist model, which was created by Stefan Huber and his colleagues (2016) aims to account for a series of effects observed in multi-digit, decimal and negative number handling. The model was successful in simulating a series of effects, such as the power compatibility effect, the filler item effect, the proportion congruity effect, the Gratton effect for multi-power processing, the string length congruity effect, and the zero facilitation effect for decimal number processing, the sign-decade compatibility effect, and blocking effect in negative number handling (for more details about these effects, see the text and Table 1 in Huber et al., 2016). The number representations applied in that model were similar to the model used by Verguts and his colleagues (Gevers et al., 2006; Verguts et al., 2005; Verguts & Fias, 2004, 2008; Verguts & Van Opstal, 2014), where numbers were functionally stored on a linear representation with a fixed noise. Importantly, like in the case of the connectionist model by Verguts, this representation cannot produce the ratio effect by itself; therefore, it cannot be considered an ANS model implementation. It is worth noting again that this kind of representation is in line with the DSS model as proposed in this paper.

Model and selected references	**Nature of the representation**	**Main effects the model accounts for**	**Relation to the DSS model**
Verbal word frame (Dehaene, 1992)	Verbal representation	Exact calculation	DSS does not address these effects
Number primitives (Pinhas & Tzelgov, 2012)	Long-term memory items	Number–physical size interference effects	Parallel interference account
Polarity and markedness (Hines, 1990; Proctor & Cho, 2006)	Verbal/conceptual labels with meta properties	Interference effects	Parallel interference account
Simple number handling connectionist model (Verguts et al., 2005; Verguts & Van Opstal, 2014)	Connectionist model	Various effects in comparison, naming, and parity tasks	Parallel account of networks
Complex number handling connectionist model (Huber et al., 2016)	Connectionist model	Various effects for handling multi-digit, decimal and negative numbers	Parallel account of networks

Table 6. Some related models of number processing.

Several of the models discussed here are compatible with the DSS model: The number primitives model, the polarity and the markedness models, and the connectionist models handling simple and composite numbers can all be considered models that account for symbolic phenomena with networks of nodes. Beyond this key feature, many of those models include similar properties, and explain phenomena in a similar manner. On the other hand, these models may contain details that lead to conflicting predictions (see the example of the interference effects to understand how the predictions of the models may differ in Krajcsi, Lengyel, & Laczkó, 2018). In some cases, they account for a partly non-overlapping sets of phenomena. However, these conflicting details do not necessarily mean contradictions between these models, because many of these models are admittedly incomplete (e.g., Huber et al., 2016; Verguts & Fias, 2008), including the DSS (Krajcsi et al., 2016), where details can be refined or modified later. Moreover, these models approach the same topic from different perspectives: While the DSS model and the hybrid ANS–DSS framework offer a broader perspective, the connectionist models provide a more technical description, and the number primitives, polarity and markedness models focus on specific aspects of number processing. To further highlight the uniqueness of the DSS compared to these related models, the DSS was created to re-evaluate the defining feature of the ANS, i.e., the ratio effect, to form an entirely different number representation, and based on those foundations to offer a comprehensive framework for a wide range of number processing phenomena. Overall, we believe that these models account for overlapping sets of phenomena and that, in the long term, these details could be integrated into a unified model. Future studies may find solutions to resolve the contradicting details of those models or the missing pieces of the present descriptions.

Contrasting the two frameworks

The previous section described the alternative DSS model and the hybrid ANS–DSS framework and stated that symbolic effects previously attributed to the ANS can also be generated by the DSS. While those arguments demonstrate why such an alternative account is viable, they cannot decide which model is more consistent with the empirical data. One main issue is that the two models have similar or indistinguishable predictions for most of the known phenomena. For example, the predictions of the ANS and the DSS for symbolic number comparison performance could be very similar (Figure 9), with around a 0.9 correlation coefficient between the two predictions (Krajcsi et al., 2016). Therefore, most of the phenomena formerly presented in the literature may not be decisive in terms of either supporting the ANS or the DSS (see a recent example in Rinaldi & Marelli, 2020). To contrast the two accounts, this section lists phenomena for which the two models have different predictions. To anticipate our conclusion, most of these contrasting phenomena challenge the ANS explanation of symbolic number processing but are in line with the DSS model or with the ANS–DSS framework.

Comparison distance and size effects

The presence of the ratio effect in various numerical tasks is the defining key feature of the ANS. Therefore, in the DSS model, it is essential to inspect whether in these tasks, such as the comparison task, the ratio effect is led by a psychophysical mechanism or by an alternative mechanism outlined above (Figure 8 and 9).

On the one hand, in a comparison task, the two models have similar predictions (Figure 9), therefore, simply fitting the model predictions to the measured data and comparing the overall fit of the models to contrast the two models is not sufficient (Krajcsi et al., 2016). This similarity also explains why former tests of the ANS prediction that were fitted to the data (Dehaene, 2007; Moyer & Landauer, 1967) cannot be conclusive: If the two predictions are highly similar and if either of them is correct, then the other model will necessarily fit relatively well. On the other hand, in a comparison task, the two accounts have at least two basic differing suppositions (Figure 11). First, while the pure ANS account supposes that both symbolic and nonsymbolic comparison is handled by the ANS, the hybrid ANS–DSS account proposes that only nonsymbolic numerosities are processed by the ANS, but symbolic numbers are handled by the DSS. Second, in the ANS representation, the distance and size effects are, in fact, different measurements of the same ratio effect, while, in the DSS representation, the distance and size effects are rooted in different mechanisms. In a series of studies, we contrasted several aspects of those predictions to discover new properties of number comparison tasks (see a more detailed summary of those studies in Krajcsi et al., submittedb).

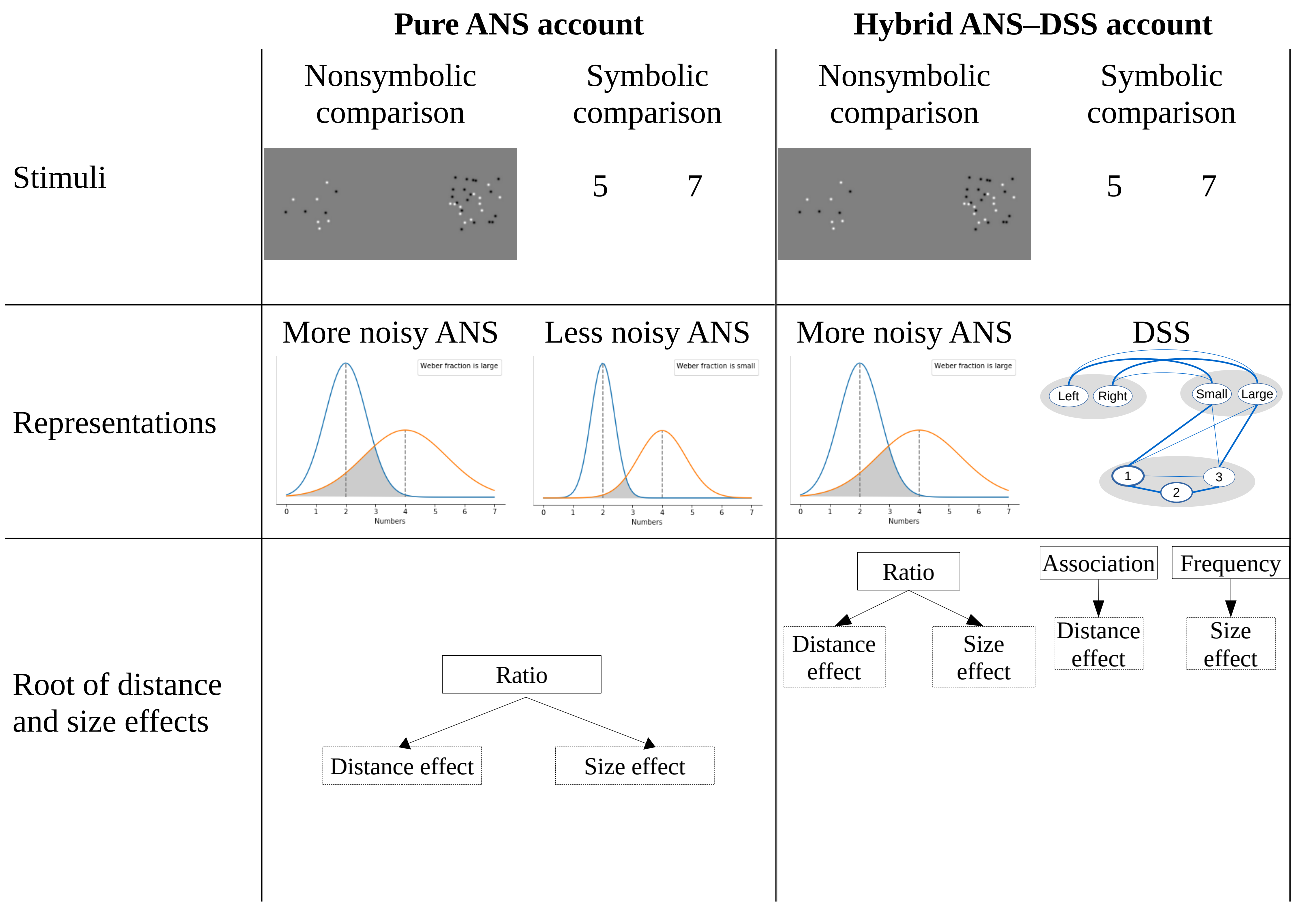


Figure 11. Differing predictions about symbolic and nonsymbolic comparisons in the pure ANS and hybrid ANS–DSS accounts.

In those studies (see a summary of those findings in Table 7), first, it was found that in the symbolic comparison task, the size effect is indeed a frequency effect as predicted by the DSS model: When manipulating the frequencies of the digits in the comparison task, the size effect followed the frequency manipulation either entirely when using new artificial symbols (Krajcsi et al., 2016) or partly when using well-known Indo–Arabic numbers (Kojouharova & Krajcsi, 2019). In addition, when the size effect was manipulated (e.g., in conditions, where frequency was uniform, the size effect disappeared), the distance effect remained unaltered (e.g., even if the size effect disappeared, the distance effect was still observable), which suggests that the two effects are independent in a symbolic comparison task.

Second, in symbolic comparison tasks, the distance effect did not follow the value of the numbers, but the frequency of the associations between the numbers and the concepts of “large” and “small” in that session, which is possible only in the DSS model and not in the ANS model (Kojouharova &

Krajcsi, 2018; Krajcsi & Kojouharova, 2017). While the values and the frequencies of associations normally strongly correlate,we omitted specific digits in a paradigm (e.g., presenting only the values of 1, 2, 3, 7, 8, and 9), dissociating the value and the associations, which makes it possible to investigate the independent effects of the two factors. It was found that, in symbolic comparison, it is the association that leads the distance effect and not the numeric value(Kojouharova & Krajcsi, 2018; Krajcsi & Kojouharova, 2017), which further confirms the DSS model.

Third, when analyzing the appropriateness of the psychophysical model for both nonsymbolic and symbolic comparison, we found that, while the model describes the nonsymbolic comparison without considerable biases, it fails to precisely describe symbolic comparison (Krajcsi, Lengyel, & Kojouharova, 2018). One important finding was that the ANS model predicts too high error rates for small-ratio number pairs (see the parallel argument why the imprecise ANS cannot account for small-ratio number pairs in the subsection Debates about phenomena supporting the ANS account). While former findings showed the overall fit of the ANS model to be satisfying (e.g., Dehaene, 2007; Moyer & Landauer, 1967), our results demonstrated that the prediction for the specific small-ratio number pairs is highly biased. Another essential finding is related to the drift rates in the diffusion analysis. In the diffusion model, it is assumed that the behavioral response is led by an accumulation mechanism, where various parameters of this process affect the behavioral response (Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). For example, the ability to solve a task efficiently is directed by the drift rate parameter, which is the speed of the accumulation. The psychophysical model predicts that as the ratio approaches the value where the two stimuli are hard to distinguish, the drift rate should approach zero (Dehaene, 2007; Palmer et al., 2005). This prediction is in line with the simple observation that as the two stimuli become more similar, it becomes harder to differentiate them and that, when the two stimuli have equal critical property, the discrimination becomes impossible. However, our detailed analyses revealed that symbolic comparison does not approach the 0 drift rate value as the ratio approaches the value representing harder task; rather it approaches the 0.2 value, which reflects the fact that humans can compare even small-ratio symbolic number pairs. This property further highlights the critical limitation of the ANS model as discussed above.

Fourth, the distance and size effects are independent in symbolic notations, while they are strongly connected in nonsymbolic stimuli, which further supports the hybrid ANS–DSS account. One finding that demonstrates this property is the frequency-based size effect manipulation study, which was described above: In those studies, while the size effect was manipulated, the distance effect remained unaltered in the symbolic comparison task, which shows that the two effects are independent (Kojouharova & Krajcsi, 2019; Krajcsi et al., 2016). Similarly, in the studies discussed above where the distance effect was manipulated (Kojouharova & Krajcsi, 2018; Krajcsi & Kojouharova, 2017), the size effect still depended on the frequencies of the digits, which further demonstrates the independence of the distance and size effects.

Fifth, the slope of the distance and size effects and their correlation were investigated in symbolic and nonsymbolic comparison tasks (Krajcsi, 2017). It was found that after correcting for the reliability-based attenuation, in nonsymbolic comparison, the correlation of the distance and size effects was practically 1, demonstrating that the two effects originate from a single source, supposedly from the ANS-based ratio effect. On the other hand, in the symbolic comparison task, the correlation of the distance and size effect slopes did not differ from 0, which shows that those effects in symbolic number processing are independent, as predicted by the independent distance and size effects in the DSS model (Krajcsi, 2017). This correlational pattern is in line only with the hybrid ANS–DSS account.

Sixth, in symbolic number comparison tasks, the distance and size effects show different flexibility in terms how strongly the statistics of the actual session may influence the effects. By utilizing Indo–Arabic numbers, for which a long period of experience may form the distance and size effects, it was found that the distance effect mainly follows the statistics of the current session and mostly ignores former experiences (Kojouharova & Krajcsi, 2018). In contrast, in the same Indo–Arabic notation, the size effect is only partly influenced by the current session statistics (Kojouharova & Krajcsi, 2019). Thus, while the distance effect is highly flexible, the size effect is only moderately flexible. Because the ANS model supposes that the two effects are rooted in the single ratio effect, that model would predict similar flexibility; therefore, the difference in flexibility of the symbolic distance and size effects is another piece of evidence that supports the DSS model.

	Symbolic		**Nonsymbolic**	
	Distance effect	**Size effect**	**Distance effect**	**Size effect**
Source of the effect	Large-small association of the numbers (Kojouharova & Krajcsi, 2018; Krajcsi & Kojouharova, 2017)	Frequency of the symbols (Kojouharova & Krajcsi, 2019; Krajcsi et al., 2016)	Psychophysics model: Ratio effect according to Weber’s law (Krajcsi, Lengyel, & Kojouharova, 2018)	
Independence (Kojouharova & Krajcsi, 2018, 2019; Krajcsi et al., 2016; Krajcsi & Kojouharova, 2017)	Dissociation of the effects		No dissociation of the effects is observed	
Correlation of the slopes (Krajcsi, 2017)	Independent		Strongly correlated	
Flexibility for the statistics of the stimuli	Highly flexible (Kojouharova & Krajcsi, 2018)	Moderately flexible (Kojouharova & Krajcsi, 2019)	Rigid?	

Table 7. Summary of some key properties of symbolic and nonsymbolic comparisons. The question mark in nonsymbolic flexibility means that the property has not been tested yet.

To summarize, all of these results (Table 7) demonstrate that (a) symbolic and nonsymbolic comparisons differ in several critical features and (b), while the distance and size effects are similar within the nonsymbolic comparison tasks, they are different within the symbolic comparison tasks. Both of those findings seriously challenge the pure ANS account, but they are in line with the hybrid ANS–DSS account (Figure 11).

These results also show why the ANS model seemed appropriate when it came to explaining the symbolic comparison task: In most cases, the psychophysical model offers a satisfying description, and special tests are needed to find phenomena where the ANS and DSS models have different predictions.

These recently described properties of the comparison distance and size effects (Table 7) highlight that the two effects do not rely directly on the meaning of the numbers but on some statistical properties of the stimuli (which may correlate with the values). Consequently, despite the widespread view, the symbolic distance and size effects do not necessarily capture the semantic processing of the numbers. For example, when the distance effect is observed, it does not necessarily mean that the participant processed the meaning of the symbol. More generally, the effects that the DSS model may account for are not necessarily a sign of semantic processing. (See a more detailed discussion of the potential role of the DSS and related representations in representing the meaning in the section The source of meaning and essential representations for numerical cognition.)

Contrasting additional properties of symbolic number processing

Symmetric priming distance effect (PDE). According to the ANS model, the size of the PDE is proportional to the overlap of the prime signal and the target signal on the ANS (Figures 5 and 7), therefore, the overlap should be asymmetric in terms of whether the prime is smaller or larger than the target (supposing that the two primes are equally distant from the target), because the larger prime has larger noise and, consequently, larger overlap with the target than the smaller prime (Dehaene, 2004). For example, 7 should be a stronger prime to target 5 than the prime 3, because the noise of 7 is larger than the noise of 3; consequently, the overlap between 7 and 5 is larger than the overlap between 3 and 5. By contrast, the DSS model is more flexible, and the PDE could be symmetric in symbolic notation as, for example, in the connectionist model of Verguts et al. (2005). In several empirical works, the PDE was consistently found to be symmetric in symbolic tasks (see a review of these findings in Verguts et al., 2005), which is in line with the DSS account. The symmetry of the PDE is another challenge for the ANS model.

Unrelated comparison distance effect (CDE) and PDE. In the ANS explanation both the comparison and the priming distance effects are the result of the overlap of the signals in the ANS. Therefore, the two effects should be related. In contrast, in the DSS model, the two effects could be unrelated (Figure 8). As an example of the unrelated CDE and PDE effects, in the connectionist model described above, the PDE comes from the noisy representation of a number within a layer, while the CDE could come from the connections between the number layer and the response layer (Verguts et al., 2005). It was found that the CDE and the PDE do not correlate (Reynvoet et al., 2009), which contradicts the ANS explanation. Note, however, that the observed correlations may be attenuated by the reliability of the variables, and because the reliability of those indexes was not reported in that work, one cannot be confident whether the lack of the correlation comes from the lack of the relation between the CDE and the PDE or from the low reliability of those indexes (where PDE is an effect with a relatively low effect size which makes it more likely to have low reliability). Because of this ambiguity in the interpretation of the results, it is too early to conclude whether the CDE and the PDE are independent in symbolic notation, but nevertheless the phenomenon is another possibility to contrast the ANS and the DSS accounts.

Numbers interfere with discrete properties. When contrasting the ANS and DSS models, one key supposition is that, in the ANS-related models, interference may emerge because interfering properties are stored in similar continuous representations (Cantlon et al., 2009; Cohen Kadosh et al., 2008; Henik et al., 2012; Walsh, 2003); however, the DSS model supposes discrete properties and related discrete representations behind the interference effects. One main issue when contrasting the two models is that all relevant properties involved in the known interference effects, such as numerosity, space, and duration, can be represented in both a continuous and discrete representation (e.g., spatial location can be either a gradual position along a line or a “left”–”right” category pair), so the existence of those interference effects cannot specify what kind of representation they may be rooted in. A possible test case is the possible existence of interference effects where discrete properties interfere with numerical information. Such interference was observed in a SNARC effect where the task is not solved by using two response locations but by using a yes–no verbal response (i.e., is the displayed digit even?) which responses cannot be continuous (Landy et al., 2008). It was also shown that numerosity and parity can interfere (Krajcsi, Lengyel, & Laczkó, 2018), which further suggests that symbolic numbers are at least partially processed by discrete representation. Note that the contrast involving the existence of numerical interference with discrete properties is valid only if the similar continuous representation supposition of the ANS model is correct. However, in a modified ANS model one can propose that the continuous ANS representation may also interfere with discrete representations. See a more detailed discussion of the relevant assumptions and their relevance in Krajcsi, Lengyel, & Laczkó (2018).

Discrete vs. continuous representation. As discussed in section The Discrete Semantic System, one critical feature that differentiates the ANS and DSS models is whether the representation is continuous, where values are stored continuously along a dimension and where properties may be generalized along a wider range, or whether the representation is discrete where, even if the values are related, they are structurally more independent than in a continuous representation. A brain imaging study found differing properties of symbolic and nonsymbolic number representations that can be interpreted in the discrete vs. continuous dichotomy: In a match-to-sample task, Lyons, Ansari and Beilock (2015) found that, while, in nonsymbolic numerosity pairs, the voxelwise correlation pattern reflects the representational overlap that the ANS model predicts, in symbolic number pairs, the correlation did not depend on the ratio of the pairs but displayed a constant value. This latter result was interpreted as a sign of a discrete, categorical representation. Another line of evidence for discrete number representation in the symbolic number phenomena, which was formerly accounted for by the ANS, is number interference with discrete properties as summarized above: If interference effects are rooted in the same type of representations (i.e., either discrete or continuous), then an interference with discrete properties, which are most probably not stored on a continuous representation, suggests that the number representation is also discrete.

Symbolic vs. nonsymbolic processing

A large number of works have investigated the relation of symbolic and nonsymbolic notations and the similarity and differences of symbolic and nonsymbolic number processing. These possible relations and differences could be important test cases for contrasting the pure ANS and the hybrid ANS–DSS accounts (Figure 10): While the pure ANS account highlights the similarities, the hybrid ANS–DSS account allows for more differences. However, it is important to be aware that not all differences or independence between symbolic and nonsymbolic number processing are evidence against the pure ANS account. Although it is usually highlighted that the ANS is a multimodal representation, which works independently of the notation of the stimuli, some of the notation-specific results can be readily explained in a modified version of the pure ANS account. First, in the pure ANS account, symbolic and nonsymbolic number processing may work with different Weber fractions, which results in some differences between symbolic and nonsymbolic tasks. Second, relatedly, there may be two independent ANSs dedicated separately to symbolic and nonsymbolic processing (which we term the dual ANS model), which may explain the independence of the symbolic and nonsymbolic numerical performance.

Lack of priming between symbolic and nonsymbolic notations. While the single ANS model assumes that all notations should prime each other (Koechlin et al., 1999), the DSS approach assumes that priming could depend on the notation mostly because symbolic and nonsymbolic numbers are stored in separate systems. Koechlin et al. (1999) found that, while Indo–Arabic and written words prime each other, arrays of dots and previous symbolic notations do not prime each other. Furthermore, in a masked priming task, none of the notations prime each other. These results are difficult to reconcile with the single ANS account (Koechlin et al., 1999), while they could be in line with the hybrid ANS–DSS approach. Note again that the dual version of the ANS model may allow for the symbolic and nonsymbolic notations to not prime each other.

Independent symbolic and nonsymbolic performance. It was found by several studies that symbolic and nonsymbolic comparisons do not correlate in either children (Holloway & Ansari, 2009; Sasanguie et al., 2014) or adults (Krajcsi, 2017). Note, however, that, in those developmental studies, set sizes from the subitizing range (i.e., 1-4) were also applied in the nonsymbolic comparisons, therefore, the validity of some of those data is questionable. Similarly, in an fMRI study, it was found that the size of the neural activations related to nonsymbolic and symbolic number processing did not correlate (Lyons et al., 2015). Like in the previous case, while these results cannot be reconciled with the single ANS model version, they may be in line with the dual ANS model.

Notation specific representations in fMRI data. Previous fMRI studies revealed areas in the intraparietal sulcus that are activated by processing numbers independent of their notations, which supports the ANS model (Eger et al., 2003; Piazza et al., 2004). However, recent works with more sensitive analysis have revealed that the areas of the intraparietal sulcus are notation-specific, and notation-independent areas were not found, which demonstrates that symbolic and nonsymbolic numerical information are processed differently (Bulthé et al., 2014, 2015; Damarla & Just, 2013). (For a critical evaluation of earlier contradicting findings, see Bulthé et al., 2015.) Again, the dual ANS model still may account for these results.

To summarize, several previous studies have demonstrated that symbolic and nonsymbolic number processing is unrelated. These results cannot be explained by the single ANS model. However, all of these results could be reconciled with a dual ANS model. Still, even if one may want to preserve the ANS model proposing a dual version, it should be noted that arguments for dual ANSs and those for a single ANS (e.g., to propose a multimodal representation) need to be coherent and cannot lead to inconsistency in the model. See additional similar examples of unrelated symbolic and nonsymbolic processing in the reviews of Leibovich & Ansari (2016) and Reynvoet & Sasanguie (2016).

Predicting general mathematical abilities

Empirical works also investigate the differing effects of symbolic and nonsymbolic processing on more general mathematical performance, such as school math performance, everyday mathematical abilities, and developmental dyscalculia. These questions are less essential from the viewpoint of the pure ANS account because, even if there are differences in how important elementary symbolic and nonsymbolic number processing is in higher-level math performance, those effects can only originate from the ANS. However, in the hybrid account these differences are substantial: As mentioned above (see section The hybrid ANS–DSS account), a relevant question is what roles the ANS and DSS representations have in the hybrid model.

Predicting overall math achievement. Although some of works have proposed that nonsymbolic performance correlates with mathematical achievement, which suggests that the ANS is one of the main roots of mathematical abilities (e.g., Halberda et al., 2008, 2012; Lourenco et al., 2012), according to a systematic review, symbolic comparison correlates with children’s mathematical achievement, but nonsymbolic comparison and mathematical achievement mostly do not (De Smedt et al., 2013). Similarly, in an extensive meta-analysis, it was concluded that nonsymbolic comparison performance correlates less with math achievement than does the symbolic comparison task (Schneider et al., 2017).

Dissociation of symbolic and nonsymbolic numerical performance in developmental dyscalculia (DD). In a review of ten empirical studies, Noël and Rousselle (2011) found that children living with DD perform worse than typically developing children in both symbolic and nonsymbolic tasks only after the ages of 9–10 years, while younger children living with DD show impaired performance only with symbolic numbers. The authors reason that symbolic and nonsymbolic processing are supported by different representations, and the impairment of nonsymbolic performance is only the consequence of the primary symbolic impairment.

Overall, studies measuring the effect of elementary numerical abilities on general mathematical performance find that symbolic basic number processing is more important than basic nonsymbolic number processing. Several considerations should be taken into account in such studies. When measuring elementary nonsymbolic numerical abilities, one should consider whether numerical, perceptual, inhibitory, or other abilities are being measured (Leibovich & Ansari, 2016), because invalid measurement could show either increased or decreased ANS effect compared to the real effect. Furthermore, while the relative role of symbolic and nonsymbolic elementary number processing is not a key issue in the pure ANS account, it is a crucial question in the hybrid ANS–DSS account.

Phenomena	**ANS (or pure ANS) prediction**	**DSS (or hybrid ANS–DSS) prediction**	**Relevant finding and selected references**
Symmetry of priming distance effect (PDE)	PDE depends on the ratio of the values, PDE should be asymmetrical	Priming effect depends on the semantic relation of the values, PDE could be flexible	PDE is symmetrical (Verguts et al., 2005)
The relation of the comparison distance effect (CDE) and the priming distance effect (PDE)	Both CDE and PDE depend on the same representational overlap; CDE and PDE are related	CDE and PDE could come from different mechanisms	CDE and PDE do not correlate (Reynvoet et al., 2009)
Interference between numbers and discrete properties	Numbers interfere only with continuous properties	Numbers can interfere with discrete properties	Numbers interfere with yes–no responses and parity (Krajcsi, Lengyel, & Laczkó, 2018; Landy et al., 2008)
Continuous vs. discrete representation	Continuous representation	Discrete representation	fMRI voxelwise pattern depends on the ratio in nonsymbolic but not in symbolic notation (Lyons et al., 2015)
Independence of symbolic and nonsymbolic number processing	Numbers are processed by the same system, independent of their notations	Symbolic and nonsymbolic numbers are processed by different systems	Symbolic and nonsymbolic notations do not prime each other (Koechlin et al., 1999)
			Symbolic and nonsymbolic performance are independent (Holloway & Ansari, 2009; Sasanguie et al., 2014)
			Notation-specific areas in fMRI studies (Bulthé et al., 2014, 2015; Damarla & Just, 2013)
Predicting general mathematical abilities	Symbolic and nonsymbolic basic numerical abilities predict general math abilities equally or the difference is irrelevant	Symbolic elementary numerical abilities may be more important than nonsymbolic abilities in symbolic general math abilities	Symbolic simple numerical performance predicts general math ability stronger than nonsymbolic simple numerical performance (De Smedt et al., 2013; Schneider et al., 2017)
			Constant symbolic impairment in developmental dyscalculia, but only late nonsymbolic impairment (Noël & Rousselle, 2011)

Table 8. Findings contrasting the ANS and DSS models. All of the listed phenomena support the DSS model and not the ANS model. (See more empirical findings about the comparison tasks favoring the DSS model in Table 7. See the text for more details.)

Contrasting the two frameworks – Conclusions

This section discusses a series of studies that demonstrated that, in symbolic number processing, several phenomena are in line only with the DSS model and not with the ANS model and that several findings can be reconciled only with the hybrid ANS–DSS model and not with the classic pure ANS model. These results suggest that processing symbolic numbers are supported by the DSS, and nonsymbolic numerosities are processed by the ANS. Furthermore, several findings demonstrate that nonsymbolic numerosity processing is less relevant in general mathematical knowledge than symbolic number processing, which means that the ANS has a more limited role than previously supposed.

Summarizing the review of the phenomena that have supported the ANS account (Table 3), those phenomena can be re-evaluated. First, some of the results believed to support the ANS are, in fact, generated by other mechanisms, often by a nonnumerical mechanism (Table 4). Second, the DSS can account for a wider range of basic symbolic number processing than the ANS (Tables 5, 7, and 8). Relatedly, the hybrid ANS–DSS account can account for a wider range of phenomena than the pure ANS account (Tables 7 and 8). Third, many nonsymbolic number processing effects are still explained by the ANS model (Table 9), although the role of the ANS is more limited than previously believed (Table 8). Fourth, there may be phenomena that can be rooted in several different mechanisms; in these cases both the ANS and the DSS (or other additional mechanisms) may play some role (Table 9). Table 9 summarizes this re-evaluation of the findings in the literature based on the considerations described in the present work.

Phenomena	**Explanation in the hybrid ANS–DSS account**
Phenomena accounted for by neither the ANS nor the DSS (see also Table 4)	
Subitizing	Pattern detection of sets
Symbolic matching task distance effect	Visual similarity
Development of subset-knower children	Unknown, but not the ANS
Multi-digit comparison	Multi-step decompositional processing
Phenomena accounted for by the DSS (see also Tables 5, 7, and 8)	
Symbolic comparison distance and size effects	Frequency of the numbers and the association of the numbers and small-large categories
Symmetry of priming distance effect	Flexible connections of nodes
Independence of the comparison and priming distance effects	Different sources for the two effects
Interference of numbers and other properties	Interference of property pairs
Phenomena accounted for by the ANS	
Nonsymbolic comparison ratio effect	Representational overlap-based ratio effect

Phenomena	**Explanation in the hybrid ANS–DSS account**
Nonsymbolic approximation	Representational overlap-based ratio effect
Phenomena with multiple and/or uncertain sources (see also Table 4)	
Source of general math abilities	Several predictors, but symbolic processing is more important
Dyscalculia	Various potential deficits, but symbolic processing is more important

Table 9. Summary of alternative explanations for key phenomena formerly attributed to the ANS.

The source of meaning and essential representations for numerical cognition

The pure ANS account suggests that the ANS is an essential base of the meaning of numbers, that it is a strong predictor of general mathematical ability, and that severe impairment of the ANS may lead to developmental dyscalculia. However, in the hybrid ANS–DSS account, the DSS is not central in providing the semantic base for number understanding. Several phenomena discussed so far have demonstrated that the DSS does not rely on the meaning of numbers but instead on the statistical properties of the stimuli. For example, the comparison size effect relies on the frequencies of the digits, and the distance effect relies on the association between the digits and the nodes of “small” and “large” (see subsection Comparison distance and size effects). Overall, this means that there is a sharp distinction between the ANS and DSS explanations: While the ANS explanation supposes that the ratio effect reflects that the meaning of the values is accessed, in the DSS explanation, the ratio effect only reflects that some nonnumerical, statistical information was processed. In other words, the ANS model assumes that a comparison task is a semantic task, while the DSS is an associative network, and the comparison task may be a mechanic task that may be blind to the meaning of numbers. (Note again that the word “semantic” in the name of the DSS does not refer to semantic processing but instead to the similarity of the representation and other networks processing semantic information, such as words and concepts.) This non-semantic nature of the DSS has several consequences. First, while the literature usually assumes that measuring the supposedly ANS-related effects assesses semantic processing, those effects may reflect simple non-semantic, statistical features. Second, if the DSS is not the source of the meaning of number understanding, then another system or systems should provide semantic component. To further illustrate the non-semantic nature of the DSS, if, in a comparison task, the associative network relies on the frequencies of the stimuli, then the system is unable to solve the task correctly before those frequencies are stored in the representation, which means that the correct responses should initially come from another system or mechanism. Overall, if the DSS is not a mechanism that can store the meaning of numbers, where does this meaning come from?

In the hybrid ANS–DSS account, one possible solution to explain where numerical meaning can come from is to keep the original ANS proposal; therefore, the ANS could be the source of meaning, even if several symbolic operations are supported by the DSS. However, there are essential details that an ANS anchoring cannot explain. First, as mentioned above (see section Predicting general mathematical abilities), according to some empirical results, the ANS is limited in accounting for higher-level mathematical achievements (De Smedt et al., 2013; Schneider et al., 2017). Additionally, the ANS may not be appropriate to support discrete values, as repeatedly argued by Susan Carey (Carey, 2004, 2009; Carey & Barner, 2019). Find additional arguments why the ANS is limited in providing the meaning of numbers in recent reviews (Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016) Therefore, we conclude that while the ANS may play some role in creating the base of number meaning, its role must be limited.

To form a hypothetical alternative explanation for how meaning can be handled in the hybrid ANS–DSS framework we offer a possible description based on linguistic and conceptual models (Figure 12). Similar to some linguistic or conceptual models, we hypothesize three possible sources of meaning. (1) First, there may be a perceptual anchoring. (1a) One obvious candidate for such anchoring could be the ANS. While we have argued above that the role of the ANS is more limited than previously proposed by the pure ANS account, this does not mean that the ANS does not have any role in creating the semantic bases of numbers. This kind of anchoring (and equivalently the connection between the ANS and DSS representations) seems to be necessary when accounting for the symbolic estimation of nonsymbolic stimuli (e.g., there are approximately 50 people on the street). (1b) A parallel perceptual anchoring could be the subitization of small sets of items, where pattern detection may support recognizing the sets (Krajcsi et al., 2013; Mandler & Shebo, 1982). (1c) According to some views, embodiment is considered to be another way to anchor the meaning of numbers. Therefore, we conclude that while the ANS may play some role in creating the base of number meaning, its role must be limited. Similarly, subitization and embodiment can have some limited role in providing perceptual anchoring for numbers.

(2) Another source of meaning can be the linguistic or conceptual system itself. (2a) Connections within the network can be a potential component of numerical meaning. One possible way of knowing some properties of numbers is to store this knowledge as the strength of the connection between the values and the properties, e.g., number 3 may have relatively strong connection with the properties of “small”, “odd” and “prime” (see an example of a possible implementation in McClelland & Rogers, 2003). For example, while deciding on the parity of a number, one does not divide the number by two and check for the remainder but parity property is recalled (Dehaene et al., 1993; Ito & Hatta, 2004). One possibility is that this knowledge is stored in the DSS. (2b) A related potentially essential contribution of language to create the meaning of numbers is that, while the information from various domain-specific modules (such as the ANS, pattern recognition, or object based information) initially cannot be combined and used together, language may serve as a domain general adhesive to conjoin the information of the domain specific modules (Spelke & Tsivkin, 2001).

(3) There may be another component of numerical meaning that gives some rich conceptual and semantic understanding of numbers in which dynamic and abstract properties of number processing are also handled. For example, how do humans know that adding an item to a set will lead to the number word in the counting list following the number word that describes the original set? Clearly, the simple connections of a DSS-like network cannot handle this situation, because no mechanism of the network can provide the answer for such question. Similarly, the ANS is not necessarily an appropriate candidate either, because it cannot handle the special status of a unit (Carey & Barner, 2019). As another example, how do we know that, if $a+b$ is c, then $b+a$ is also c? Again, a DSS-like

network does not seem to be a suitable mechanism for this kind of knowledge. Regarding the ANS, one might say that the model can offer this knowledge because, assuming that the ANS can handle addition, both a+b and b+a will result in c. However, in the current form of the model, the ANS does not know that the same operands are used in the two additions and that the two additions also give the same result. A higher-level metaknowledge is required to arrive at the rule a+b equals b+a, which is a different mechanism from the model of overlapping representations. This rich semantic knowledge is not easy to capture not only in numerical cognition but also more generally in psycholinguistics and also in cognitive science. While there are no comprehensive models for this rich semantic knowledge, some models may be considered as initial starting points. (3a) In her description, Susan Carey proposes that enriched parallel individuation (i.e., memory for objects extended with long-term memory templates for sets) with counting principles could be the basis of symbolic number meaning, instead of the ANS explanation (Carey & Barner, 2019; Le Corre & Carey, 2007). (See a similar account extended with order handling mechanism in Reynvoet & Sasanguie, 2016.) Discussing the cause of developmental dyscalculia, Butterworth (2010) proposes a similar set-based mechanism as a source of mathematical understanding and a deficit in that system as an explanation for dyscalculia. (3b) Several works of the embodiment movement could be considered a dynamic conceptual analogy for the properties of numbers, where conceptual-semantic properties of body use is transferred to understanding number processing, such as how using the hand may contribute to understanding how larger powers consist of smaller powers. (3c) Another related model that attempts to explain why handling multi-power place-value number notations (such as the Indo–Arabic number system) is difficult proposes that multi-power numbers are considered as the number of items, groups, and super-groups, and the difficulty of multi-power notations depends on how straightforward the translation between a number notation and this group-based representation is (Krajcsi & Szabó, 2012). Similarly, in a study where children were able to handle zero in operations but they were not sure whether zero was a number, it was proposed that natural number understanding may be rooted in the properties of objects: While the lack of objects can be handled in operations, the lack of the objects may not have an assigned value, zero cannot therefore be thought of as a number (Krajcsi et al., submitteda).

Overall, we speculate that the meaning of numbers and related mathematical concepts may be rooted in several components; in this framework, the ANS is neither the main nor an essential component, and the DSS is neither the only nor the most important component of meaning either.

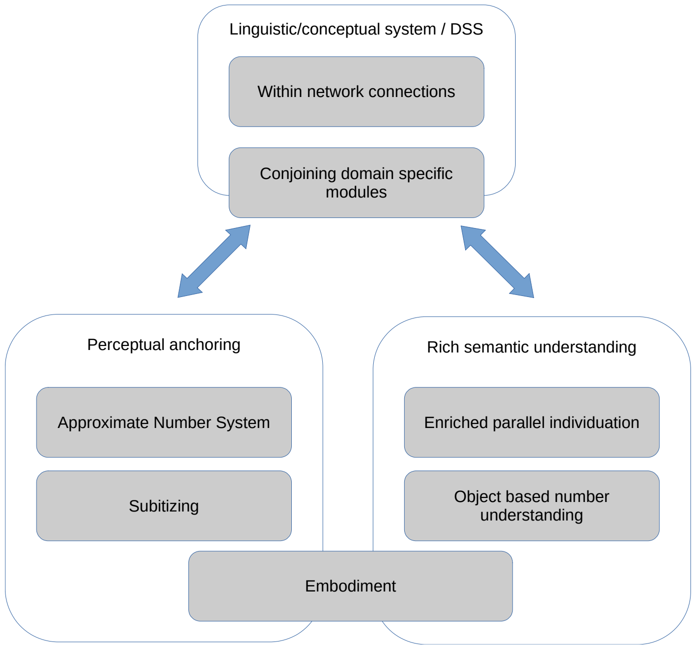


Figure 12. Hypothetical outline for possible sources of numerical meaning.

Another consideration on the role of the DSS in supporting meaning leads to a seeming contradiction. Several previous works have suggested that simple symbolic numerical performance is a more important predictor of more complex mathematical performance than simple nonsymbolic performance (see section Predicting general mathematical abilities). We argue that, because simple symbolic numerical performance is supported by the DSS, those former results may be interpreted as the DSS being more important in predicting more general math performance than the ANS. Combining this interpretation with the limited role of the DSS in supporting meaning may lead to a seeming paradox, which is that the asemantic DSS is more important in higher-level math performance than the ANS.

If the DSS is not essential in representing meaning, why are simple symbolic effects, which are supposed to be DSS indexes, stronger predictors of more general math abilities than simple nonsymbolic effects, and why could be the DSS an essential component of mathematical performance? While there is currently no straightforward answer to these questions, we list a few possibilities. (1) There could be methodological ambiguities. For example, in correlational studies, the observed correlation is always attenuated by the reliability of the variables. If the reliability of the symbolic and nonsymbolic performance variables differ, then the size of the attenuation may be different, and without knowing the reliability (this is often not reported in the current practice) the correlations of symbolic and nonsymbolic regressors cannot be compared. (2) While several works illustrate that simple symbolic performance is a stronger predictor of higher-level math performance compared to the simple nonsymbolic performance, this does not mean that symbolic performance is a *strong* predictor, i.e., both symbolic and nonsymbolic performance may be weak (note again that various additional methodological factors influence how precise the measured absolute correlational indexes are). (3) While several DSS-related phenomena, such as the symbolic comparison distance and size effects rely on statistical features of the stimuli, this kind of statistical processing of the external world may be essential in more general problem solving as well, such as in detecting the frequencies of stimuli, for example, when one should decide whether to learn the multiplication table instead of performing an online calculation depending on the frequency of specific multiplications (Ashcraft, 1992). In other words, the symbolic numerical tasks may reflect abilities that are related to processing statistical properties of the environment. (4) While we argue that simple symbolic tasks previously attributed to the ANS should be attributed to the DSS, not necessarily all symbolic tasks or indexes measure the DSS performance. In other words, it is possible that simple symbolic tasks or some of the indexes are not valid DSS indexes, and the measured correlations or differences should not be interpreted as the function of individual differences in the DSS. (5) Some other numerical systems may influence or modulate the functioning of the DSS. A similar example is when acquiring the cardinality principle improves the ANS sensitivity (Shusterman et al., 2016). In this way, the performance of another symbolic numerical representation may change the observed DSS performance, which, in turn, correlates with more general math performance. Therefore, it is not the DSS that correlates with general math performance, but it is some other third factor that correlates with the DSS and that is not measured directly. Overall, there are many open questions about the methods, validity and interpretations of former results, and further research is needed to clarify and test these ambiguities.

General conclusions

We propose that all of the symbolic number processing phenomena that have been previously attributed to the ANS can also be explained with a model that assumes a network of discrete units. This alternative view can be formulated as the DSS model as outlined in this paper; this view is also compatible with several other numerical models that explain symbolic number processing phenomena with symbolic units and with their connections. More generally, in the hybrid ANS–DSS account, we propose that, while nonsymbolic phenomena may be explained by the ANS, simple symbolic effects can be accounted for by the DSS. We summarized a considerable amount of empirical evidence that the classic pure ANS account cannot handle, but whose results are in line with the hybrid ANS–DSS account.

We also argued that, while the classic ANS model supposed that the ANS is the main root of numerical meaning and general mathematical abilities, in the hybrid ANS–DSS account, neither the ANS nor the DSS has an extensive fundamental role in supporting the meaning of numbers or general math knowledge. This idea may be essential, because it indicates that additional key representations should be introduced to uncover the semantic processes behind mathematical understanding. This also opens up new possibilities to find effective ways to improve math education, to find more comprehensive diagnosis and development tools for developmental dyscalculia, and to improve everyday general mathematical understanding in society and times where and when rational thinking is much needed.

The literature on numerical thinking and mathematical understanding is enormous. We believe that it is impossible to take into account all relevant details. In the present work, while we intended to give a reasonably large review of the relevant phenomena, we intentionally omitted several issues, and, there are no doubt additional relevant results and models of which we are unaware. Even so, we hope that the extensive list of phenomena and the new perspective that we offer in this paper can generate more discussion on the appropriate description of numerical models, and the present framework can inspire new research questions and new empirical works.

Acknowledgments

We would like to thank Daniel Algom, Gergely Csibra, József Fiser, Ildikó Király, Zoltán Nádasdy, Dezső Németh, Marie-Pascale Noël, Manuela Piazza, Csaba Pléh, Bert Reynvoet, and Tom Verguts for discussing with us several details of the hybrid ANS–DSS framework and other alternative solutions. We would also like to thank Krisztián Kasos for his recommendations about a previous version of the manuscript. The work of Attila Krajcsi was supported by the National Research, Development and Innovation Fund (NKFI 132165), CELSA Research Fund (CELSA/19/011), ELTE Eötvös Loránd University, Faculty of Education and Psychology, and Central European University, Department of Cognitive Sciences.

References

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

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(3), 435–447. https://doi.org/10.1037/0096-1523.2.3.435

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *NeuroImage*, *87*, 311–322. https://doi.org/10.1016/j.neuroimage.2013.10.049

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2015). Visual Number Beats Abstract Numerical Magnitude: Format-dependent Representation of Arabic Digits and Dot Patterns in Human Parietal Cortex. *Journal of Cognitive Neuroscience*, *27*(7), 1376–1387. https://doi.org/10.1162/jocn_a_00787

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, *18*(6), 425–428. https://doi.org/10.1016/j.cub.2008.02.052

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*(12), 534–541. https://doi.org/16/j.tics.2010.09.007

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91. https://doi.org/10.1016/j.tics.2008.11.007

Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus*, 59–68.

Carey, S. (2009). *The Origin of Concepts* (1st ed.). Oxford University Press, USA.

Carey, S., & Barner, D. (2019). Ontogenetic Origins of Human Integer Representations. *Trends in Cognitive Sciences*, *23*(10), 823–835. https://doi.org/10.1016/j.tics.2019.07.004

Cohen, D. J. (2009). Integers do not automatically activate their quantity representation. *Psychonomic Bulletin & Review*, *16*(2), 332–336. https://doi.org/10.3758/PBR.16.2.332

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*(2), 132–147. https://doi.org/10.1016/j.pneurobio.2007.11.001

Cordes, S., & Brannon, E. M. (2008). Quantitative competencies in infancy. *Developmental Science*, *11*(6), 803–808. https://doi.org/10.1111/j.1467-7687.2008.00770.x

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557. https://doi.org/10.1073/pnas.1113195108

Damarla, S. R., & Just, M. A. (2013). Decoding the representation of numerical values from brain activation patterns. *Human Brain Mapping*, *34*(10), 2624–2634. https://doi.org/10.1002/hbm.22087

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Defever, E., Sasanguie, D., Vandewaetere, M., & Reynvoet, B. (2012). What can the same–different task tell us about the development of magnitude representations? *Acta Psychologica*, *140*(1), 35–42. https://doi.org/10.1016/j.actpsy.2012.02.005

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (1996). The Organization of Brain Activations in Number Comparison: Event-Related Potentials and the Additive-Factors Method. *Journal of Cognitive Neuroscience*, *8*(1), 47–68. https://doi.org/10.1162/jocn.1996.8.1.47

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dehaene, S. (2004). The neural bases of subliminal priming. In N. Kanwisher & J. Duncan (Eds.), *Functional Neuroimaging of visual cognition* (pp. 205–224).

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Vol. XXII* (pp. 527–574). Harvard University Press.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*(2), 314–326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and mental number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Dehaene, S., & Cohen, L. (1994). Dissociable mechanisms of subitizing and counting: Neuropsychological evidence from simultagnosic patients. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(5), 958–975.

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355–361.

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*(3), 626–641.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Dehaene, S., Spelke, E. S., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974.

Desoete, A., Ceulemans, A., Roeyers, H., & Huylebroeck, A. (2009). Subitizing or counting as possible screening variables for learning disabilities in mathematics education or learning? *Educational Research Review*, *4*(1), 55–66. https://doi.org/10.1016/j.edurev.2008.11.003

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265. https://doi.org/10.1016/j.cognition.2015.05.016

Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal representation in human intraparietal cortex. *Neuron*, *37*, 719–725.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Gallistel, C., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

García-Orza, J., Perea, M., Mallouh, R. A., & Carreiras, M. (2012). Physical similarity (and not quantity representation) drives perceptual comparison of numbers: Evidence from two Indian notations. *Psychonomic Bulletin & Review*, *19*(2), 294–300. https://doi.org/10.3758/s13423-011-0212-8

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Gebuis, T., & Reynvoet, B. (2012a). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*(4), 642–648. https://doi.org/10.1037/a0026218

Gebuis, T., & Reynvoet, B. (2012b). Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology*, *49*(11), 1649–1659. https://doi.org/10.1111/j.1469-8986.2012.01461.x

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 32–44. https://doi.org/10.1037/0096-1523.32.1.32

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465. https://doi.org/10.1037/a0012682

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Hauser, M. D., & Spelke, E. (2004). Evolutionary and Developmental Foundations of Human Knowledge: A Case Study of Mathematics. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences (3rd ed.)* (pp. 853–864). MIT Press.

Henik, A., Leibovich, T., Naparstek, S., Diesendruck, L., & Rubinsten, O. (2012). Quantities, amounts, and the numerical core system. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00186

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395. https://doi.org/10.3758/BF03202431

Hines, T. M. (1990). An odd effect: Lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*(1), 40–46. https://doi.org/10.3758/BF03202644

Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory & Cognition*, *10*(5), 487–495.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children’s mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Huber, S., Nuerk, H.-C., Willmes, K., & Moeller, K. (2016). A general model framework for multisymbol number comparison. *Psychological Review*, *123*(6), 667–695. https://doi.org/10.1037/rev0000040

Humphreys, M. S. (1975). The derivation of endpoint and distance effects in linear orderings from frequency information. *Journal of Verbal Learning and Verbal Behavior, 14*(5), 496–505. https://doi.org/10.1016/S0022-5371(75)80027-2

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*(4), 785–813. https://doi.org/10.3758/BF03196544

Ito, Y., & Hatta, T. (2004). Spatial structure of quantitative representation of numbers: Evidence from SNARC effect. *Memory and Cognition*, *32*(4), 662–673.

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: An entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1845–1864. https://doi.org/10.1037/a0016892

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*(4), 498–525.

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1882–1905. https://doi.org/10.1037/0096-1523.25.6.1882

Kojouharova, P., & Krajcsi, A. (2018). The Indo-Arabic distance effect originates in the response statistics of the task. *Psychological Research*. https://doi.org/10.1007/s00426-018-1052-1

Kojouharova, P., & Krajcsi, A. (2019). Two components of the Indo-Arabic numerical size effect. *Acta Psychologica*, *192*, 163–171. https://doi.org/10.1016/j.actpsy.2018.11.009

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., Fedele, M., & Reynvoet, B. (in preparation). *Approximate Number System cannot account for the number acquisition by subset-knowers*.

Krajcsi, A., & Igács, J. (2010). Processing negative numbers by transforming negatives to positive range and by sign shortcut. *European Journal of Cognitive Psychology*, *22*(7), 1021–1038. https://doi.org/10.1080/09541440903211113

Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02013

Krajcsi, A., Kojouharova, P., & Lengyel, G. (submitteda). *Preschoolers' development of understanding zero*.

Krajcsi, A., Kojouharova, P., & Lengyel, G. (submittedb). *Processing symbolic numbers: The example of distance and size effects*.

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analogue number system: Different symbolic and nonsymbolic numerical distance and size effects. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00124

Krajcsi, A., Lengyel, G., & Laczkó, Á. (2018). Interference between number magnitude and parity: Discrete representation in number processing. *Experimental Psychology*, *65*(2), 71–83. https://doi.org/10.1027/1618-3169/a000394

Krajcsi, A., & Szabó, E. (2012). The role of number notation: Sign-value notation number processing is easier than place-value. *Frontiers in Psychology, 3*(463). https://doi.org/10.3389/fpsyg.2012.00463

Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology, 60*(4), 227–234. https://doi.org/10.1027/1618-3169/a000191

Kucian, K., & von Aster, M. (2015). Developmental dyscalculia. *European Journal of Pediatrics, 174*(1), 1–13.

Landy, D. H., Jones, E. L., & Hummel, J. E. (2008). Why Spatial-Numeric Associations Aren't Evidence for a Mental Number Line. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 357–362). Cognitive Science Society.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438. https://doi.org/10.1016/j.cognition.2006.10.005

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *70*(1), 12–23. https://doi.org/10.1037/cep0000070

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From “sense of number” to “sense of magnitude”: The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *40*. https://doi.org/10.1017/S0140525X16000960

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics, Herzliya, Israel*, 357–362.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*(1), 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*(46), 18737–18742. https://doi.org/10.1073/pnas.1207212109

Lyons, I. M., Ansari, D., & Beilock, S. L. (2015). Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. *Human Brain Mapping*, *36*(2), 475–488. https://doi.org/10.1002/hbm.22641

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

McClelland, J. L., & Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*(4), 310–322. https://doi.org/10.1038/nrn1076

Molko, N., Cachia, A., Rivière, D., Mangin, J.-F., Bruandet, M., Le Bihan, D., Cohen, L., & Dehaene, S. (2003). Functional and Structural Alterations of the Intraparietal Sulcus in a

Developmental Dyscalculia of Genetic Origin. *Neuron*, *40*(4), 847–858. https://doi.org/10.1016/S0896-6273(03)00670-6

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*, 177–190. https://doi.org/10.1038/nrn1626

Nieder, A., Freedman, D. J., & Miller, E. K. (2002). Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, *297*, 1708–1711.

Noël, M.-P., & Rousselle, L. (2011). Developmental Changes in the Profiles of Dyscalculia: An Explanation Based on a Double Exact-and-Approximate Number Representation Model. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00165

Nuerk, H.-C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC - and the MARC - (linguistic markedness of response codes) effect. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(5), 835–863. https://doi.org/10.1080/02724980343000512

Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., Chillemi, G., & Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience Letters*, *438*(3), 308–311. https://doi.org/10.1016/j.neulet.2008.04.051

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 1–1. https://doi.org/10.1167/5.5.1

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Patro, K., Nuerk, H.-C., Cress, U., & Haman, M. (2014). How number-space relationships are assessed before formal schooling: A taxonomy proposal. *Developmental Psychology*, *5*, 419. https://doi.org/10.3389/fpsyg.2014.00419

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., & Dehaene, S. (2004). From number neurons to mental arithmetic: The cognitive neuroscience of number sense. In M. Gazzaniga (Ed.), *The cognitive neuroscience* (3rd ed., pp. 865–875). MIT Press.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Piazza, M., & Izard, V. (2009). How Humans Count: Numerosity and the Parietal Cortex. *The Neuroscientist*, *15*(3), 261–273. https://doi.org/10.1177/1073858409333073

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an amazonian indigene group. *Science*, *306*, 499–503.

Pinhas, M., & Tzelgov, J. (2012). Expanding on the Mental Number Line: Zero Is Perceived as the “Smallest.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1187–1205. https://doi.org/10.1037/a0027390

Poltrock, S. E., & Schwartz, D. R. (1984). Comparative judgments of multidigit numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 32–45. https://doi.org/10.1037/0278-7393.10.1.32

Price, G. R., Holloway, I., Räsänen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, *17*, R1042–R1043.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614.

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*, 191–201.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Reynvoet, B., & Sasanguie, D. (2016). The Symbol Grounding Problem Revisited: A Thorough Evaluation of the ANS Mapping Account and the Proposal of an Alternative Account Based on Symbol–Symbol Associations. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01581

Rinaldi, L., & Marelli, M. (2020). The use of number words in natural language obeys Weber's law. *Journal of Experimental Psychology: General*, *149*(7), 1215–1230. https://doi.org/10.1037/xge0000715

Rubinsten, O., & Henik, A. (2009). Developmental Dyscalculia: Heterogeneity might not mean different mechanisms. *Trends in Cognitive Sciences*, *13*(2), 92–99. https://doi.org/10.1016/j.tics.2008.11.002

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *The Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12372

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the Cardinal Principle Coincides with Improvement in Approximate Number System Acuity in Preschoolers. *PLOS ONE*, *11*(4), e0153072. https://doi.org/10.1371/journal.pone.0153072

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Spelke, E. S., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In *Language acquisition and conceptual development* (pp. 70–100). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Stoianov, I., & Zorzi, M. (2012). Emergence of a “visual number sense” in hierarchical generative models. *Nature Neuroscience*, *15*(2), 194–196. https://doi.org/10.1038/nn.2996

Tzelgov, J., Ganor-Stern, D., & Maymon-Schreiber, K. (2009). The representation of negative numbers: Exploring the effects of mode of processing and notation. *Quarterly Journal of Experimental Psychology*, *62*(3), 605–624. https://doi.org/10.1080/17470210802034751

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Verguts, T., & Fias, W. (2008). Symbolic and Nonsymbolic Pathways of Number Processing. *Philosophical Psychology*, *21*(4), 539–554. https://doi.org/10.1080/09515080802285545

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of “The Number Race”, an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*(20). http://www.behavioralandbrainfunctions.com/content/2/1/20

Wong, B., & Szűcs, D. (2013). Single-digit Arabic numbers do not automatically activate magnitude representations in adults or in children: Evidence from the symbolic same–different task. *Acta Psychologica*, *144*(3), 488–498. https://doi.org/10.1016/j.actpsy.2013.08.006

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. https://doi.org/10.1016/0010-0277(90)90003-3

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. https://doi.org/10.1016/0010-0285(92)90008-P

Wynn, K. (1996). Infants' Individuation and Enumeration of Actions. *Psychological Science*, *7*(3), 164–169. https://doi.org/10.1111/j.1467-9280.1996.tb00350.x

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Zorzi, M., & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the twenty first annual conference of the Cognitive Science Society* (pp. 778–783). Erlbaum. http://eprints.ucl.ac.uk/91024/

Chapter 3: The source of the symbolic numerical distance and size effects

The present chapter has been published:

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.01795

The source of the symbolic numerical distance and size effects

Attila Krajcsi[1], Gábor Lengyel[2], Petia Kojouharova[1, 3]

[1]Cognitive Psychology Department, Institute of Psychology, Eötvös Loránd University

[2]Department of Cognitive Science, Central European University

[3]Doctoral School of Psychology, Eötvös Loránd University

Abstract

Human number understanding is thought to rely on the analogue number system (ANS), working according to Weber's law. We propose an alternative account, suggesting that symbolic mathematical knowledge is based on a discrete semantic system (DSS), a representation that stores values in a semantic network, similar to the mental lexicon or to a conceptual network. Here, focusing on the phenomena of numerical distance and size effects in comparison tasks, first we discuss how a DSS model could explain these numerical effects. Second, we demonstrate that the DSS model can give quantitatively as appropriate a description of the effects as the ANS model. Finally, we show that symbolic numerical size effect is mainly influenced by the frequency of the symbols, and not by the ratios of their values. This last result suggests that numerical distance and size effects cannot be caused by the ANS, while the DSS model might be the alternative approach that can explain the frequency-based size effect.

Keywords: Numerical cognition; Numerical distance effect; Numerical size effect; Analogue number system; Discrete semantic system

An alternative to the analogue number system

According to the current models understanding numbers is supported by an evolutionary ancient representation shared by many species (Dehaene, Dehaene-Lambertz, & Cohen, 1998; Gallistel & Gelman, 2000; Hauser & Spelke, 2004), the analogue number system (ANS). One defining feature of the ANS is that it works similarly to some perceptual representations in which the ratio of the stimuli's intensity determines the performance (Weber's law) (Cantlon, Platt, & Brannon, 2009; Moyer & Landauer, 1967; Walsh, 2003). Two critical phenomena supporting the ratio based performance are the distance and the size effects: when two numbers are compared, the comparison is slower and more error prone when the distance between the two values is smaller (distance effect) or when the two numbers are larger (size effect), (Moyer & Landauer, 1967) (Figure 13 and 14). Thus, in the literature, the numerical distance and size effects are considered to be the sign of an analogue noisy numerical processing system working according to Weber's law. The distance and the size effects are observable both in non-symbolic and symbolic number processing, reflecting that the same type of system processes numerical information, independent of the number notations (Dehaene, 1992; Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003).

However, the distance and size effects in symbolic comparison can also be explained by a different representation. Quite intuitively, one might think that symbolic and abstract mathematical concepts, like numbers could be handled by a discrete semantic system (DSS), similar to conceptual networks or to the mental lexicon, i.e., representations that process symbolic and abstract concepts. In this DSS model, numbers are stored in a network of nodes, and the strength of their connections is proportional to the strength of their semantic relations. We propose that this DSS account could be responsible for symbolic number processing, whereas non-symbolic number processing is still supported by the ANS (see some additional details about the relation of the two models below). The main aim of the present study is to investigate the feasibility of the DSS model as a comprehensive explanation of the symbolic numerical effects, and to contrast it with the ANS model.

DSS explanation for the distance and size effects

How can a DSS explain the symbolic numerical distance and size effects? (1) Regarding the distance effect, the strength of the connections between the nodes can produce an effect which is proportional to their strength, and since in a network storing numbers the strength of the connections is proportional to the numerical values and numerical distance, this system could produce a numerical distance effect. In fact, a similar semantic distance effect was shown in a picture naming task (Vigliocco, Vinson, Damian, & Levelt, 2002): Naming time slowed down when the picture of the previous trial was semantically related to the present picture, and a small semantic distance between the previous and the actual word caused stronger effect than a large semantic distance, similar to the numerical distance effect.[1] This semantic distance effect cannot be the result of a continuous representation similar to the ANS, because the stimuli were categorical (e.g., finger, car, smile, etc.)[2]. Thus, a discrete representation potentially can produce a numerical distance effect. Several mechanisms can be imagined how a numerical distance effect is generated. One can imagine that the semantic distance information, that can be revealed in a semantic priming, could generate a distance effect. Alternatively, it is possible that the strength of the association between the numbers and the large-small categories create the numerical distance effect (Verguts & Fias, 2004; Verguts, Fias, & Stevens, 2005). Here, we do not want to specify the exact mechanism behind the numerical distance effect, but only propose that several possible mechanisms are already available in the literature. (2) Turning to the size effect, this effect also could be generated by a DSS. It is known that smaller numbers are more frequent than larger numbers, and the frequency of a number is proportional to the power of its value (Dehaene & Mehler, 1992). Since the numbers observed more frequently could be processed faster, the size effect could result from this frequency

1 Comparison distance effect (e.g., which of two numbers is larger) and priming distance effect (whether previous stimulus influences the actual stimulus processing based on the distance of the two stimuli) are known to be two different mechanisms (Reynvoet, De Smedt, & Van den Bussche, 2009; Verguts, Fias, & Stevens, 2005). While we want to find a DSS explanation for the comparison distance effect, the cited semantic distance effect is more similar to a priming distance effect. Importantly, we are not stating that these two effects are the same, but we suggest that a distance-based effect is possible in a DSS, independent of the exact mechanism behind that effect.

2 A similar proposal is that the numerical distance effect might emerge from the order property of numbers, and a distance effect can be observed not only in numbers, but also in non-numerical orders, e.g., days or letters (Potts, 1972; Verguts & Van Opstal, 2014). However, (a) it might be possible that in those examples the non-numerical orders are transformed to the numerical representation, which is not possible for the categorical words in the cited picture naming task (Vigliocco, Vinson, Damian, & Levelt, 2002), and (b) the DSS model has less strict constrains, i.e., no order structure is presupposed, but a more general series of associations is sufficient to explain the distance effect.

pattern.[3] Thus, the DSS model can also explain the appearance of distance and the size effects (Figure 13).

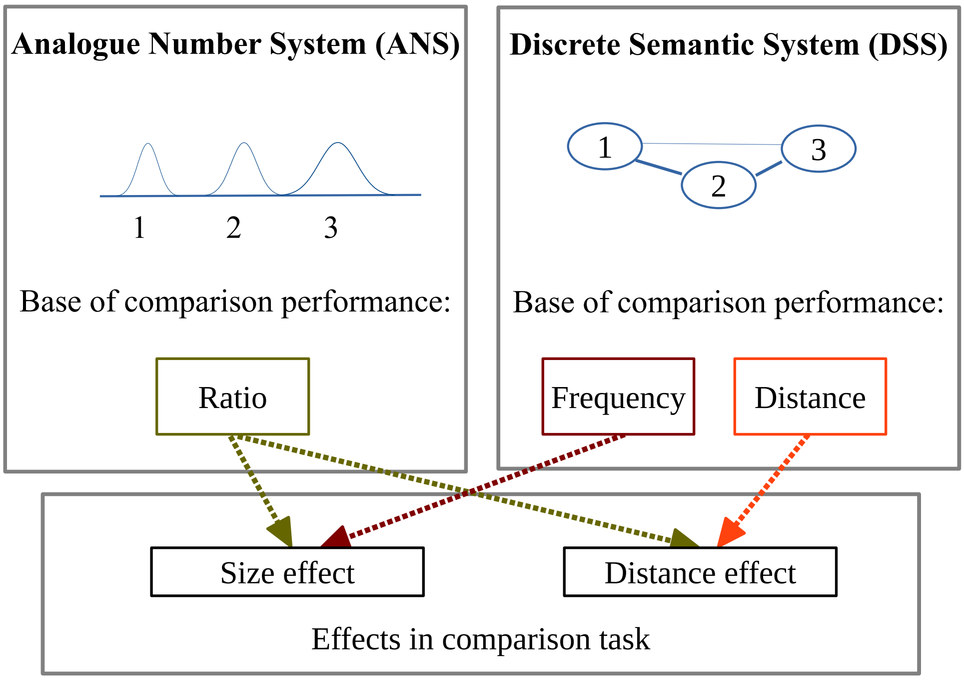


Figure 13. The sources of the distance and size effects according to the two models.

DSS explanation for other numerical effects

Whereas in the present work we focus on the DSS explanation of the distance and size effects, the DSS explanation can be readily extended to other effects, too, and it can be a comprehensive model of symbolic number processing. The following details can demonstrate that despite its radical difference from the ANS model, DSS might be a viable option to explain symbolic numerical phenomena. Many of these explanations have already been proposed in the literature, although these explanations usually focused on single specific phenomena, and they did not offer a comprehensive model.

Several interference effects can be explained in the DSS framework. For example, the SNARC effect (interference between numerical value and response location in a task) was originally interpreted as the interference of the ANS's spatial property and the response locations (Dehaene, Bossini, & Giraux, 1993), however, it is also possible that the effect is the result of the interference of the left-right and large-small nodes in a semantic network similar to the DSS (Krajcsi, Lengyel, & Laczkó, submitted; Leth-Steensen, Lucas, & Petrusic, 2011; Patro, Nuerk, Cress, & Haman, 2014; Proctor & Cho, 2006). Similarly, while the size congruency effect (Stroop-like interference between the numerical value and the physical size of symbols; Henik & Tzelgov, 1982) can be thought of as an interference between the ANS and a representationally similar analogue size

3 Frequency is essential in other numerical tasks to produce size effect (Zbrodoff & Logan, 2005), and the role of frequency in size effect was also proposed in other alternative models of number comparison (Verguts & Van Opstal, 2014).

representation, it can also be thought of as an interference between the many-few and the physically large-physically small nodes.

While there are many empirical and theoretical works in the literature that support the ANS model, in fact there are only a handful effects that are cited to support the ANS model, and we propose that most of these effects (in fact to our knowledge all of them at the moment) can also be explained by the DSS. While mostly it would not be too difficult to find DSS explanations for different phenomena, in the present work we only focus on the numerical distance and size effects in comparison tasks.

Different representations for symbolic and non-symbolic numbers

As it was mentioned above, the DSS model can only account for symbolic number processing. Clearly, there are cases when the DSS cannot handle numerical information, for example, when the symbolic mental tools are not available, like in the case of infants (Feigenson, Dehaene, & Spelke, 2004), animals (Hauser & Spelke, 2004), or adults living in a culture without number words (Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004), therefore, the ANS seems to be a sensible model to explain these non-symbolic phenomena. It also seems reasonable that because of their representational structure, the two systems could be specialized for different forms of numbers: The DSS could be responsible for the precise and symbolic numbers, while the ANS could process the approximate non-symbolic stimuli.

This idea of different representations for symbolic and non-symbolic numbers is supported by the increasing number of findings in the literature, suggesting that symbolic and non-symbolic number processing are supported by different representations. For example, it has been shown that performance of the symbolic and non-symbolic number comparison tasks do not correlate in children (Holloway & Ansari, 2009; Sasanguie, Defever, Maertens, & Reynvoet, 2014), and in an fMRI study the size of the symbolic and non-symbolic number activations did not correlate (Lyons, Ansari, & Beilock, 2015). As another example, whereas former studies found common brain areas activated by both symbolic and non-symbolic stimuli (Eger et al., 2003; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004), later works with more sensitive methods found only notation-dependent activations (Bulthé, De Smedt, & Op de Beeck, 2014, 2015; Damarla & Just, 2013). According to an extensive meta-analysis, although it was repeatedly found that simple number comparison task (the supposed sensitivity of the ANS) correlates with mathematical achievement, it seems that nonsymbolic comparison correlates much less with math achievement, than symbolic comparison (Schneider et al., 2016). In another example, Noël and Rousselle (2011) found that whereas older than 9- or 10-year-old children with developmental dyscalculia (DD) perform worse in both symbolic and non-symbolic tasks than the typically developing children, younger children with DD perform worse than control children only in the symbolic tasks, but not in the non-symbolic tasks. The authors concluded that the deficit in DD can be explained in the terms of two different representations: The deficit is more strongly related to the symbolic number processing, and the impaired non-symbolic performance is only the consequence of the symbolic processing problems. See a more extensive review of similar findings in Leibovich & Ansari (2016). All of these findings are in line with the present proposal, suggesting that symbolic and non-symbolic numbers are processed by different systems.

Related models for symbolic number processing

There are former models in the literature that are potential alternatives to the ANS model, and some of those models can be fitted into a DSS framework, or they could be considered as implementations of more specific aspects of the DSS account.

Tom Verguts and his colleagues proposed a connectionist model describing several phenomena of number processing and more generally several phenomena of ordinal information processing (Verguts et al., 2005; Verguts & Van Opstal, 2014). According to their simulations and experiments, this model offers a superior description of number naming, parity judgment and number comparison than the ANS model, and their model can also explain non-numerical order processing phenomena. Their model includes a hidden layer representing the values of the numbers in a place-code with a fixed width of noise. This means that the nodes of the hidden layer represent numbers on a linear scale, and a number most strongly activates the node mainly representing that number, but additional activation also can be found in the neighboring nodes. The distance these additional activations can reach to does not depend on the source number, i.e., the noise has a fixed width. Although the authors suggest that this model implements an analogue representation, it contradicts the ANS model, because on a linear inner scale the size of the noise is not proportional to the size of the number, and relatedly it could not generate ratio-based performance. In line with this representational issue, the model in itself cannot produce a size effect, and an uneven frequency of numbers should be introduced to generate the numerical size effect (Verguts & Fias, 2004; Verguts et al., 2005), questioning whether this model can be seen as an ANS-like model. However, we propose here that the model can be interpreted as a discrete symbolic representation: Activation in the neighboring nodes is not the noise of that representation but it is a spreading activation in the hidden layer. With this alternative interpretation the model can be seen as a specific implementation of the discrete symbolic system when stimuli are arranged as an ordered list. Note that in their model the comparison distance effect is not explained by the spreading activation, but by the connection weights between the value nodes and the response nodes (Verguts et al., 2005; Verguts & Van Opstal, 2014). This model as a potential DSS implementation can give a more precise description for a whole range of phenomena, the ANS model could not account for, thus, strengthening the discrete semantic system explanation of symbolic number processing.

Tracking a different line, Joseph Tzelgov and his colleagues investigated automatic processing of numbers with the size congruency effect (interference between the physical size and numerical value properties of the stimuli) (Henik & Tzelgov, 1982). Based on their results they suggested that some basic elements (primitives) are stored in the long term memory, e.g., integers from 1 to 9 and the number 0 (Pinhas & Tzelgov, 2012), while other numbers are not stored as basic elements, e.g., negative numbers and ratios (Kallai & Tzelgov, 2009; Tzelgov, Ganor-Stern, & Maymon-Schreiber, 2009). The basic elements or primitives can be considered as the nodes of the DSS: These basic elements could be the values that are stored in the nodes of the network, while other numbers are the combination of the primitives, somewhat similar to the relation of words and sentences. Also, the size congruency effect can be used as a method to find whether a number is stored as a unit in the DSS.

Possible quantitative descriptions of symbolic comparison performance in the DSS model

While the DSS model can explain why the numerical distance and size effects appear in a comparison task, the ANS model not only suggests that there should be numerical distance and size effects, but it offers a quantitative description for the performance. For example, Moyer and Landauer (1967) proposed that the reaction time of a comparison task is proportional to the following function: *K × log (large_number / (large_number - small_number))*. (See Dehaene, 2007 for a more detailed description of the ANS predictions for behavioral numerical decisions.)

One of the next challenges for the DSS model is to find a quantitative description similar to the ANS model. As in the ANS model where the details of the model were borrowed from psychophysics models, we borrow the details of the DSS model from psycholinguistics and semantic network models. Unfortunately, whereas in many cases the psychophysics models offer quantitative descriptions of the performance (Dehaene, 2007; Kingdom & Prins, 2010), the bases of the DSS model do not have consensual quantitative descriptions. Additionally, our description does not build upon a detailed working model with specific mechanisms (e.g., as it was mentioned, there could be different candidates that could generate the distance effect), but a functional description of these potential effects are given here. Thus, our quantitative proposal is unavoidably speculative, although there are some constrains we can build upon. First, one term of this quantitative description should depend on the distance between the two values. Second, another term should depend on the frequencies of the values, where the frequency of the number is the power of that number (Dehaene & Mehler, 1992). Current theoretical considerations do not specify what distance and size functions should be used, how the frequency of the two numbers should be combined, and how exactly the two terms create performance, thus these details are unavoidably speculative at the moment, and future work can refine the versions offered here. However, based on these few starting points, a number of alternative versions of the DSS model can be created, and many of them display a qualitatively similar pattern of number comparison performance. One simple example is displayed on Figure 14, where, as the mathematically simplest version, the distance effect is a linear function, the frequencies of the numbers are summed up, and the distance and size components are added up. This DSS-motivated function creates a qualitatively very similar pattern to the function of the ANS model: Looking at the patterns, the two models are rather similar, also reflected in the high correlation between the two models (r = -0.89). Thus, one can create a hypothetical quantitative description based on the DSS account, that seemingly can explain the comparison performance in a similar way as the ANS model.[4]

4 After creating additional versions of the DSS quantitative prediction with considering the constrains described here, we found qualitatively similar patterns. See another example in the Methods section of Experiment 1.

A

Analogue Number System (ANS) model

$$RT = a \times \log(\text{large}/\text{distance}) + b$$

Discrete Semantic System (DSS) model

$$RT = a_1 \times \text{distance} + a_2 \times (x_1^{-1} + x_2^{-1}) + b$$

B

Number 1

Number 2	1	2	3	4	5	6	7	8	9
1		0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1
2	0.3		0.5	0.3	0.2	0.2	0.1	0.1	0.1
3	0.2	0.5		0.6	0.4	0.3	0.2	0.2	0.2
4	0.1	0.3	0.6		0.7	0.5	0.4	0.3	0.3
5	0.1	0.2	0.4	0.7		0.8	0.5	0.4	0.4
6	0.1	0.2	0.3	0.5	0.8		0.8	0.6	0.5
7	0.1	0.1	0.2	0.4	0.5	0.8		0.9	0.7
8	0.1	0.1	0.2	0.3	0.4	0.6	0.9		1.0
9	0.1	0.1	0.2	0.3	0.4	0.5	0.7	1.0	

Number 1

Number 2	1	2	3	4	5	6	7	8	9
1		1.9	2.1	2.5	2.8	3.2	3.5	3.9	4.3
2	1.9		1.2	1.6	1.9	2.3	2.6	3.0	3.4
3	2.1	1.2		1.0	1.3	1.7	2.1	2.5	2.8
4	2.5	1.6	1.0		0.9	1.2	1.6	2.0	2.4
5	2.8	1.9	1.3	0.9		0.8	1.1	1.5	1.9
6	3.2	2.3	1.7	1.2	0.8		0.7	1.1	1.5
7	3.5	2.6	2.1	1.6	1.1	0.7		0.7	1.1
8	3.9	3.0	2.5	2.0	1.5	1.1	0.7		0.6
9	4.3	3.4	2.8	2.4	1.9	1.5	1.1	0.6	

Figure 14. (**A**) Reaction time (RT) function for the ANS model (based on Crossman, 1955; Moyer & Landauer, 1967) (left) and a hypothetical RT function for the DSS model where the reaction time is proportional to a combination of the specific forms of the distance and the frequencies of the numbers (right). Notations: large: larger number; distance: distance between the two numbers; x_1 and x_2: the two numbers; a, a_1, a_2 and b are free parameters. (**B**) The prediction of the models on a full stimulus space in a number comparison task of numbers between 1 and 9. Number 1 and 2 are the two values to be compared. Green denotes fast responses, red denotes slow responses (note that numerically the ANS function increases, and the DSS function decreases towards the high ratio, but the direction is irrelevant in the linear fit below). The distance effect can be seen as the gradual change when getting farther from the top-left bottom-right diagonal, and the size effect is seen as the gradual change from top-left to bottom-right. In the figures the parameters a and a_2 are set to 1, a_1 is 0.4, and parameter b is set to 0.

In the first section, so far we have introduced the DSS model, an alternative to the ANS explanation of number processing, where the basic building blocks of the representation are nodes with appropriate connections. We have reasoned that the DSS framework can be a comprehensive explanation of symbolic number processing. While focusing on the comparison distance and size effects, we have demonstrated that the DSS model is capable of giving as appropriate a description of the comparison performance as the ANS model. In the following parts we turn to empirical tests. First, we investigate which model describes better an Indo-Arabic comparison task. Then, we investigate a very specific aspect of number comparison where the two models have clearly different predictions: Whether the size effect depends on the frequency of the numbers (predicted by the DSS model) or on the ratio of the numbers (predicted by the ANS model).

Experiment 1 – Goodness of the two quantitative description of the models in Indo-Arabic comparison

After creating a quantitative description for the DSS model, we can contrast the two models, testing which model (Figure 14) fits better the empirical data in an Indo-Arabic number comparison task.

Although the two models strongly correlate, and the differences between them are subtle, still, there are differences between them, and it is possible that those differences are detectable in a simple comparison task, supposing that the noise is relatively low.

Methods

Participants. Twenty university students participated in the study. Pilot studies with Indo-Arabic and new symbols (see also the second experiment) aiming to refine the applied paradigms revealed that the main effects to be observed can be detected reliably with a sample size of around 20. After excluding two participants because of a higher than 5% error rate, the sample included 18 participants (15 females, mean age 21.5 years, standard deviation 2.8 years).

Stimuli and procedure. The participants compared Indo-Arabic number pairs. In a trial two numbers between 1 and 9 were shown until response, and the participants chose the larger one. All possible number pairs including numbers between 1 and 9 were shown 10 times, excluding ties, resulting in 720 trials. Presentation of the stimuli and measurement of the responses were managed by the PsychoPy software (Peirce, 2007).

Analysis methods. In the analysis, we contrasted the two models with analyzing the reaction times, the error rates, and the diffusion analysis drift rates. (1) Reaction time analysis was used, because response latency may be a more sensitive measurement than the error rate, and the results are comparable with many former results, including the seminal Moyer & Landauer (1967) paper. However, there is no strong consensus which function could describe the ANS model (see the applied version below). (2) Error rate analysis was chosen, because the function describing error rate performance is well established (Dehaene, 2007; Kingdom & Prins, 2010), even if the measurement is not as sensitive as the reaction time data. (3) Finally, drift rate was applied, because diffusion analysis is thought to be more sensitive than the error rate or the reaction time, although its parameter recover methods could be debated. In the recent decades, the diffusion model and related models became increasingly popular to describe simple decision processes (Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). In the diffusion model, decision is based on a gradual accumulation of evidence offered by perceptual and other systems. Decision is made when an appropriate amount of evidence is accumulated. Reaction time and error rates partly depend on the quality of the information (termed the drift rate) upon which the evidence is built. Importantly for our analysis, observed reaction time and error rate parameters can be used to recover the drift rates (Ratcliff & Tuerlinckx, 2002; Wagenmakers, van der Maas, & Grasman, 2007). Drift rates can be more informative than the error rate or reaction time in themselves, because drift rates reveal the sensitivity of the background mechanisms more directly (Wagenmakers et al., 2007).

Because different versions of the ANS models and the DSS models can be proposed, multiple versions of the models were tested, when it was necessary. For the ANS model the following functions were used in the analysis. (1) Regarding the reaction time analysis, although there are several considerations how to describe the reaction time function of continuous perceptual comparisons (Crossman, 1955; Dehaene, 2007; Welford, 1960), it is not straightforward which version should be applied to describe the ANS model (Kingdom & Prins, 2010). First, we used the version used by Moyer & Landauer (1967), displayed in Figure 14. Second, we applied the $RT \propto 1 / (\log(large/small))$ function suggested by Crossman (1955), which function he finds to be more superior compared to the previous function. (2) For the error rate analysis we used the ANS model described in Dehaene (2007, equation 10), which supposes a linear scaling in the ANS,

$$p_{correct}(n_1, n_2) = \int_0^{+\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-(r-1)}{w\sqrt{1+r^2}}\right)^2}}{\sqrt{2\pi}\, w\sqrt{1+r^2}}\, dx$$

where n_1 and n_2 are the two numbers to be compared, r is the ratio of the larger and the smaller number, and w is the Weber ratio. (3) Regarding the drift rates, in the ANS model the stored values to be compared can be conceived as two random Gaussian variables, and the difficulty of the comparison might depend on the overlap of the two random variables: Larger overlap leads to worse performance (see the detailed mathematical description in Dehaene, 2007). It is supposed that in a comparison task the drift rate depends purely on the overlap of the two random variables (Dehaene, 2007; Palmer, Huk, & Shadlen, 2005). According to the current theories, *drift_rate = k × task_difficulty*, (Dehaene, 2007; Palmer et al., 2005), or it could also include a power term as a generalization, *drift_rate = k × task_difficulty*$^{\beta}$, although the exponent is often close to 1, thus the first, proportional model approximates the second, power model. Task difficulty is measured as stimulus strength, which is calculated with the *distance/large_number* function as suggested by Palmer et al. (2005) for psychophysics comparison. Because in an analogue representation as the task becomes more difficult (i.e., the two stimuli become indistinguishable) the drift rate tends to zero, in the linear fit this means that the intercept is forced to be zero. To summarize, the *drift_rate = k × distance/large_number* function was used in the drift rate analysis fit for the ANS model.

For the DSS model, two versions were used in the analysis. First, the simple linear version was applied, as described in Figure 14. Additionally, a logarithmic version of the DSS model was also used, in which the logarithm of the two terms are used, i.e., RT ∝ log(distance) + log($x_1^{-1} \times x_2^{-1}$). This logarithmic version seems reasonable, because strictly speaking the distance effect cannot be linear, since that would result in negative reaction time or error performance for sufficiently large distances (even if the linear version could be an appropriate approximation). Additionally, the logarithmic distance effect is partly confirmed by the second experiment and by the inspection of the residuals (results not presented here).

Detecting the distance and size effects. The present analysis is not relevant in contrasting the ANS and DSS models, but in the second and third experiments the existence of the numerical distance and size effects was tested, and the same analysis was run in the present experiment, to be able to use these results as a point of reference. The slopes of the specific effects were tested (1) with multiple linear regressions, and (2) with simple linear regressions.

1. Methods for multiple linear regression. Average error rates and median reaction times of the correct responses were calculated for each number pair for each participant. Error rates and reaction times were fitted with two regressors for all participants: (a) distance effect (the absolute difference of the two values), (b) size effect (the sum of the two values). See the values of the regressors for the whole stimulus space on Figure 15. (The end effect regressor is used only in the second and third experiments.) This analysis gives a more stable result compared to the more commonly applied simple linear regression analysis (see below). The weights of the regressors were calculated for each participant in both error rates and reaction times, and all regressors' values were tested against zero.

Distance effect

	1	2	3	4	5	6	7	8	9
1	0	1	2	3	4	5	6	7	8
2	1	0	1	2	3	4	5	6	7
3	2	1	0	1	2	3	4	5	6
4	3	2	1	0	1	2	3	4	5
5	4	3	2	1	0	1	2	3	4
6	5	4	3	2	1	0	1	2	3
7	6	5	4	3	2	1	0	1	2
8	7	6	5	4	3	2	1	0	1
9	8	7	6	5	4	3	2	1	0

Size effect

	1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9	10
2	3	4	5	6	7	8	9	10	11
3	4	5	6	7	8	9	10	11	12
4	5	6	7	8	9	10	11	12	13
5	6	7	8	9	10	11	12	13	14
6	7	8	9	10	11	12	13	14	15
7	8	9	10	11	12	13	14	15	16
8	9	10	11	12	13	14	15	16	17
9	10	11	12	13	14	15	16	17	18

End effect

	1	2	3	4	5	6	7	8	9
1	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1
2	0.5	0	0	0	0	0	0	0.5	1
3	0.5	0	0	0	0	0	0	0.5	1
4	0.5	0	0	0	0	0	0	0.5	1
5	0.5	0	0	0	0	0	0	0.5	1
6	0.5	0	0	0	0	0	0	0.5	1
7	0.5	0	0	0	0	0	0	0.5	1
8	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	1
9	1	1	1	1	1	1	1	1	1

Figure 15. Values of the three regressors applied in the multiple linear regression in the whole stimulus space.

2. Methods for simple linear regression. To test our data with a more commonly applied simple linear regression, all multiple linear regression analyses were retested. For the distance effect the trials were grouped according to distance (absolute difference between the two numbers) for all participants. For the size effect the trials were grouped according to the sum of the two numbers, excluding trials with distance larger than 3. The latter was necessary, because otherwise the specific shape of the stimulus space and the distance effect might cause an artifact size effect: Cells from the middle part of the size range include more large-distance cells than cells from the end part of the size range do. Linear slope was fitted both on the error rates and on the reaction times for both the distance and size effects for all participants, then the slopes were tested against zero. Because the simple linear regression analysis gave the very same pattern as the multiple linear regression for all experiments of the present work, the results of this analysis are not presented here.

Results and discussion

Fitting the functions of the ANS and the DSS models to the reaction times. For the reaction time analysis median reaction time of the correct responses for each number pair and for each participant was calculated. The mean of the participants data for all number pairs (Figure 16) were fit linearly with the least square method. Four models were fit to the group mean: The Moyer & Landauer version of the ANS function, the Crossman version of the ANS function, the linear DSS function, and the logarithm DSS function (see methods for their descriptions).

For the Moyer & Landauer version the data showed a quite appropriate fit, with $R^2 = 0.884$, $AIC = 613.8$, while the Crossman version of the ANS function fit was somewhat worse, although similar, with $R^2 = 0.769$ and $AIC = 663.5$. Regarding the DSS models, the fit for the linear version was $R^2 = 0.808$, $AIC = 652.4$, and the fit for the logarithm version was $R^2 = 0.893$, and $AIC = 610.3$.

Overall, fitting the functions of the four versions of the two models resulted in similar AICs within the same range, therefore no clear preference for any model can be pronounced. It seems that either the appropriate function is not precise enough to have a higher fit (which could be true for either the ANS or the DSS model), and/or with the current noise of the data the subtle differences between the models cannot be investigated. Thus, reaction time analysis with the current functions and the available signal-to-noise ratio could not be decisive in contrasting the ANS and DSS model.

Error rate

Number on the left \ Number on the right	1	2	3	4	5	6	7	8	9
1		1%	0%	0%	1%	1%	0%	0%	0%
2	2%		2%	3%	0%	1%	3%	0%	1%
3	2%	1%		4%	2%	1%	0%	0%	1%
4	1%	1%	5%		7%	1%	4%	2%	1%
5	1%	1%	1%	4%		7%	7%	2%	2%
6	1%	1%	0%	3%	2%		9%	6%	2%
7	0%	2%	2%	4%	7%	12%		11%	6%
8	1%	0%	0%	0%	2%	6%	7%		6%
9	0%	0%	1%	0%	4%	3%	6%	11%	

Reaction time

Number on the left \ Number on the right	1	2	3	4	5	6	7	8	9
1		494	487	481	462	449	462	444	448
2	502		556	538	503	489	491	474	469
3	480	513		591	537	517	527	487	483
4	474	523	560		576	515	555	498	505
5	459	487	528	561		577	589	524	521
6	457	467	497	513	588		633	560	554
7	457	493	509	530	576	624		605	559
8	447	473	472	488	518	600	580		622
9	457	456	483	482	514	561	541	604	

Drift rate

Number on the left \ Number on the right	1	2	3	4	5	6	7	8	9
1		0.33	0.29	0.34	0.35	0.38	0.37	0.4	0.41
2	0.32		0.27	0.29	0.34	0.34	0.36	0.37	0.37
3	0.29	0.32		0.27	0.29	0.3	0.34	0.34	0.36
4	0.32	0.3	0.27		0.22	0.3	0.31	0.29	0.31
5	0.36	0.32	0.31	0.28		0.24	0.26	0.31	0.32
6	0.36	0.36	0.32	0.25	0.24		0.21	0.24	0.27
7	0.32	0.31	0.31	0.27	0.23	0.2		0.21	0.26
8	0.39	0.34	0.36	0.34	0.29	0.23	0.23		0.24
9	0.36	0.37	0.35	0.31	0.29	0.25	0.25	0.22	

Figure 16. Error rates (left), response times in ms (middle) and drift rates (right) in the Indo-Arabic digits number comparison for the whole stimulus space. Green denotes fast and error-free responses, red denotes slow and erroneous responses. Results show distance and size effects.

Fitting the functions of the models to the error rates. For the error rate analysis, the mean error rate for each number pair and for each participant was calculated, then the average of the participants was computed (Figure 16). To test the ANS model, first, we looked for the Weber ratio that gives the same mean error rate for the stimulus space used here (all possible number pairs for numbers between 1 and 9, ties excluded) as it was measured in our data (2.5%). The found 0.11 Weber ratio was used to generate the predictions of the ANS model for all cells of the stimulus space (see methods for the function), and the model was linearly fit to the error rate data with the least square method. The goodness of fit was R^2 = 0.625, *AIC* = -371. In testing the DSS model, the goodness of fit for the linear version was R^2 = 0.505, *AIC* = -341, and the logarithmic DSS model gave a goodness of fit of R^2 = 0.667, *AIC* = -377.

Like in the case of the reaction time, the goodness of fit of the ANS and the DSS models are indistinguishable in the error rates data. This again shows that with the signal-to-noise ratio of the present data, the two models are indistinguishable, or the DSS model is not precise enough to show a higher fit.

Fitting the functions of the models to the drift rates. To recover the drift rates for all number pairs in the two notations, the EZ diffusion model was applied, which can be used when the number of trials per cells is relatively small (Wagenmakers et al., 2007). For edge correction we used the half trial solution (see the exact details about edge correction in Wagenmakers et al., 2007). The scaling within-trials variability of drift rate was set to 0.1 in line with the tradition of the diffusion analysis literature. Drift rates for each number pair and participant were calculated in both notations. The mean drift rates of the participants for the full stimulus space are displayed in Figure 16.

According to the goodness of fit of the models, the ANS model is worse (*AIC* = -140.1) than the DSS model (*AIC* = -332.4 and *AIC* = -348.1 for the linear and logarithmic DSS model versions, respectively). (Because in a linear fit with zero intercept, the R^2 is much higher than in a linear fit with non-zero intercept (as a consequence of some of the mathematical properties of R^2), and because the ANS model uses 0 intercept, but the DSS model does not, the R^2 values are not reported here.)

Looking at the drift rates of the comparison task (Figure 16) might reveal why the ANS model is worse than the DSS model: While the ANS model predicts that the drift rate tends to zero as the stimuli become indistinguishable (e.g., 8 vs. 9), the recovered drift rates are in fact much larger, tending to the 0.2 values. This problem is analogous to a conceptual problem: How is it possible that an imprecise representation solves a precise comparison task? In other words, if the Weber fraction of the ANS is around 0.11, how is it possible that small ratio number pairs, e.g., 8 vs. 9, can still be differentiated with relatively high precision.

Thus, in the diffusion model analysis the DSS model seems to offer a better prediction than the ANS model, however, it is important to note that (a) the EZ diffusion model analysis and more generally any diffusion models have some constrains (Wagenmakers et al., 2007), and consequently, it is possible that in this case the recovered parameters are not entirely reliable, and (b) task difficulty can be defined in different ways (Dehaene, 2007; Palmer et al., 2005), and it might be debated which definition is appropriate. Thus, while the present diffusion model analysis reveals the advantage of the DSS model over the ANS model, the uncertainties of the methods might question how reliable these results are. (The methods and the models are investigated in more details in Krajcsi, Lengyel, & Kojouharova, submitted.)

Presence of the distance and the size effects. According to the multiple linear regression analysis, both the distance and the size effects were present both in the error rates and in the reaction times, 95% CI for the slope was [-1.16%, -0.65%], $t(17) = -7.42$, $p < 0.001$ for the distance effect in error rates, and CI of [-23.6 ms, -15.5 ms], $t(17) = -10.1$, $p < 0.001$ in reaction times, CI with [0.3%, 0.59%], $t(17) = 6.57$, $p < 0.001$ for the size effect in error rates, and CI with [4.8 ms, 9.1 ms], $t(17) = 6.78$, $p < 0.001$ in reaction times.

Summary. First, we found that reaction time and error rate patterns in Indo-Arabic number comparison (Figure 16) could not be decisive in contrasting the ANS and the DSS models. Even if the two models correlate, the correlation is not perfect, and there was a chance that the present test could have decided. Still, with the present models and/or signal-to-noise ratio, the test was not decisive. On the positive side, this means that the DSS model is a viable alternative to the ANS model, because the goodness of fit of the DSS model is in the same range as the goodness of fit of the ANS model. Second, we found that in a diffusion model analysis the drift rate pattern is more in line with the DSS model than with the ANS model, although the uncertainties about the method may question the reliability of these results. Overall, while the performance in the Indo-Arabic comparison task suggests that the DSS model is a viable model, this paradigm could not decide firmly which model is preferred. Thus, in the next experiment a new approach is utilized in which we investigate the role of the frequency in the size effect.

Experiment 2 – Role of the frequency in the size effect

In a different approach, we tested whether the distance and the size effects are strongly related as suggested by the ratio-based ANS model, or whether the two effects can dissociate. In the present experiment we investigated whether size effects can dissociate from distance effect if the frequency of the symbols is manipulated. (See another type of test for the dissociation of the two effects in Krajcsi, 2016) To manipulate the frequency of the symbols, it might be more appropriate to use new symbols, instead of the well-known Indo-Arabic symbols, because the frequency of the already known symbols might be well established and learned.

Thus, to investigate the role of the frequency in the size effect, participants learned new number symbols in a simple number comparison task, and the frequency of the symbols was manipulated in the experiment. According to the DSS model, the size effect could be changed as a function of the symbol frequencies (Figure 13), if the reaction time depends on the frequency of the symbol, and not the frequency of the concept. For example, if the distribution of the frequencies is uniform, then according to the DSS model, the size effect should vanish. In contrast, according to the ANS model, even with uniform frequency distribution the size effect should be visible, because the size effect is rooted in the ratio of the two values, independent of the frequency (Figure 17). It is important to stress that although according to the ANS model it might be possible that the frequency of the symbols have an effect on the performance, the effect should be relatively weak: Although in the ANS model the role of the frequency is not discussed, it states that the largest part of the performance variance should be explained by the ratio (Dehaene, 2007; Moyer & Landauer, 1967), which means that any other factors could have only a minor effect on the performance.

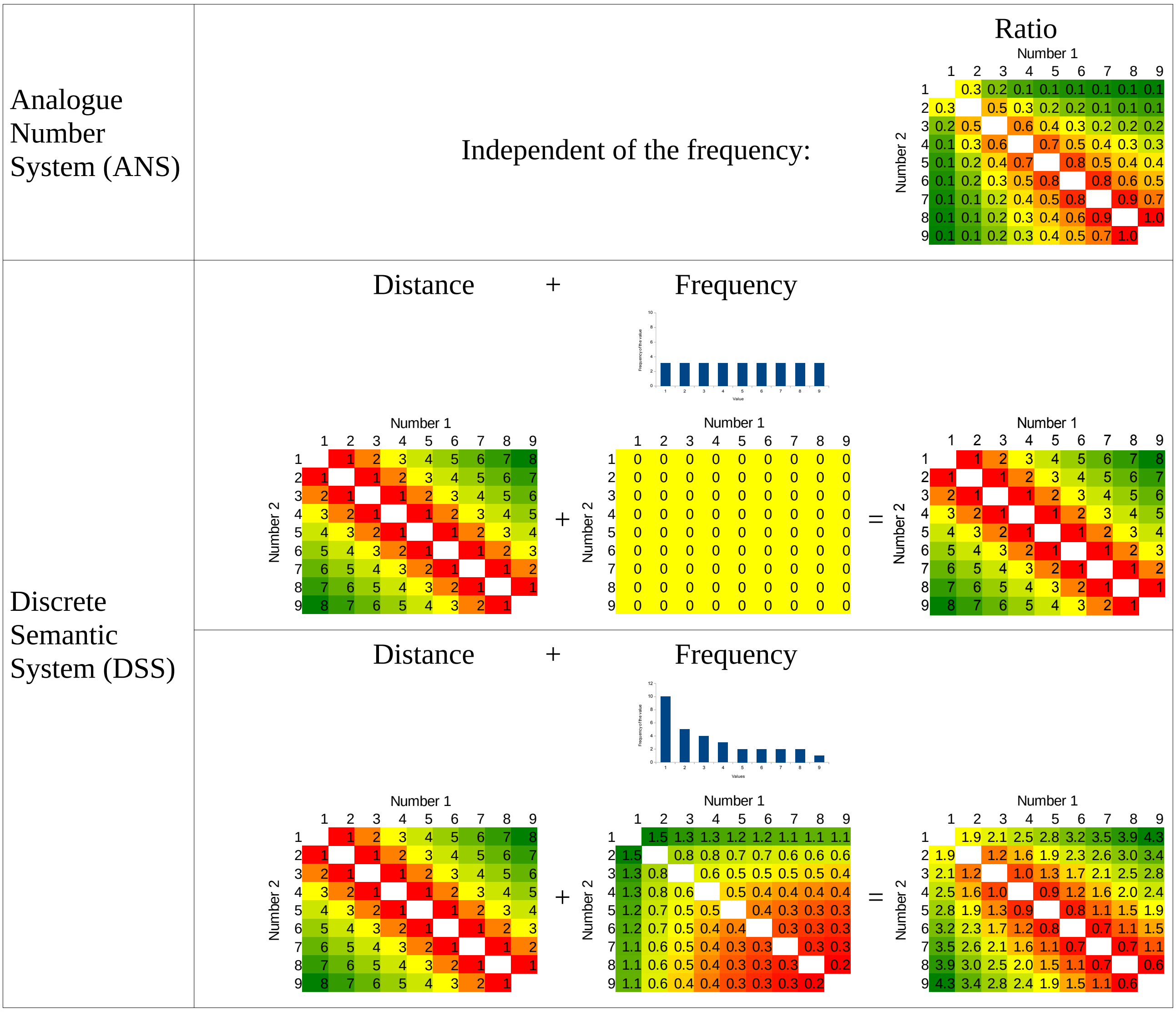


Figure 17. Prediction of the two models for the symbol frequency manipulation in Experiment 2. Bar charts show the frequency of the stimuli used in the uniform distribution condition and in the Indo-Arabic-like distribution condition. (In the Indo-Arabic-like distribution the resulting performance is computed as 0.4×Distance+Frequency.)

Methods

Participants learned new symbols (Figure 18) for the numbers between 1 and 9 to compare, while the frequency of all symbols was manipulated in two conditions.

It is possible that the new symbols are not connected to the numerical values they represent, and they may be processed only as a non-numerical ordered series. This could cause a problem, because the ANS could not process this non-numerical order[5]. To ensure that the new symbols were connected to the numerical values they represent, at the end of the experiment we used a priming task to measure the priming distance effect between the newly learned symbols and familiar Indo-Arabic digits (Figure 18). In a priming distance effect (PDE) the reaction time to the target is faster when the numerical distance between the prime and the target is smaller, reflecting a semantic relation between the prime and the target (Koechlin, Naccache, Block, & Dehaene, 1999; Reynvoet & Brysbaert, 1999; Reynvoet, De Smedt, & Van den Bussche, 2009).

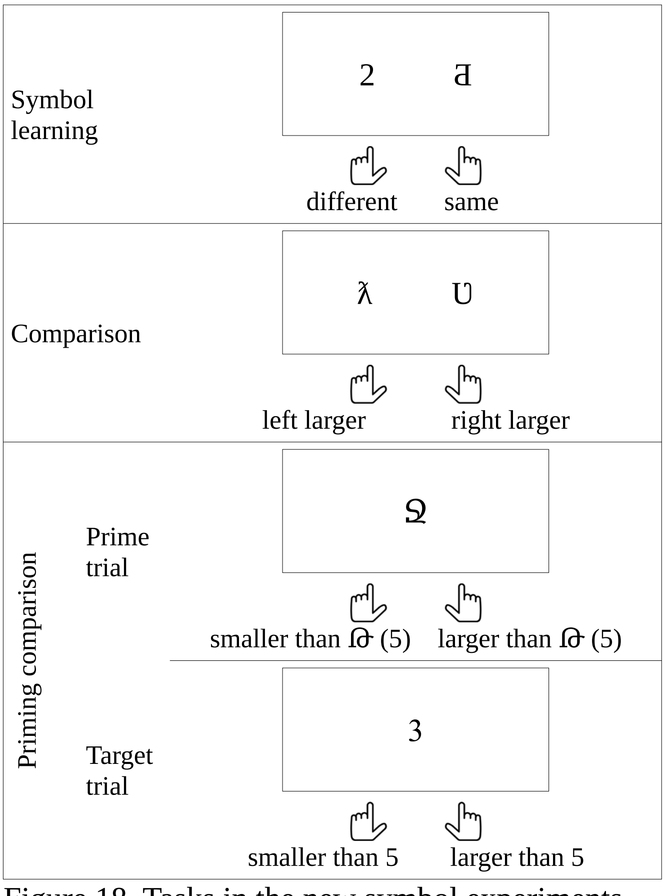


Figure 18. Tasks in the new symbol experiments.

Participants. Eighteen university students participated in the uniform frequency distribution condition. After excluding 2 of them because the error rate did not fall below 5% even after the 5th block, and excluding 2 further participants showing higher than 5% error rates in the main comparison task, the data of 14 participants was included (11 females, mean age 20.6 years, standard deviation 2.1 years). Fifteen university students participated in the Indo-Arabic-like

5 Note, however, that several works suggest that order processing and quantity processing rely on the same mechanisms (Leth-Steensen & Marley, 2000; Marshuetz, Reuter-Lorenz, Smith, Jonides, & Noll, 2006; Verguts & Van Opstal, 2014), thus, ANS should be activated even when the new symbols are non-numerical orders.

frequency distribution condition. After excluding two participants because their error rates were higher than 5% either in the main comparison or in the priming comparison task, the data of 13 participants was analyzed (13 females, mean age 24.3 years, standard deviation 6.9 years).

Stimuli and procedure. The participants first learned new symbols for the numbers between 1 and 9. Then, in a comparison task they decided which number is larger in a simultaneously presented symbol pair. Finally, in a priming comparison task the participants decided whether one-digit numbers are smaller or larger than 5 (Figure 18).

New symbols were introduced to represent values between 1 and 9. The new symbols were chosen from writing systems that were mostly unknown to the participants, and the characters had similar vertical and horizontal size. The symbols were randomly assigned to values for all participants, i.e., the same symbol could mean a different value for different participants, from the following characters: Ꮆ, Ꭹ, Ꮂ, Ꮼ, Ꮬ, Ł, Ɵ, Đ, И, Я, Ꮧ, Ħ, Œ, Ꮷ, ƛ, Ꮭ, Ʊ, Ʊ.

To ensure that the participants have learned the symbols, in the symbol learning phase, the symbols were practiced until a threshold hit rate was reached. In a trial of the new symbol learning phase a new symbol and an Indo-Arabic digit were shown simultaneously, and the participant decided whether the two symbols denote the same value. The stimuli were visible until response. After the response, auditory feedback was given. In a block all symbols were presented 10 times (90 trials in a block) in a randomized order. In half of the trials the symbols denoted the same values. The symbol learning phase ended if the error rate in a finished block was smaller than 5%, or the participant could not reach that level in five blocks.

In the main comparison task, the same procedure was used as in the first experiment, but here the numbers were denoted with the new symbols. In the uniform frequency distribution condition the number of the presentation of a digit were the same as in the first experiment (all possible number pairs were shown 10 times). In the Indo-Arabic-like frequency distribution condition the frequencies of the specific values followed the frequencies of the numbers in everyday life (Dehaene & Mehler, 1992), specified with the following formula: $frequency_{value} = value^{-1} \times 10$. This formula generated the following frequencies (value:frequency): 1:10, 2:5, 3:4, 4:3, 5:2, 6:2, 7:2, 8:2, 9:1 (Figure 17). The 2-permutations of these numbers excluding ties were presented, resulting in 794 trials.

In the priming comparison task in odd (prime) trials a new symbol was visible, and the participant decided whether it was smaller or larger than 5. Two hundred ms after the response in an even (target) trial an Indo-Arabic digit was shown, and the participant decided whether it was smaller or larger than 5. Two thousand ms after the response a new odd (prime) trial was shown. The stimuli were visible until response. The instruction included the value of 5 in both notations: For even trials Indo-Arabic notation (“5”), for odd trials the new notation (e.g., “Ꮂ”) was used. All possible new symbols were presented with all possible Indo-Arabic digits three times, resulting in 192 trials.

Results and discussion

To summarize the main results, in the uniform distribution comparison task the distance effect was present, but the size effect was not (Figure 19A). This result is in line with the DSS model, but not with the ANS model. On the other hand, in the Indo-Arabic-like, biased frequency comparison task both the distance and the size effects were visible (Figure 19B) in a similar pattern as observable in Indo-Arabic number comparison (Figure 16), suggesting that it is the frequency manipulation that is responsible for the size effect.

A Equal frequencies condition

Number on the right (columns), Number on the left (rows)

	1	2	3	4	5	6	7	8	9
1		4%	4%	2%	0%	0%	0%	0%	0%
2	6%		5%	1%	0%	0%	0%	0%	0%
3	1%	4%		2%	0%	1%	1%	0%	0%
4	2%	1%	1%		5%	2%	2%	0%	1%
5	1%	0%	2%	4%		2%	1%	0%	0%
6	0%	0%	1%	2%	4%		7%	1%	3%
7	0%	0%	0%	2%	2%	6%		7%	2%
8	0%	0%	0%	1%	0%	1%	2%		4%
9	0%	0%	0%	0%	0%	0%	1%	4%	

Number on the right (columns), Number on the left (rows)

	1	2	3	4	5	6	7	8	9
1		1274	1196	1026	989	927	1088	844	810
2	1283		1309	1275	1163	1153	1085	1000	885
3	1270	1432		1591	1292	1261	1141	1021	888
4	1176	1354	1556		1381	1328	1202	1024	887
5	1066	1225	1326	1502		1387	1211	1124	908
6	998	1134	1278	1343	1358		1435	1239	956
7	1075	1154	1230	1237	1247	1463		1229	960
8	945	1003	1009	1147	1040	1288	1179		1008
9	782	825	821	823	780	885	846	911	

B Biased frequencies condition

Number on the right (columns), Number on the left (rows)

	1	2	3	4	5	6	7	8	9
1		2%	1%	1%	0%	0%	0%	0%	0%
2	4%		5%	1%	1%	0%	2%	2%	0%
3	0%	5%		5%	2%	2%	0%	1%	2%
4	0%	2%	2%		1%	1%	3%	0%	0%
5	0%	2%	4%	4%		13%	8%	4%	0%
6	0%	0%	0%	8%	19%		8%	0%	0%
7	0%	1%	0%	1%	4%	6%		13%	0%
8	0%	0%	1%	0%	4%	10%	13%		19%
9	0%	0%	2%	0%	8%	0%	4%	0%	

Number on the right (columns), Number on the left (rows)

	1	2	3	4	5	6	7	8	9
1		856	815	822	811	817	800	784	737
2	926		1206	1127	1168	1174	1117	937	885
3	880	1234		1451	1287	1322	1190	1094	924
4	923	1191	1411		1508	1461	1270	1222	896
5	882	1171	1359	1484		1756	1362	1114	994
6	871	1241	1299	1202	1845		1569	1184	1037
7	818	1054	1237	1362	1370	1554		1342	1037
8	843	938	989	1257	1169	1131	1272		1002
9	770	819	860	847	935	1274	1151	1098	

Figure 19. Error rates (left) and response times in ms (right) in the new symbol number comparison for the whole stimulus space. Green denotes fast and error-free responses, red denotes slow and erroneous responses. (**A**) Equal frequencies condition, showing distance and end effects. (**B)** Biased frequencies condition, showing distance, size and end effects.

Distance and size effects in the uniform frequency distribution. The same analysis methods were applied as in the first experiment with two exceptions. Descriptive data clearly shows an end effect (Leth-Steensen & Marley, 2000). Thus, an end effect regressor was also included in the multiple linear regression (Figure 15) with a value of 1 if any of the presented numbers were 9, 0.5 if any of the numbers were 8 or 1, and 0 otherwise. These values were specified with first calculating the average reaction time for all presented numbers, then the distance effect (distance from 5) of the middle number range (i.e., without end effect) was extrapolated, and finally, the deviation from this extrapolation at the end of the number range was estimated.

In the multiple linear regression the slope of the distance effect deviated from zero, 95% CI was [-1.04%, -0.48%], $t(13) = -5.84$, $p < 0.001$ for error rates, and CI was [-73.6 ms, -26.1 ms], $t(13) = -4.53$, $p = 0.001$ for reaction time. On the other hand, the slope of the size effect did not differ from zero, CI with [-0.15%, 0.06%], $t(13) = -0.933$, $p = 0.368$ for error rates, and CI with [-26.6 ms, 13.9 ms], $t(13) = -0.679$, $p = 0.509$ for reaction time. The end effect was present for the reaction time, CI of [-430.6 ms, -147.6 ms], $t(13) = -4.41$, $p = 0.001$, and more unstably for the error rates, CI with [-0.23%, 2%], $t(13) = 1.71$, $p = 0.111$.

These results also demonstrated an end effect (the most extreme values in the set are easier to respond than other values) (Leth-Steensen & Marley, 2000), however, while the end effect can be in line with the DSS model (Leth-Steensen & Marley, 2000), it is also possible that the effect is irrelevant in the description of the representation processing the numerical values (Balakrishnan & Ashby, 1991; Piazza, Giacomini, Le Bihan, & Dehaene, 2003), consequently, the presence of this effect is not decisive in the present question.

Distance and size effects in the Indo-Arabic-like frequency distribution. The slope of the distance effect differed from zero in both the error rates, CI with [-1.56%, -0.5%], $t(12) = -4.25$, $p = 0.001$, and in reaction times, [-55.7 ms, -28.9 ms], $t(12) = -6.87$, $p < 0.001$. The non-zero slope of the size effect was also observable, [0.20%, 0.68%], $t(12) = 3.99$, $p = 0.002$ for the error rate, and CI with [28.4 ms, 50.2 ms], $t(12) = 7.85$, $p < 0.001$ for the reaction time. Additionally, the end effect was observable in the reaction times, CI with [-622.5 ms, -294.9 ms], $t(12) = -6.1$, $p < 0.001$, but not in the error rates, CI with [-2.76%, 0.7%], $t(12) = -1.3$, $p = 0.217$.

We tested directly whether the size effects of the two frequency conditions differed. The size effect slopes between the uniform frequency distribution and the Indo-Arabic-like frequency distribution conditions differed significantly in both the error rates, $U = 13$, $p < 0.001$, and in the reaction times, $U = 15$, $p < 0.001$.

Priming distance effect. In this analysis the error rates and median reaction times of the correct responses of the target Indo-Arabic numbers were analyzed as a function of the prime (new symbols) – target (Indo-Arabic digit) distance (Figure 20). Only the trials in which the response was the same for the prime and distance (i.e., both numbers were smaller than 5, or both numbers were larger than 5) were analyzed (Koechlin et al., 1999; Reynvoet & Brysbaert, 1999; Reynvoet et al., 2009). Linear slope was calculated for the priming distance effect.

In the uniform frequency distribution the data of one participant was not recorded due to technical problems. Because in the symbol learning task participants practiced the new symbol – Indo-Arabic pairs, the zero distance pairs could have this extra practice gain, and not depend purely on the semantic priming effect. Thus, the 0 distance pairs were not included in the analysis. While the descriptive data showed increasing priming effect with smaller distance (Figure 20), the effect was not significant: In the uniform frequency condition CI is [-1.62%, 2.69%], $t(12) = 0.54$, $p = 0.599$ for the error rate, and CI is [-1.4 ms, 39.7 ms], $t(12) = 2.03$, $p = 0.065$ for the reaction time, and in the Indo-Arabic frequency condition CI is [-0.13%, 1.63%], $t(12) = 1.85$, $p = 0.089$ for the error rates, and CI is [-8.9 ms, 27.2 ms], $t(12) = 1.11$, $p = 0.290$ for the reaction time. The lack of significance could mean the lack of PDE, or it could reflect the lack of statistical power, or both. Looking at the gradual increase of error rate and reaction time as the function of priming distance (Figure 20) and the biased CIs, it seems more probable that the PDE could be statistically significant with larger statistical power. To extend the reasoning that the lack of the significance might be the result of insufficient statistical power, we also analyzed three unpublished similar experiments conducted in our laboratory, where in the same design new symbols were learned with the same stimuli and procedure as in the present works (in the third unpublished experiment the learning and the comparison were repeated for 5 days). In those experiments the PDE was measured with similar sample sizes as in the experiments presented here. We found that in all cases the confidence interval was biased to the direction the PDE predicts, although mostly it was only close to be significant. In the first two unpublished experiments 95% CI is [-6.2 ms, 17.7 ms], $N = 12$, $p = 0.312$, and [29.1 ms, 75.1 ms], $N = 10$, $p < 0.001$. In the third unpublished experiment the PDE was measured for 5 consecutive days which is especially informative about the consistency and fluctuation of the PDE in a relatively small sample: 95% CI [10.25 ms, 34.54 ms], $N = 13$, $p = 0.002$, [-2.09 ms, 28.21 ms], $p = 0.085$, [-1.05 ms, 14.39 ms], $p = 0.084$, [-9.34 ms, 8.13 ms], $p = 0.882$, [-0.49 ms, 13.24 ms], $p = 0.066$, for the five days, respectively. A meta-analysis on the five available experiments (second and third experiments of the present paper and three unpublished experiments; only day 1 was used from the last unpublished experiment; meta-analysis of means in original units with random effect) revealed 95% CI [6.7 ms, 34.3 ms], $p = 0.004$ (Cumming, 2013). The analysis also confirms that the effect size would require much larger sample to have a significant result reliably in a single experiment: The estimated effect size could be as small as $d = 0.3$ (with around 25 ms standard deviation), which would require a magnitude of 100 participants to reach 95% statistical power (Faul, Erdfelder, Lang, & Buchner, 2007). Taken together, based on (a) the expected gradual pattern of the PDE (Figure 20), and (b) the consistently biased CIs across experiments, (c) confirmed with the meta-analysis, it is most reasonable to conclude that the PDE is present, even if our usual sample size around 15 does not guarantee the preferred 95% statistical power for a single experiment.

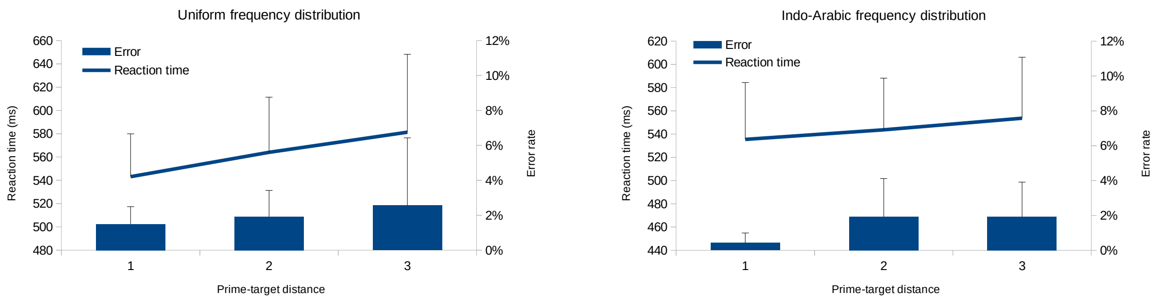


Figure 20. Prime distance effect (PDE) measured in error rates (bars) and reaction time (lines), in equal frequency condition (left) and in biased frequency condition (right) in Experiment 2. Error bars represent 95% confidence interval.

Effect of the frequency. To further demonstrate the effect of the frequency (because it cannot be observed readily on Figure 19), the mean reaction time was calculated for all cells that include a specific value in both conditions (right of Figure 21). The reaction time changes in line with the frequencies of the values: The more frequent a number is in one condition compared to the other condition (left of Figure 21), the faster it is to process (right of Figure 21). In other words, the differences of the two conditions for the values in the reaction time data are inversely proportional to the differences of the two conditions for the values in the frequency. Note that the reaction time data do not include purely the frequency effect, because (a) middle values are gradually slower to process because of the interaction of the distance effect and the shape of the stimulus space, and (b) end values are faster to process because of the end effect.

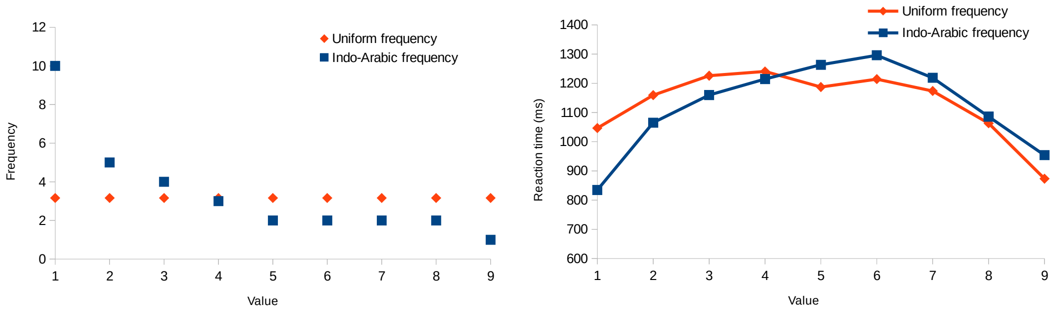


Figure 21. Frequencies of the specific values (left) and response latencies for those values (right) in Experiment 2.

Summary. In the second experiment the numerical distance and size effects dissociated. More specifically, the numerical size effect was missing when the frequency distribution was uniform, and the size effect was present with the biased frequencies of symbols, suggesting that size effect was guided by the frequencies of the symbols. These results cannot be explained by the ANS model, whereas they can be in line with the DSS model. We highlight again that according to the ANS model although the frequency might slightly modulate the performance, it cannot change a large proportion of the variance in the performance. However, the present result reveal that largest part of the variance of the size effect is directed by the frequency, while the ratio has no observable effect (as revealed by the statistical lack of the size effect), contradicting the ANS model prediction.

Results also show that the new numbers semantically primed the Indo-Arabic digits as revealed by the priming distance effect, demonstrating that the new symbols were connected to the values they represent. Thus, the lack of the size effect in the second experiment cannot be the result of potentially non-numeric new symbols which could not be processed by the ANS.

Experiment 3 – Role of the semantic congruency effect in the size effect

As another potential confound, it is possible that in the second experiment there was a size effect in the uniform distribution condition, however, a semantic congruency effect (SCE) extinguished it. According to the SCE, large numbers are responded to faster than small numbers when the task is to choose the larger number, resulting in a reversed size-like effect, and small numbers are faster to decide on when the smaller number should be chosen, resulting in a regular size-like effect (Leth-Steensen & Marley, 2000). If the SCE was present in the second experiment, this anti-size effect could have extinguished a potentially existing size effect. To test this possibility, the uniform frequency condition of the second experiment was rerun, but this time the participants had to choose the smaller number. If the SCE was present in the second experiment as a reversed size-like effect, then it should be observed in the present experiment as a regular size-like effect, increasing the size effect. However, the size effect was not present in this control experiment, demonstrating that the SCE did not mask a potentially existing size effect.

Methods

The methods of the second experiment was applied, however, participants had to choose the smaller number, not the larger, in the comparison task. The priming comparison task was not run.

Eighteen university students participated in the study. Two participants were excluded, because their error rates were higher than 5% after the 5th learning block, and two participants were excluded because they had higher than 5% error rate in the comparison task. The data of 14 participants were analyzed, 10 females, with mean age of 25.4 years, and standard deviation of 6.9.

Results

Distance and size effects. In the multiple linear regression analysis the distance effect was present in both the error rate and the reaction time, CI [-1.54%, -0.48%], $t(13) = -4.13$, $p = 0.001$, and CI [-77.0 ms, -39.6 ms], $t(13) = -6.73$, $p < 0.001$, respectively. More critically, the size effect was not observable neither in the error rate nor in the reaction time, CI [-0.18%, 0.15%], $t(13) = -0.184$, $p = 0.857$, and CI [-25.0 ms, 26.2 ms], $t(13) = 0.0482$, $p = 0.962$, respectively. Comparing the slopes of the uniform frequency condition of the second and the present experiments, the slopes of the size effects did not differ significantly, neither in error rate nor in reaction time, t(26) = 0.33, $p = 0.744$, and $U = 91.5$, $p = 0.783$, respectively. Thus, choosing the smaller number did not change the size effect, consequently, the SCE did not influence essentially the size effect in the second experiment.

General discussion

We introduced the DSS model as a comprehensive alternative account to the ANS model to explain symbolic number processing. First, we have shown that the DSS model can explain many symbolic numerical effects, and we demonstrated that the DSS model could give a similar quantitative prediction for symbolic number comparison performance as the ANS model. Second, we tried to contrast the two models in Indo-Arabic comparison task. However, because of the relatively high noise and the uncertainties of the diffusion analysis method, it was not possible to find a straightforward preference for any models. On the other hand this result could show that the DSS model prediction empirically fits the Indo-Arabic number comparison as good as the ANS model prediction. Finally, the second and third experiments revealed that in new symbol comparison tasks the numerical size effect is the consequence of the frequency manipulations of the symbols, as proposed by the DSS model, and not the consequence of the ratios of the values, as predicted by the ANS model. These data also show that the numerical distance and size effects are not straightforward signs of the ANS, because an alternative mechanism could produce them as well.

While the second and the third experiments utilized new symbols, it is possible to extend our conclusion about other symbolic number comparisons, for example, the Indo-Arabic number comparison. Because all known numerical effects that were observable in the new symbol comparison show the very same pattern as in Indo-Arabic comparison (i.e., distance effect, size effect and priming distance effect), it is parsimonious to suppose that the same mechanisms work behind new symbol comparison and Indo-Arabic comparison, and our findings can also be generalized to the Indo-Arabic and other symbolic number processing, unless additional data show the opposite.

We argue that the ANS model is not in line with our results. While one can try to modify the ANS model to align with the present result, ratio-based performance is a defining feature of the ANS, and changing that feature leads not only to a modified ANS model, but to a completely new model. Additionally, adding frequency effect to the ANS model cannot modify it to explain the frequency-based size effect, because the ANS critically suggests that the performance should mainly be driven by the ratio, which ratio effect in fact was statistically invisible in the second and third experiments.

We argue that the ANS model is not in line with the present results, and the DSS can be an appropriate alternative. However, one might question how strongly our results support a DSS model. Obviously, one can only tell if a model is in line with the empirical results, and whether the model is coherent. We argue that the DSS is in line with the present and previous results (e.g., it can explain the independent distance and size effects, why symbolic and non-symbolic comparisons are relatively independent, or how arbitrarily precise comparison can be made), and it is a coherent model. Additionally, based on current cognitive models, it is reasonable to suppose that abstract symbolic operations are processed by a system that is otherwise known to be used for other symbolic operations, such as the mental lexicon or a conceptual network. On the other hand, no one can exclude that an alternative, third model could account for these results, and not the ANS or the DSS models. Further research can tell whether the DSS framework is an appropriate explanation for the symbolic number processing or another alternative should be found. Furthermore, it is possible that it is not a single representation that is responsible for the discussed effects, but cooperation of several representations is required, and although the ANS cannot explain the distance effect in comparison task, still there could be other symbolic numerical phenomena that could be rooted in the ANS. Additional works can find out whether such a partial role can be attributed to the ANS in symbolic number processing.

The DSS model in its current form relies on models about mental lexicon or conceptual networks. These starting points could offer many properties of the models, while at the same time, many other details are seemingly missing, e.g., the exact quantitative description of the comparison performance. While these shortcomings might make the impression that the DSS model is less detailed than the ANS model, these differences are the consequence of changing the base of the explanations. While the ANS model is a low-level perceptual model in its nature, the DSS model is more like a linguistic or conceptual network model. Models describing higher level functions are usually less quantitative than models describing lower level functions, partly because of methodological reasons, and from this viewpoint it seems reasonable that a DSS model is less quantitative than an ANS model. However, from a different—and more relevant—viewpoint, the DSS model is as efficient as the ANS model, because seemingly all relevant symbolic numerical effects and phenomena can also be explained in the DSS model, and a few examples can already be found where the DSS model can give a better explanation than the ANS model.

The ANS model is a widely accepted and deeply grounded explanation for number processing. However, despite the huge amount of papers discussing and supporting the ANS view, they are relying on surprisingly few effects and findings that demonstrate an ANS activation. In fact, the few phenomena can also be explained in the alternative DSS model as well. Additionally and more importantly, an increasing number of findings are not in line with the ANS model. For example, symbolic and non-symbolic performance seems to be independent on many behavioral (Holloway & Ansari, 2009; Sasanguie et al., 2014; Schneider et al., 2016) and neural level (Bulthé et al., 2014, 2015; Damarla & Just, 2013; Lyons et al., 2015). In a correlational study it has been shown that distance and size effects dissociate in Indo-Arabic comparison task (Krajcsi, 2016). Some results show that the numerical representation is not analogue: Functional activation in the brain while processing symbolic numbers seems to be discrete (Lyons et al., 2015), and symbolic numbers can also interfere with the discrete yes-no responses (Landy, Jones, & Hummel, 2008). The present finding showing the frequency dependence of the size effect also extends the list of results contradicting the ANS model. Future research can tell whether the ANS can be reformulated to account for these findings, or an alternative model, such as the DSS, can characterize symbolic number processing better.

Acknowledgments

We thank Réka Eszter Börcsök, Dávid Czikora, József Fiser, Ildikó Király and Krisztina Peres for their comments on an earlier version of the manuscript.

References

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception and Psychophysics*, *50*, 555–564.

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *NeuroImage*, *87*, 311–322. https://doi.org/10.1016/j.neuroimage.2013.10.049

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2015). Visual Number Beats Abstract Numerical Magnitude: Format-dependent Representation of Arabic Digits and Dot Patterns in Human Parietal Cortex. *Journal of Cognitive Neuroscience*, *27*(7), 1376–1387. https://doi.org/10.1162/jocn_a_00787

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91. https://doi.org/10.1016/j.tics.2008.11.007

Crossman, E. R. F. W. (1955). The measurement of discriminability. *Quarterly Journal of Experimental Psychology*, *7*(4), 176–195. https://doi.org/10.1080/17470215508416692

Cumming, G. (2013). *Understanding The New Statistics: Effect Sizes, Confidence Intervals, and Meta-Analysis*. Routledge.

Damarla, S. R., & Just, M. A. (2013). Decoding the representation of numerical values from brain activation patterns. *Human Brain Mapping*, *34*(10), 2624–2634. https://doi.org/10.1002/hbm.22087

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition* (Vol. XXII, pp. 527–574). Cambridge: Harvard University Press.

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

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355–361.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal representation in human intraparietal cortex. *Neuron*, *37*, 719–725.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, Instruments and Computers*, *39*, 175–191.

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Gallistel, C., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496–499.

Hauser, M. D., & Spelke, E. (2004). Evolutionary and Developmental Foundations of Human Knowledge: A Case Study of Mathematics. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences (3rd ed.)* (pp. 853–864). Cambridge, MA, US: MIT Press.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395. https://doi.org/10.3758/BF03202431

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: An entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1845–1864. https://doi.org/10.1037/a0016892

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: a practical introduction*. London: Academic.

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1882–1905. https://doi.org/10.1037/0096-1523.25.6.1882

Krajcsi, A. (2016). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, 1–8. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., Lengyel, G., & Kojouharova, P. (submitted). Symbolic number comparison is not processed by the analogue magnitude system: different symbolic and nonsymbolic numerical distance and size effects.

Krajcsi, A., Lengyel, G., & Laczkó, Á. (submitted). Interference between numerosity and parity: discrete representation in number processing.

Landy, D. H., Jones, E. L., & Hummel, J. E. (2008). Why Spatial-Numeric Associations Aren't Evidence for a Mental Number Line. In *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 357–362).

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *70*(1), 12–23. https://doi.org/10.1037/cep0000070

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of Fechner Day*, *27*(1), 357–362.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*(1), 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Lyons, I. M., Ansari, D., & Beilock, S. L. (2015). Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. *Human Brain Mapping*, *36*(2), 475–488. https://doi.org/10.1002/hbm.22641

Marshuetz, C., Reuter-Lorenz, P. A., Smith, E. E., Jonides, J., & Noll, D. C. (2006). Working memory for order and the parietal cortex: An event-related functional magnetic resonance imaging study. *Neuroscience*, *139*(1), 311–316. https://doi.org/10.1016/j.neuroscience.2005.04.071

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Noël, M.-P., & Rousselle, L. (2011). Developmental Changes in the Profiles of Dyscalculia: An Explanation Based on a Double Exact-and-Approximate Number Representation Model. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00165

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 1–1. https://doi.org/10.1167/5.5.1

Patro, K., Nuerk, H.-C., Cress, U., & Haman, M. (2014). How number-space relationships are assessed before formal schooling: A taxonomy proposal. *Developmental Psychology*, *5*, 419. https://doi.org/10.3389/fpsyg.2014.00419

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Piazza, M., Giacomini, E., Le Bihan, D., & Dehaene, S. (2003). Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging. *Proc. R. Soc. Lond. B*, *270*, 1237–1245.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an amazonian indigene group. *Science*, *306*, 499–503.

Pinhas, M., & Tzelgov, J. (2012). Expanding on the Mental Number Line: Zero Is Perceived as the “Smallest.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1187–1205. https://doi.org/10.1037/a0027390

Potts, G. R. (1972). Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning & Verbal Behavior, 11*(6), 727–740. https://doi.org/10.1016/S0022-5371(72)80007-0

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*, 191–201.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children’s representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *The Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, n/a-n/a. https://doi.org/10.1111/desc.12372

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Tzelgov, J., Ganor-Stern, D., & Maymon-Schreiber, K. (2009). The representation of negative numbers: exploring the effects of mode of processing and notation. *Quarterly Journal of Experimental Psychology*, *62*(3), 605–624. https://doi.org/10.1080/17470210802034751

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: a neuronal model. *Journal of Cognitive Neuroscience*.

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics*, *3*, 189–230. https://doi.org/10.1080/00140136008930484

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

Chapter 4: Symbolic numerical distance effect does not reflect the difference between numbers

The present chapter has been published:

Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. Frontiers in Psychology, 8. https://doi.org/10.3389/fpsyg.2017.02013

Symbolic numerical distance effect does not reflect the difference between numbers

Attila Krajcsi[1], Petia Kojouharova[1, 2]

[1]Department of Cognitive Psychology, Institute of Psychology, ELTE Eötvös Loránd University

[2]Doctoral School of Psychology, ELTE Eötvös Loránd University

Both authors contributed equally to this work.

In a comparison task, the larger the distance between the two numbers to be compared, the better the performance, a phenomenon termed the numerical distance effect. According to the dominant explanation, the distance effect is rooted in a noisy representation, and performance is proportional to the size of the overlap between the noisy representations of the two values. According to alternative explanations, the distance effect may be rooted in the association between the numbers and the small-large categories, and performance is better when the numbers show relatively high differences in their strength of association with the small-large properties. In everyday number use the value of the numbers and the association between the numbers and the small-large categories strongly correlate, thus, the two explanations have the same predictions for the distance effect. To dissociate the two potential sources of the distance effect, in the present study participants learned new artificial number digits only for the values between 1 and 3, and between 7 and 9, thus, leaving out the numbers between 4 and 6. It was found that the omitted number range (the distance between 3 and 7) was considered in the distance effect as 1, and not as 4, suggesting that the distance effect does not follow the values of the numbers predicted by the dominant explanation, but it follows the small-large property association predicted by the alternative explanations.

Keywords: symbolic number processing; numerical distance effect; analogue number system; discrete semantic system;

The numerical distance effect and its explanations

In a symbolic number comparison task, performance is better (i.e., error rates are lower and reaction times are shorter) when the numerical distance is relatively large, e.g., comparing 1 vs. 9 is easier than comparing 5 vs. 6 (Moyer & Landauer, 1967). There are several explanations for this phenomenon termed the numerical distance effect.

According to the dominant model, numbers are stored on a continuous (analogue) and noisy representation, the Analogue Number System (ANS). The numbers are stored as noisy signals, and the closer two numbers are on the ANS, the larger the overlap of the two respective signal distributions is. Because comparison performance is better when the overlap is relatively small, large distance number pairs are easier to process because of the smaller overlap between the signals . More specifically, the ANS works according to Weber's law, therefore, comparison performance depends on the ratio of the two numbers to be compared (Moyer & Landauer, 1967). In fact, the distance effect is the consequence of this ratio effect, because larger distance also means higher ratio. The ratio effect is also thought to be the cause of the numerical size effect: Comparison performance is better for smaller numbers than for larger numbers, because smaller number pairs have larger ratio than larger number pairs with the same distance (Moyer & Landauer, 1967). The ANS is thought to be the essential base of numerical understanding (Dehaene, 1992), and numerical distance effect is believed to be a diagnostic signal of the ANS activation while solving a numerical task.

However, there could be another explanation for the cause of the distance effect. Recently, it has been proposed that symbolic numerical effects, such as the distance and size effects, can be explained by a representation similar to the mental lexicon or conceptual networks, where nodes of the network represent the digits, and connections between them are formed according to their semantic and statistical relations (Krajcsi, Lengyel, & Kojouharova, 2016). In this model, termed the Discrete Semantic System (DSS) model, numerical distance and size effects are rooted in two different mechanisms, even if the combination of these effects looks similar to the formerly supposed ratio effect. According to the model the size effect might depend on the frequencies of the numbers: Smaller numbers are more frequent than larger numbers (Dehaene & Mehler, 1992), therefore, smaller numbers are easier to process, producing the numerical size effect. See a similar frequency-based explanation of the size effect in the model of Verguts, Fias and Stevens (2005). At the same time, numerical distance effect could be based on the relations of the numbers, for example, similarly to the phenomenon in a picture naming task, where priming effect size depended on the semantic distance between the prime and target pictures (Vigliocco, Vinson, Damian, & Levelt, 2002). There are several other alternative number processing models with partly overlapping suppositions and predictions as the DSS model (Leth-Steensen, Lucas, & Petrusic, 2011; Nuerk, Iversen, & Willmes, 2004; Pinhas & Tzelgov, 2012; Proctor & Cho, 2006; Verguts & Fias, 2004; Verguts et al., 2005; Verguts & Van Opstal, 2014). See the comparison of those models in Krajcsi et al. (2016) and in Krajcsi, Lengyel and Laczkó (in press). Supporting the alternative DSS model, it has been found that the size effect followed the frequency of the digits in an artificial number notation comparison task (Krajcsi et al., 2016). Also, it has been shown in a correlational study that in symbolic number comparison task, the distance and the size effects were independent (Krajcsi, 2017), reflecting two independent mechanisms generating the two effects. (See a similar prediction for independent distance and size effects in Verguts et al., 2005; Verguts & Van Opstal, 2014.)

Because of the DSS model and the empirical findings demonstrating that the size effect is a frequency effect and that the distance and size effects are independent, it is essential to reconsider how the distance effect is generated. According to the DSS model, different explanations consistent with the supposed network architecture are feasible. First, it is possible that based on the values of the numbers connections with different strengths between the numbers are formed – numbers with closer values have stronger connections – and stronger connections create interference in a comparison task, resulting in a distance effect. This explanation is similar to the ANS model in a sense that value-based semantic relations are responsible for the distance effect. As an alternative explanation, it is also possible that based on previous experiences numbers are associated with the “small” and the “large” properties, e.g., large digits, such as 8 or 9, are more strongly associated with “large”, and small digits, such as 1 or 2, are more strongly associated with “small”. These associations could influence the comparison decision, and number pairs with larger distance might be easier to process, because the associations of the two numbers with the small-large properties differ to a larger extent. A similar explanation has been proposed formerly in a connectionist model, which model predicted several numerical effects successfully, and one key component of this model was that distance effect relies on the connection between the number layer and the “larger” nodes, where relatively large numbers are associated with the “larger” node more strongly than relatively small numbers (Verguts et al., 2005).

Thus, the explanations of the numerical distance effect suppose two different sources for the effect: According to the ANS model and to the value-based DSS explanation, the effect is rooted in the *values or the distance of the numbers*, while in the association-based DSS explanation and in the connectionist model, the effect is rooted in the *strength of the associations between the number and the small-large properties*. The two explanations are not exclusive, it is possible that both information sources contribute to the distance effect.

The two critical properties of the two explanations, i.e., the values or distance of the numbers, and the association between the numbers and the small-large properties, strongly correlate in the number symbols used in everyday numerical tasks, therefore, in those cases one cannot specify their role in the distance effect. However, in a new artificial number notation the two factors (the distance of the values and the association) could be manipulated independently. This is only possible if the distance effect is notation specific, otherwise the new symbols would get the association strengths of the already known numbers, instead of forming new association strengths between the new symbols and the small-large properties. It is possible that the numerical effects are notation specific, as has been already demonstrated in the case of the numerical size effect: In an artificial number notation comparison task, the size effect followed the frequency of the digits, which also means that the size effect is notation specific (Krajcsi et al., 2016).

The aim of the study

The present study investigates whether in a new artificial number notation, where the values of the digits and the small-large associations do not necessarily correlate, the distance effect is influenced by the distance of the values or by the small-large associations, or both. One way to dissociate the two properties is to use a number sequence in which some of the values are omitted (Figure 22). If the distance effect is directed by the distance of the values, then the measured distance effect should be large around the gap (in this example the effect should be measured as 4 units large), while if the distance effect is directed by the small-large associations, then the measured distance effect should be small around this gap, measured as an effect with a single unit distance, supposing that the new digits were used in a comparison task with equal probability. If both mechanisms contribute to the distance effect, then the distance effect should be measured somewhere between the single unit and the many units (in this example 4 units) distance.

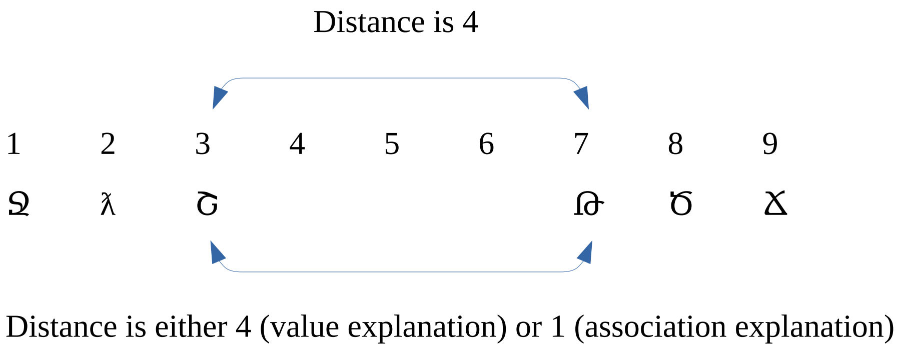


Figure 22. An example of the symbols and their meanings in the present study. Arrows show the predicted distance effect size based on the predictions of the two explanations.

Why does the association explanation predict a distance effect of 1 distance around the gap? In a comparison task, the association between a digit and the small-large properties may depend on how many times the digit were judged as smaller or larger. If the new digits are used with equal probability in the comparisons (and if the distance effect is notation specific), then the probability of being smaller or larger than the other number can be specified easily (see Table 10). In our example (Figure 22), the number 1 is always smaller, thus the association frequency is 100% with the small property, and 0% with the large property. The number 2 is smaller when compared with 3, 7, 8 and 9, and larger when compared with 1, thus the association frequency is 80% small and 20% large. Continuing the example, the association frequency is directly proportional to the order of the symbols and not to their value. If the distance effect depends on the order, then the distance between 3 and 7 (i.e., the two digits around the gap) is the same as any other neighboring digits (see the specific values in Table 10).

Example symbols	Ձ	ƛ	Ꞡ		Ꭷ	Ծ	Ⲭ
Meaning of the symbols	1	2	3		7	8	9
Chance of being smaller in a comparison	100%	80%	60%		40%	20%	0%
Chance of being larger in a comparison	0%	20%	40%		60%	80%	100%

Table 10. The chance of being smaller or larger in a comparison task when the symbols are presented with equal probability.

The two explanations predict different effect sizes for the distance effect not only for the two numbers next to the gap (e.g., for 3 vs 7 on Figure 22 and Table 10), but for any number pairs in which the two numbers are on the opposing side of the gap. The possible number pairs of the new symbols seen on Figure 22 and their hypothetical distance effect sizes according to the two explanations can be seen on Figure 23: Columns and rows denote the two numbers to be compared, and the cells show the distances of the value pairs (darker cells mean smaller distance). In the value explanation (left side) the predicted distance is the difference of the two numbers, while in the association explanation (right side) the predicted distance is computed based on the strength of the association with the small-large properties when the numbers are presented with equal probability, which is simply the order of those symbols in that series. The comparison performance should be proportional to the distance, thus, these figures show the performance pattern predictions according to the two explanations. The results will be displayed in a similar way as seen here, because (a) displaying the full stimulus space is more informative than other indexes of distance effects, since any systematic deviation from the expected patterns could be observed, and because (b) with the relatively large number of cells any systematic pattern could be a convincing and critical information independent of the statistical hypotheses tests.

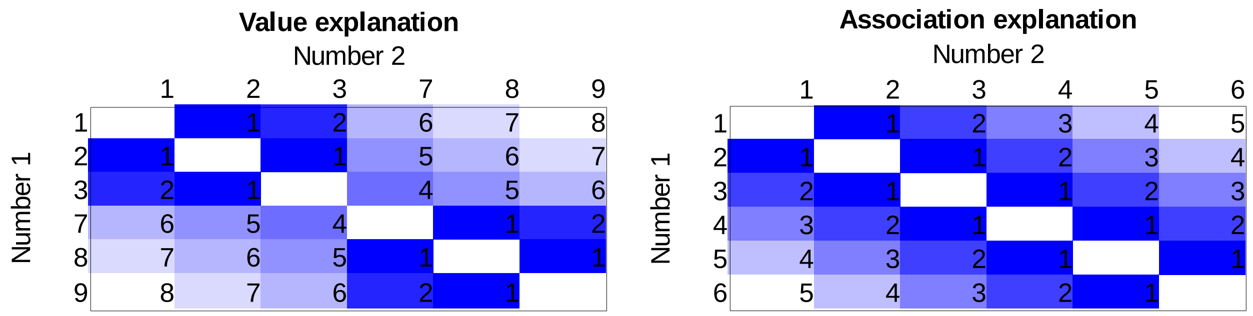


Figure 23. The expected distance effect pattern for the stimulus space used in the present study based on the value explanation (left side) and based on the association explanation (right side). Specific values in the cells are the difference of the values (value model) or the difference of the order (association model) of the numbers to be compared on an arbitrary scale. Darker color indicates worse performance.

In the present test it is critical that the new symbols should represent their intended values, and not a series that is independent of the intended number meanings, otherwise the new symbols could be considered by the participants as numbers e.g., from 1 to 6 because of their order in the new symbol series, which in turn could generate the performance predicted by the association explanation, even if the effect would be based on their values. One way to ensure that the new symbols are sufficiently associated to their intended values is to ensure that the priming distance effect works between the new and a well-known (for example, Indo-Arabic) notations. In numerical comparison tasks, the decision about the actual trial might be influenced by the stimulus of the previous trial, and the size of the influence is proportional to the numerical distance of the previous and actual stimuli, termed the priming distance effect (PDE) (Koechlin, Naccache, Block, & Dehaene, 1999; Reynvoet & Brysbaert, 1999). The PDE is considered to be a sign of the relation between the symbols or the overlap of their representations (Opstal, Gevers, Moor, & Verguts, 2008). Former experiments have shown that new artificial symbols can cause PDE in Indo-Arabic numbers (Krajcsi et al., 2016), suggesting that the new digits are not a series of symbols independent of their intended values, but they can be considered as a notation for the respective numbers. In the ANS framework, the PDE reflects the representational overlap between the numbers, thus the PDE demonstrates that both notations appropriately activate the same representation, the ANS.

To summarize, the present study investigates whether the distance effect follows the distance of the values of the numbers (left of Figure 23) or the association of the small-large properties (right of

Figure 23) or both, in the case of a newly learned notation (Figure 22), where some of the symbols are omitted. If both explanations are true, then we expect a pattern in-between the two figures, i.e., we should observe a break between 3 and 7 similar to the value explanation, however, the difference between the two sides of the gap should not be as large as in that explanation. All of these predictions only hold if the distance effect is notation specific, otherwise, the distance effect reflects the already well-known numbers, where the value and the association strongly correlates, and the pattern seen on the value model prediction can be expected. Consequently, only a pattern seen on the right in Figure 23 can decide about the models, because a pattern seen on the left can either mean a value-based distance effect, or it can mean that the distance effect is notation independent.

Methods

In the present experiment participants learned new symbols (Figure 24), with the meaning of the numbers between 1 and 3, and between 7 and 9 (Figure 22). Then a number comparison task was performed with the new symbols (Figure 24).

Stimuli and procedure. The new symbols were chosen from writing systems that were mostly unknown to the participants (e.g., ƌ, ƛ, Ʊ, ⵕ). The characters had similar vertical and horizontal size, and similar visual complexity, and the height of the symbols were approximately 2 cm. (Because most probably apparent size does not influence the effects we investigate here, the visual angle was not controlled strictly.) Numbers were displayed in white on gray background. The symbols were randomly assigned to values for all participants, i.e., the same symbol could mean a different value for different participants.

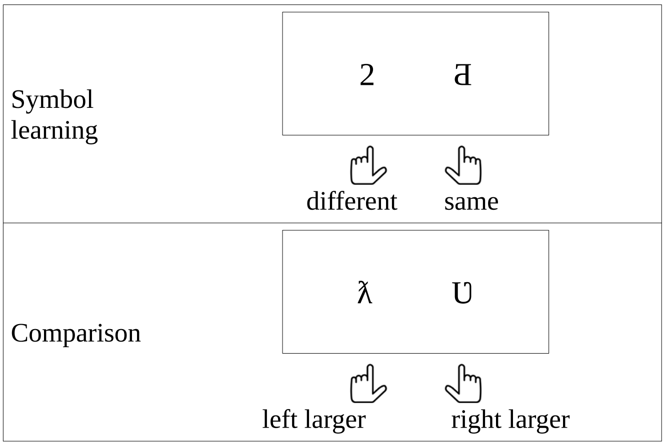


Figure 24. Tasks in the new symbol experiment.

The participants first learned new symbols for the numbers between 1 and 3, and between 7 and 9 (Figure 24). To ensure that the participants have learned them in the learning phase, symbols were practiced until a threshold hit rate was reached. In a trial a new symbol and an Indo-Arabic digit were shown simultaneously, and the participant decided whether the two symbols denoted the same value by pressing the R or I key. The stimuli were visible until response. After the response, auditory feedback was given. In a block all symbols were presented 10 times (60 trials in a block) in a randomized order. In half of the trials the symbols denoted the same values. The symbol learning phase ended if the error rate in a completed block was smaller than 5%, or the participant could not reach that level in five blocks.

In the following comparison task the participants decided which number is larger in a simultaneously presented new symbol pair by pressing the R or I key (Figure 24). In a trial two numbers were shown until response, and the participants chose the larger one. Numbers to be compared could be between 1 and 3, and between 7 and 9. After the response, auditory feedback was given. All possible number pairs including the applied numbers, excluding ties, were shown 15 times, resulting in 450 trials.

Presentation of the stimuli and measurement of the responses were managed by the PsychoPy software (Peirce, 2007).

Participants. Twenty-three university students participated in the experiment for partial course credit. After excluding 4 participants showing higher than 5% error rates (higher than the mean + the standard deviation of the error rates in the original sample) in the comparison task, the data of 19 participants was analyzed (16 females, mean age 22.2 years, standard deviation 4.6 years).

Results

All participants successfully reached a lower than 5% error rate within 3 blocks in the symbol learning task, therefore, no participants were excluded for not learning the symbols within 5 blocks.

For all participants the mean error rates and the mean reaction times for correct responses were calculated for all number pairs. Data of participants with a higher-than-5% mean error rate were excluded (higher than the mean + the standard deviation of the error rates in the original sample). The mean error rates and reaction times of the group are displayed in Figure 25 for the whole stimulus space. Visual inspection of the error rate pattern suggests that partly the value model can be observed, although the data are rather noisy, as reflected in some outlier cells. In the case of reaction time, it is more straightforward that the pattern is more in line with the association model (see the two expected pure patterns in Figure 23). In the reaction time data one can also observe the end effect: number pairs including the largest number in the range (i.e., 9) are faster to process (Leth-Steensen & Marley, 2000; Scholz & Potts, 1974). (There are different possibilities concerning what causes the end effect. It is possible that participants learn that 9 is the largest number in the actual session, therefore, when 9 is displayed, no further consideration is required in a comparison task. Alternatively, according to the ANS model, it is possible that because in the session, number 9 has neighboring number only on one side, the overlap between the noisy signal distributions should be smaller, leading to a faster response (Balakrishnan & Ashby, 1991).)

Number on the right (rows: Number on the left)

Number on the left	1	2	3	7	8	9
1		4.2%	2.5%	0.0%	0.4%	0.7%
2	3.2%		3.2%	1.4%	0.0%	1.1%
3	0.7%	4.6%		3.5%	1.4%	1.8%
7	0.0%	0.7%	3.2%		5.3%	0.7%
8	0.7%	0.0%	0.7%	4.6%		2.8%
9	0.4%	1.1%	1.1%	0.4%	2.1%	

Number on the right (rows: Number on the left)

Number on the left	1	2	3	7	8	9
1		1416	1319	1216	1142	860
2	1445		1566	1350	1252	897
3	1387	1698		1532	1227	945
7	1239	1403	1692		1503	957
8	1081	1158	1303	1484		993
9	885	901	889	959	963	

Figure 25. Error rates (left) and reaction times (in ms, right) in the whole stimulus space.

To test the results statistically, we first fit the two predictions of the models (Figure 23) to the group average of the error rate and the reaction time data (Figure 25) with a simple linear regression, where one of the model prediction was the explanatory variable, and one of the behavioral performance measurements was the dependent variable. Then the goodness of the fit measured as R^2 was calculated (R^2 columns in Table 11), and the correlations of two models were compared with the method described in Steiger (1980) for every performance measurement (Difference of the group fits column in Table 11). As an alternative method, we calculated the R^2 values for every single participants for both the value and association models, and the R^2 of these model fits, as ordinal variables, were compared pairwise with Wilcoxon signed-rank test (Better model for the participants columns in Table 11).

To fit the distance effect appropriately, the end effect should also be considered, and its variance should be removed from the data. Inspection of the descriptive data on Figure 25 suggests that number pairs including the number 9 were involved in the end effect in the present study. One possibility to remove the end effect is to apply multiple linear regression, and beyond the distance effect regressor, an end effect regressor (e.g., 1 if the number pair includes 9, 0 otherwise) also should be utilized. The problem with this solution is that the end effect not only shortened response latency for number pairs including 9, but it also decreased the slope of the distance effect in those cells (see the less steep distance effect in the row and column with 9 than in other rows and columns). Because the end effect is not added linearly to the distance effect, a multiple linear regression could not describe this nonlinear aspect of the end effect, which in turn would distort the distance effect results. As an alternative method, to remove the end effect all cells with number pairs including 9 were removed from the analysis (i.e., the bottom row and the right column on Figure 25), and only the distance effect regressors were used. Therefore, for all linear fits (Table 11) in both the group average and the participants level, the number pairs including 9 were removed.

Regarding the possible difference between the goodness of fit of the two models, we note that the difference is limited by the fact that the two models correlate, e.g., the value model can be considered as a modified association model with an additional increase of the values in the top-right and bottom-left part of the stimulus space seen in Figure 23. Therefore, if one model is appropriate, the other, inappropriate model should show some non-zero R^2 value, too, although the R^2 should be smaller than the R^2 of the appropriate model.

Results for the goodness of fits (Table 11, linear model columns on the left) show that whereas in the error rates the two models are indistinguishable, in the reaction time patterns the association model seems to describe the data better, in line with the visual inspection of the data.

Although error rate and reaction time data are highly informative, the recently becoming more popular diffusion model analysis could draw a more sensitive picture (Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). In the diffusion model, decision is based on a gradual accumulation of evidence offered by perceptual and other systems, and decision is made when appropriate amount of evidence is accumulated. Reaction time and error rates partly depend on the quality of the information (termed the drift rate) upon which the evidence is built. Drift rate is considered to be the most important parameter that influences the number comparison performance and the task difficulty . Importantly, observed reaction time and error rate parameters can be used to recover the drift rates (Ratcliff & Tuerlinckx, 2002; Wagenmakers, van der Maas, & Grasman, 2007). Drift rates can be more informative than the error rate or the reaction time, because drift rates reveal the sensitivity of the background mechanisms more directly (Wagenmakers et al., 2007). To recover the drift rates for all number pairs, the EZ diffusion model was applied (Wagenmakers et al., 2007). The EZ model supposes that some of the parameters do not play a role in the response generation, and the model investigates and recovers only the drift rate, the decision threshold and the non-decision time parameters. If one can suppose that only these three parameters play a role in the responses, the EZ model can be utilized. Importantly, one essential advantage of this method is that unlike most other diffusion parameter recovery methods, EZ can be used when the number of trials per cells are relatively small. For edge correction we used the half trial solution, i.e., for error rates of 0%, 50% or 100%, the actual error rate was modified with the percent value of 0.5 trial, e.g., in a cell with 15 trials and 0% error rate, the corrected error rate was 0.5/15 which is 3.33% (see the exact details about edge correction in Wagenmakers et al., 2007). The scaling within-trials variability of drift rate was set to 0.1 in line with the tradition of the diffusion analysis literature. Drift rates for all number pairs and participants were calculated. The mean drift rates of the participants (Figure 26) show a similar pattern observed above for the former descriptive data. Fitting the two predictions of the models, the association model shows again a better fit (Table 11). Additionally, (a) the largest difference between the goodness of fit of the two models can be observed for the drift rates (compared to the error rate and the reaction time data), and (b) the highest R^2 value is found for the drift rates, suggesting that the drift rate indeed captures the difficulty of the comparison tasks more sensitively than the error rates or the reaction times do.

Number on the right

Number on the left	1	2	3	7	8	9
1		0.141	0.156	0.168	0.171	0.208
2	0.144		0.147	0.155	0.171	0.188
3	0.160	0.135		0.134	0.151	0.185
7	0.169	0.155	0.127		0.143	0.183
8	0.189	0.173	0.165	0.140		0.181
9	0.200	0.188	0.197	0.187	0.188	

Figure 26. Drift rate values in the whole stimulus space.

The analysis above supposed that the distance effect (either coming from the value model or from the association model) is linear. However, a logarithmic, or a similar function with decreasing change as the distance increases might be a better option to describe the data. First, one cannot suppose a linear distance effect, because after a sufficiently large distance the reaction time should be unreasonably short or even negative, which would not make sense. Second, in a former artificial symbol comparison task, where the missing size effect did not influence the distance effect, the distance effect was better described with the logarithm function than with a linear function (unpublished results in Krajcsi et al., 2016). For these reasons, the analysis of goodness of fit was repeated with logarithmic distance effect models, in which the regressors were the natural logarithm of the values of the previously used linear models seen in Figure 23. The results (Table 11, logarithm model columns on the right) show that (a) for all three data types (error rate, reaction time and drift rate) the association model fits better than it did with the linear regressor models, and (b) the differences of the two models are larger than they were for the linear regressor models. Overall, the largest difference between the value and the association models can be seen in the logarithm model versions for the drift rates.

	Linear model (Figure 23)				Logarithm model			
	Value model R^2	Association model R^2	Difference of the group fits	Better model for the participants	Value model R^2	Association model R^2	Difference of the group fits	Better model for the participants
Error rate	0.709	0.708	Z = 0.008, p = .993	T = 73, p = .376	0.714	0.821	Z = -1.091, p =.275	T = 92, p = .904
Reaction time	0.543	0.790	Z = -2.294, p = .022	T = 44, p = .040	0.457	0.817	Z = -3.646, p < .001	T = 34, p = .014
Drift rate	0.526	0.861	Z = -3.647, p < .001	T = 39, p = .024	0.425	0.874	Z = -5.748, p < .001	T = 18, p = .002

Table 11. Goodness of fit of the models (measured as R^2) and comparison of the correlations (Difference column) for the error rates, reaction times and drift rates patterns based on the group average data, and hypothesis tests for choosing the better model based on the participants' data.

Whereas our present main interest is the nature of the distance effect, it is worth to note that no size effect can be found in the data: The regressor formed as the sum of the two numbers to be compared (e.g., the regressor value for the 3 vs. 4 number pairs is 7) does not fit either the error rates (R^2 = 0.001), or the reaction time (R^2 = 0.01), or the drift rate (see below) data (R^2 = 0.001). These data replicate the results of Krajcsi et al. (2016), confirming that in new symbols with equal frequency of numbers in a comparison task, the size effect does not emerge, and also confirm that distance and size effects may dissociate. Relatedly, we note that the size effect could not influence the fit of the distance effect not only because the size effect could not be demonstrated in the present data, but also because the size effect regressor (sum of the numbers to be compared) does not correlate with distance effect regressor (difference of the numbers to be compared) at all.

Reliability of the results

To investigate the reliability of the present results two additional experiments are summarized here: (a) the whole experiment was repeated with another sample, and (b) the data of a follow-up study was analyzed where the same paradigm was used with Indo-Arabic numbers instead of new symbols, to see if the distance effect can follow the associations of the numbers and small-large responses in an already well-established notation (Kojouharova & Krajcsi, submitted). (a) In the replication study, 41 university students participated. Four of them were excluded, either because they did not reach the required maximum 5% error rate after 5 blocks of symbol learning, or because they used wrong response keys. Five additional participants were excluded, because they had higher than 6.5% error rate (which was the mean + standard deviation error rate in that sample) in the comparison task. As a result, the data of 32 participants were analyzed (mean age was 21.0 years, 3 males). The error rate, reaction time and drift rate means for the whole stimulus space can be seen in Figure 27, and the R^2s of the models with the appropriate hypothesis tests are displayed in Table 12. While the reaction time and drift rate means replicate the results of the main study (although the difference was significant only with the comparison of the group fits, but not with the hypothesis test choosing the better fit for the participants), the error rates show the superiority of the value model. (b) In the Indo-Arabic comparison task 23 university students participated. One participant was dyscalculic whose data was excluded from further analysis, and 2 further participants were excluded for having an error rate higher than 5%. Thus, the data of 20 participants were analyzed (mean age was 20.15 years, 4 males). The goodness of fit of the logarithmic models and their contrast can be seen in Table 13. The Indo-Arabic study replicated the results of the main study, and also in the error rates the association model fitted significantly better than the value model.

Number on the right

Number on the left	1	2	3	7	8	9
1		8.5%	2.9%	1.0%	1.0%	0.4%
2	9.2%		5.6%	1.9%	0.4%	0.4%
3	2.3%	8.5%		2.9%	2.1%	1.7%
7	1.0%	1.3%	3.8%		6.9%	2.5%
8	0.4%	0.6%	2.5%	8.5%		8.1%
9	0.6%	0.8%	1.3%	4.0%	9.4%	

Number on the right

Number on the left	1	2	3	7	8	9
1		1250	1252	1038	970	810
2	1374		1429	1212	1035	845
3	1263	1509		1301	1050	856
7	1118	1270	1386		1245	922
8	985	1047	1136	1194		996
9	823	839	818	967	925	

Number on the right

Number on the left	1	2	3	7	8	9
1		0.132	0.151	0.175	0.182	0.225
2	0.123		0.132	0.157	0.173	0.205
3	0.154	0.130		0.136	0.170	0.196
7	0.164	0.146	0.134		0.130	0.175
8	0.188	0.173	0.152	0.128		0.146
9	0.206	0.199	0.207	0.161	0.148	

Figure 27. Error rates (top left), reaction times (in ms, top right) and drift rates (bottom) in the whole stimulus space in the replication study.

	Linear model (Figure 23)				Logarithm model			
	Value model R^2	Association model R^2	Difference of the group fits	Better model for the participants	Value model R^2	Association model R^2	Difference of the group fits	Better model for the participants
Error rate	0.791	0.629	*Z* = 2.041, *p* = .041	*T* = 130, *p* = 0.012	0.862	0.724	*Z* = 2.081, *p* = .037	*T* = 150, *p* = 0.033
Reaction time	0.610	0.719	*Z* = -1.258, *p* = 0.208	*T* = 233, *p* = 0.562	0.517	0.713	*Z* = -2.236, *p* = .025	*T* = 196, *p* = 0.204
Drift rate	0.768	0.914	*Z* = -2.727, *p* = .006	*T* = 232, *p* = 0.550	0.695	0.929	*Z* = -4.284, *p* < .001	*T* = 191, *p* = 0.172

Table 12. Goodness of fit of the models (measured as R^2) and comparison of the correlations (Difference column) for the error rates, reaction times and drift rates patterns based on the group average data, and hypothesis tests for choosing the better model based on the participants' data in the replication study.

	Logarithm model			
	Value model R^2	Association model R^2	Difference of the group fits	Better model for the participants
Error rate	0.634	0.825	*Z* = -2.766, *p* = .006	*T* = 17, *p* = .001
Reaction time	0.749	0.917	*Z* = -3.737, *p* < .001	*T* = 14, *p* < .001
Drift rate	0.681	0.864	*Z* = -3.080, *p* = .002	*T* = 31, *p* = .006

Table 13. Goodness of fit of the models (measured as R^2) and comparison of the correlations (Difference column) for the error rates, reaction times and drift rates based on the group average data, and hypothesis tests for choosing the better model based on the participants' data in the Indo-Arabic study (Kojouharova & Krajcsi, submitted).

Looking strictly at the significance of the results, the replication shows a somewhat different result pattern as the first measurement, because in error rate, significant differences support the value model instead of the association model, and in reaction time and drift rate, not all hypothesis tests are significant. Clearly, some non-significant effects might reflect not only the lack of an effect, but also the lack of statistical power, and significant effects can also be type-I errors (there is especially a chance for this, when replication studies find opposing significant effects). To evaluate the accumulated data, a mini meta-analysis was run on the three set of data (Maner, 2014). Binary random-effects with the DerSimonian-Laird method (Viechtbauer, 2010; Wallace et al., 2012) was performed on the logarithm model fit data measuring the ratio of participants where the association model was better than the linear model. While the error rate does not show a clear preference for any models (45.9% mean preference for the association model with 95% CI of [17.5%, 74.2%]) reaction time and drift rate clearly prefers the association model (76.6% with CI of [65.0%, 88.2%] for reaction time, and 72.9% with CI [58.2%, 87.7%] for drift rate). Taken together, while the reaction time and drift rate show the superiority of the association model, the results of the error rates are ambiguous. It is important to highlight that from the viewpoint of the present question, reaction time and especially drift rates are more relevant. First, reaction time data are usually considered to be more reliable and sensitive than error rate, because whereas error rate and reaction time data measure two strongly correlating constructs, error rate measures it in a dichotomous scale, while reaction time is a continuous scale, thus, the latter having more information about the trial performance. Second, drift rate measures the difficulty of the task more sensitively than error rates or reaction times in themselves (Wagenmakers et al., 2007) (this is also confirmed by the usually higher R^2 values for drift rates than for reaction times or error rates). Therefore, we consider that reaction times and drift rates reliably reflect the superiority of the association model over the value model. At the same time, it might be a question of future research whether heterogeneous error rates are the result of random noise or whether there are aspects of the performance that partly reflects the functioning of the value model.

To summarize the results, it was found that (a) the association model described the distance effect better than the value model did, measured with reaction time and drift rate, while error rate displayed an inconsistent pattern, (b) drift rate draws more straightforward picture than the reaction time or the error rate data, (c) logarithmic type distance effect describes the data more precisely than the linear distance effect, and finally, (d) size effect is absent in the present paradigm with uniform number frequency distribution.

Discussion

The present work investigated whether the numerical distance effect is rooted in the values of the numbers to be compared or in the association between the numbers and the small-large properties. In a new artificial number notation with omitted numbers, the distance effect measured with reaction time and drift rate did not follow the values of the numbers, as it would have been suggested by the mainstream ANS model or by the value-based explanation of the DSS model, instead, the effect reflected the association between the numbers and the small-large categories, as proposed by the association-based explanation of the DSS model or by the delta-rule connectionist model of numerical effects (Verguts et al., 2005). Measured with error rate, the results were not conclusive, so it is the question of additional studies whether the inconsistency in the error rate data is simply noise or whether there are additional aspects of the distance effect that should be investigated with more sensitive methods.

Together with the present results several findings converge to the conclusion that the symbolic number comparison task cannot be explained by the ANS. First, unlike the prediction of that model suggesting that distance and size effects are two ways to measure the single ratio effect, symbolic distance and size effects are independent (Krajcsi, 2017), and the distance effect can be present even when no size effect can be observed (shown in the present results and in Krajcsi et al., 2016). Second, the size effect follows the frequency of the numbers as demonstrated in Krajcsi et al. (2016), and also in the present results where the uniform frequency of the digits induced no size effect (i.e., the slope of the size effect is zero). Third, the present data demonstrated that the distance effect is not directed by the values of the digits as predicted by the ANS model, but they are influenced by the frequency of the association with the small and large categories (see also the extension of the present findings for Indo-Arabic numbers in Kojouharova & Krajcsi, submitted).

The present and some former results also characterize the symbolic numerical comparison task, an alternative model should take into consideration: (a) symbolic distance and size effects are independent (Krajcsi, 2017; Krajcsi et al., 2016), (b) the effects are notation independent (the present results and Krajcsi et al., 2016), (c) the size effect depends on the frequency of the numbers (the present results and Krajcsi et al., 2016), (d) the distance effect depends on the association between the numbers and the small-large categories (present results), and (e) the distance effect can be described with a logarithm of the difference of the values (present results).

We highlight again that these results are not the consequence of the possibility that the new symbols are not related to their intended values and that the independent series of symbols would create a performance pattern similar to the association model prediction, because it was already shown that the new symbols prime the Indo-Arabic numbers, revealing that the new symbols denote their intended values (Krajcsi et al., 2016), and the present finding were also replicated with Indo-Arabic numbers (Kojouharova & Krajcsi, submitted).

From a methodological point of view it is worth to note that in the present comparison task, the drift rate seemed to be the most sensitive index to describe performance, which strengthens the role of the diffusion model analysis, among others in cases when sensitivity and statistical power is essential.

To summarize, the results revealed that in an artificial number notation where some omitted numbers might create a gap, the distance effect followed the association with the small-large properties, and not the values of the numbers. This result contradicts the Analogue Number System model and the value-based DSS explanation, that would suggest that distance effect is directed by the values or the ratio of the numbers. On the other hand, the result is in line with the alternative association-based DSS explanation and the delta-rule connectionist model, in which the distance effect is directed by the association between the number nodes and the small-large nodes.

Acknowledgments

We thank Ákos Laczkó and Gábor Lengyel for their comments on an earlier version of the manuscript.

References

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception and Psychophysics*, *50*, 555–564.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1882–1905. https://doi.org/10.1037/0096-1523.25.6.1882

Kojouharova, P., & Krajcsi, A. (submitted). The Indo-Arabic distance effect originates in the response statistics of the task.

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Laczkó, Á. (in press). Interference between numerosity and parity: discrete representation in number processing. *Experimental Psychology.*

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics*, 357–362.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review*, *107*(1), 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Maner, J. K. (2014). Let’s Put Our Money Where Our Mouth Is: If Authors Are to Change Their Ways, Reviewers (and Editors) Must Change With Them. *Perspectives on Psychological Science*, *9*(3), 343–351. https://doi.org/10.1177/1745691614528215

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Nuerk, H.-C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC - and the MARC - (linguistic markedness of response codes) effect. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(5), 835–863. https://doi.org/10.1080/02724980343000512

Opstal, F. V., Gevers, W., Moor, W. D., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*(2), 419–425. https://doi.org/10.3758/PBR.15.2.419

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pinhas, M., & Tzelgov, J. (2012). Expanding on the Mental Number Line: Zero Is Perceived as the “Smallest.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1187–1205. https://doi.org/10.1037/a0027390

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*, 191–201.

Scholz, K. W., & Potts, G. R. (1974). Cognitive processing of linear orderings. *Journal of Experimental Psychology*, *102*(2), 323–326. https://doi.org/10.1037/h0035873

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*(2), 245–251.

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Wallace, B. C., Dahabreh, I. J., Trikalinos, T. A., Lau, J., Trow, P., & Schmid, C. H. (2012). Closing the Gap between Methodologists and End-Users: R as a Computational Back-End. *Journal of Statistical Software*, *049*(i05). Retrieved from http://econpapers.repec.org/article/jssjstsof/v_3a049_3ai05.htm

Chapter 5: Symbolic number comparison is not processed by the analogue number system: different symbolic and nonsymbolic numerical distance and size effects

The present chapter has been published:

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analog number system: different symbolic and nonsymbolic numerical distance and size effects. Frontiers in Psychology, 9. https://doi.org/10.3389/fpsyg.2018.00124

Symbolic number comparison is not processed by the analogue number system: different symbolic and nonsymbolic numerical distance and size effects

Attila Krajcsi[1], Gábor Lengyel[2], Petia Kojouharova[3, 4]

[1]Cognitive Psychology Department, Institute of Psychology, Eötvös Loránd University

[2]Department of Cognitive Science, Central European University

[3]Doctoral School of Psychology, Eötvös Loránd University

[4]Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences

Dominant numerical cognition models suppose that both symbolic and nonsymbolic numbers are processed by the Analogue Number System (ANS) working according to Weber's law. It was proposed that in a number comparison task the numerical distance and size effects reflect a ratio-based performance which is the sign of the ANS activation. However, increasing number of findings and alternative models propose that symbolic and nonsymbolic numbers might be processed by different representations. Importantly, alternative explanations may offer similar predictions to the ANS prediction, therefore, former evidence usually utilizing only the goodness of fit of the ANS prediction is not sufficient to support the ANS account. To test the ANS model more rigorously, a more extensive test is offered here. Several properties of the ANS predictions for the error rates, reaction times and diffusion model drift rates were systematically analyzed in both nonsymbolic dot comparison and symbolic Indo-Arabic comparison tasks. It was consistently found that while the ANS model's prediction is relatively good for the nonsymbolic dot comparison, its prediction is poorer and systematically biased for the symbolic Indo-Arabic comparison. We conclude that only nonsymbolic comparison is supported by the ANS, and symbolic number comparisons are processed by other representation.

Keywords: Analogue Number System; number comparison; Weber's law; diffusion model; symbolic numbers;

Highlights:

- We test whether symbolic number comparison is handled by an analogue noisy system
- Analogue system model has systematic biases in describing symbolic number comparison
- This suggests that symbolic and non-symbolic numbers are processed by different systems

Representation behind symbolic number processing

Analogue Number System

In their seminal work Moyer and Landauer (1967) described that in an Indo-Arabic single digit number comparison task the performance is worse (i.e., reaction time is slower and error rate is higher) when the difference between the two numbers is relatively small (numerical distance effect) or when the numbers are relatively large (numerical size effect). They proposed that the effects are the expression of a general ratio-based effect in which number pairs with smaller ratio are harder to process. This ratio-based performance was thought to be the result of a simple representation working according to Weber's law, termed the Analogue Number System (ANS), similar to the representations working behind simple physical feature comparison tasks. Since then, the ratio-based performance (usually measured only with the distance effect) is thought to be the signal of a noisy analogue representation working in the background.

The ratio-based performance was also specified with quantitative descriptions. Originally, Moyer and Landauer (1967) demonstrated that the reaction time pattern can be described appropriately with a function used at that time in physical property comparison tasks: a $K \times \log(large\_number / (large\_number - small\_number))$ function correlates well with the measured reaction time, $r = 0.75$. Later, more precise mathematical descriptions were offered (see Dehaene, 2007 for an extensive mathematical description of the model). According to one of the implementations of these descriptions, the numbers are stored as noisy representation following a Gaussian distribution, and the noise is proportional to the value of the number. This increasing noise can produce the ratio-based performance. For example, the overlap between the representations of two numbers predicts the error rate in a comparison task, or more generally, this overlap predicts the difficulty of the task, expressed as drift rate in the diffusion model (see more details in the Methods section). (This proportionally increasing noise can also be implemented in a logarithmic representation with constant noise on a logarithmic scale.)

The ANS is supposed to work behind any number comparison, independent of the notation of the numbers (Dehaene, 1992; Nieder, 2005; Piazza, 2010), because the same ratio-based performance can be observed behind symbolic and nonsymbolic tasks (Dehaene, 2007; Moyer & Landauer, 1967), and because overlapping brain areas are activated in symbolic and nonsymbolic number processing (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003; Nieder, 2005). Although there could be differences between the symbolic and nonsymbolic number processing, and even there could be two different representations working with different sensitivity (i.e., Weber fraction), both of these stimuli are processed by the same *type* of representations, which representations work according to Weber's law, producing a ratio-based performance (Dehaene, 2007; Piazza, 2010).

The common mechanism and the strong relation between symbolic and nonsymbolic processing is also reflected by several findings showing that, for example, the sensitivity of the ANS measured in a nonsymbolic dot comparison task correlates with symbolic math achievement (Halberda, Mazzocco, & Feigenson, 2008), or training nonsymbolic number processing improves the symbolic number processing (Park & Brannon, 2013). To summarize, it is widely supposed that number processing is supported by a noisy, analogue representation, working according to Weber's law, and therefore producing a ratio-based performance in comparison tasks. Also, this type of mechanism works behind both symbolic and nonsymbolic number processing, as reflected by many similarities and relations between symbolic and nonsymbolic numerical tasks.

Different symbolic and nonsymbolic number processing

However, there are increasing number of findings in the literature suggesting that the symbolic and nonsymbolic number processing is not backed by the same representation or by the same type of representations. For example, it has been shown that performance of the symbolic and non-symbolic number comparison tasks do not correlate in children (Holloway & Ansari, 2009; Sasanguie, Defever, Maertens, & Reynvoet, 2014). As another example, while former studies found that common brain areas are activated by both symbolic and non-symbolic stimuli (Eger et al., 2003; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004), later works with more sensitive methods found only notation-specific activations (Bulthé, De Smedt, & Op de Beeck, 2014, 2015; Damarla & Just, 2013). In another fMRI study, the size of the symbolic and non-symbolic number activations did not correlate, and more importantly, the activation for the symbolic number processing seemed to be discrete and not analogue (Lyons, Ansari, & Beilock, 2015). According to an extensive meta-analysis, while it was repeatedly found that the simple number comparison task (the supposed index for the sensitivity of the ANS) correlates with mathematical achievement, it seems that nonsymbolic comparison correlates much less with math achievement, than symbolic comparison (Schneider et al., 2017). In another example, Noël and Rousselle (2011) found that while older than 9- or 10-year-old children with developmental dyscalculia (DD) perform worse in both symbolic and non-symbolic tasks than the typically developing children, younger children with DD perform worse than control children only in the symbolic tasks, but not in the non-symbolic tasks, meaning that the deficit is more strongly related to the symbolic number processing, and the impaired non-symbolic performance is only the consequence of the symbolic processing problems. See a more extensive review of similar findings in Leibovich & Ansari (2016). All of these findings are in line with the present proposal, suggesting that symbolic and non-symbolic numbers are processed by different systems.

Additionally, there are a few alternative models that are in line with these later findings showing that symbolic and nonsymbolic number processing is not backed by the same representation or by the same type of systems. In a connectionist model of symbolic number processing, the model successfully explains many phenomena the ANS model cannot handle (Verguts, Fias, & Stevens, 2005; Verguts & Van Opstal, 2014). Although this model is interpreted as a version of the ANS

(Dehaene, 2007; Verguts & Fias, 2004), critically, it does not show the defining feature of the ANS: the model does not produce inherently the ratio-based performance, instead, introduction of the uneven frequency of the digits is necessary to produce the size effect (Verguts & Fias, 2004; Verguts et al., 2005). Thus, the model proposes different type of mechanisms for symbolic and nonsymbolic number processing. Another model assumes that primitives (simple representational units) are stored in the long term memory only for the digits (numbers between 0 and 9) (Pinhas & Tzelgov, 2012), but not for other values (Kallai & Tzelgov, 2009; Tzelgov, Ganor-Stern, & Maymon-Schreiber, 2009), suggesting a symbolic-only representation. In a third model it was proposed that symbolic numbers can be stored in a Discrete Semantic System (DSS), similar to the mental lexicon or a semantic network. In this system numbers are represented by nodes, and the connections of the nodes reflect the semantic relations of the nodes mostly directed by the numerical distance of the number pairs (Krajcsi, Lengyel, & Kojouharova, 2016). The distance effect might be originated in the semantic relation of the nodes, as was seen in the similar semantic distance effect in a picture naming task (Vigliocco, Vinson, Damian, & Levelt, 2002). The numerical size effect could be rooted in the fact that smaller numbers are more frequent than larger numbers (Dehaene & Mehler, 1992), and more frequent numbers can be processed more easily. The DSS model can be easily extended to account for symbolic numerical interference effects as well (Leth-Steensen, Lucas, & Petrusic, 2011; Patro, Nuerk, Cress, & Haman, 2014; Proctor & Cho, 2006). Thus, the DSS can account for symbolic numerical effects, independent of the nonsymbolic number processing.

Importantly, in the DSS account a performance pattern similar to the ANS model can be offered. For example, it is possible that the reaction time could be proportional to the sum of the linear distance effect and the size effect originated in the frequency of the values, which in turn is related to the power of those values (see the justification for this function and similar possibilities in Krajcsi et al., 2016). Figure 28 shows two possible implementations of the ANS and the DSS models, and it reveals that the DSS model might generate a very similar pattern to the one supposed by the ANS model (the correlation of the two presented performance predictions is -0.89).

The similarity of the ANS and the DSS model predictions means that the DSS model could be potentially an appropriate alternative explanation for the observed distance and size effects. Even more importantly, this means that former works investigating whether the ANS model is correct might have found high correlation between the ANS model and the observed performance either because the ANS model is correct, or because it is the DSS model that is correct, and as the ANS model prediction correlates highly with the DSS model prediction, the correlation between the ANS prediction and the performance was only illusory.

Analogue Number System (ANS) model

$$RT = a \times \log(\text{large}/\text{distance}) + b$$

Number 1

Number 2	1	2	3	4	5	6	7	8	9
1		0.3	0.2	0.1	0.1	0.1	0.1	0.1	0.1
2	0.3		0.5	0.3	0.2	0.2	0.1	0.1	0.1
3	0.2	0.5		0.6	0.4	0.3	0.2	0.2	0.2
4	0.1	0.3	0.6		0.7	0.5	0.4	0.3	0.3
5	0.1	0.2	0.4	0.7		0.8	0.5	0.4	0.4
6	0.1	0.2	0.3	0.5	0.8		0.8	0.6	0.5
7	0.1	0.1	0.2	0.4	0.5	0.8		0.9	0.7
8	0.1	0.1	0.2	0.3	0.4	0.6	0.9		1.0
9	0.1	0.1	0.2	0.3	0.4	0.5	0.7	1.0	

Discrete Semantic System (DSS) model

$$RT = a_1 \times \text{distance} + a_2 \times (x_1^{-1} + x_2^{-1}) + b$$

Number 1

Number 2	1	2	3	4	5	6	7	8	9
1		1.9	2.1	2.5	2.8	3.2	3.5	3.9	4.3
2	1.9		1.2	1.6	1.9	2.3	2.6	3.0	3.4
3	2.1	1.2		1.0	1.3	1.7	2.1	2.5	2.8
4	2.5	1.6	1.0		0.9	1.2	1.6	2.0	2.4
5	2.8	1.9	1.3	0.9		0.8	1.1	1.5	1.9
6	3.2	2.3	1.7	1.2	0.8		0.7	1.1	1.5
7	3.5	2.6	2.1	1.6	1.1	0.7		0.7	1.1
8	3.9	3.0	2.5	2.0	1.5	1.1	0.7		0.6
9	4.3	3.4	2.8	2.4	1.9	1.5	1.1	0.6	

Figure 28. Two possible predictions for the reaction time pattern by the ANS model (based on Crossman, 1955; Moyer & Landauer, 1967) and the DSS model (based on Krajcsi et al., 2016) in a comparison task. The prediction of the models on a full stimulus space in a number comparison task of numbers between 1 and 9. Number 1 and 2 are the two values to be compared. White denotes fast responses, red denotes slow responses (note that numerically the ANS function increases, and the DSS function decreases towards the high ratio, but the direction is irrelevant in the linear fit below). The distance effect can be seen as the gradual change when getting farther from the top-left bottom-right diagonal, and the size effect is seen as the gradual change from top-left to bottom-right. Notations: *large*: larger number; *distance*: distance between the two numbers; x_1 and x_2: the two numbers; a, a_1, a_2 and b are free parameters. In the figures the parameters a and a_2 are set to 1, a_1 is 0.4, and parameter b is set to 0. See also the Methods section for the interpretation of these heat map graphs.

To summarize, an increasing body of evidence indicates that symbolic and nonsymbolic numbers might be processed by different types of representations, and there could be appropriate alternative models to explain symbolic number processing, which may also question the suitability of former tests.

The aim of the study

The aim of the present study is to test the appropriateness of the ANS model in comparison tasks more extensively. The appropriateness of the ANS model for both symbolic and non-symbolic notations have been investigated several times, finding that the prediction of the ANS model is similar to the observed performance (e.g., Dehaene, 2007; Moyer & Landauer, 1967). Former studies usually investigated the goodness of fit of the ANS model for the observed performance. However, these former tests are insufficient, because similarity between the ANS model prediction and the observed performance may be caused by alternative models with similar predictions, such as the DSS model. For example, it is possible that in the Moyer & Landauer (1967) study, the r = 0.75 correlation between the observed reaction time and the ANS model prediction is the result of a stronger than r = 0.75 correlation between the DSS model prediction and the observed performance, and the strong correlation between the DSS model and the ANS model predictions (e.g., r = -0.89). Therefore, it is not enough to show that the ANS model's prediction is similar to the observed data, but a more extensive test is needed.

Here we test the appropriateness of the ANS model by investigating whether the ANS model can explain both symbolic and nonsymbolic comparison tasks equally well, or whether there are critical differences between symbolic and nonsymbolic comparison tasks.[6] If the ANS account is correct, then one should expect that the ANS model can describe both symbolic and non-symbolic equally well, as suggested repeatedly in the literature (Dehaene, 2007; Eger et al., 2003; Moyer & Landauer, 1967; Nieder, 2005). However, if there are differences between the symbolic and nonsymbolic notations, one might suppose that the ANS can describe the nonsymbolic comparison appropriately, in line with the fact that nonsymbolic stimuli are visual-perceptual as other physical properties processed by other representations working according to Weber's law (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Gebuis & Reynvoet, 2012; Moyer & Landauer, 1967; Stoianov & Zorzi, 2012), while the ANS model cannot account for the symbolic comparison, as suggested by the alternative symbolic number processing models.

One might question whether this type of test is meaningful, because symbolic and non-symbolic comparison do not necessarily work in the same way, even if the ANS model is correct. For example, there could be additional notation-specific mechanisms that could change behavioral performance, therefore, one cannot expect that the two notations should show the same performance pattern. However, if someone believes that there could be additional components that might influence the behavioral performance, then one must also question whether the findings suggesting ratio-based performance in any comparisons are valid: even if ratio-based performance is observed, the contribution of the hypothesized additional components should be removed, and if that additional component is unspecified, then nothing could be known about the real mechanism in the background. According to this view, the findings of Moyer and Landauer (1967) or any similar results cannot lead to the conclusion that a ratio-based mechanism is working in the background. Overall, one can believe that the current test is invalid, but at the same time it should also be supposed that all tests demonstrating a ratio-based comparison performance are invalid. Even if this viewpoint might seem unusual, it still could be valid. In this case, another types of tests should be found (see for alternative approaches for these tests in Krajcsi, 2017; Krajcsi et al., 2016). But if one thinks that the works that have proposed that ratio-based performance were valid, the present test should be considered to be valid, too.

In the present work we systematically examine whether ANS predicts both symbolic and nonsymbolic number comparison performance equally well. Specifically, we examine (1) whether the error rates can be described equally well by the functions derived from the ANS model, (2) whether the reaction time pattern of the two notations fit each other linearly, and (3) whether the

6 A similar investigation of symbolic and nonsymbolic comparisons testing against the ANS model was done by Dehaene (2007), however, in that study multi-digit Indo-Arabic numbers were utilized. When comparing multi-digit symbolic numbers one might process the numbers power by power, and holistic ANS processing of the number cannot be guaranteed (Hinrichs, Berie, & Mosell, 1982; Huber, Nuerk, Willmes, & Moeller, 2016; Krajcsi & Szabó, 2012; Poltrock & Schwartz, 1984), therefore, in such a test multi-digit symbolic numbers should be avoided. The present work utilizes only single digit Indo-Arabic numbers.

diffusion model drift rates of the two notations can be described by the same analogue representation. According to the widely accepted version of the ANS model, the model should predict any comparison equally well, because the same ANS-type mechanism processes any numbers independent of their notations. On the other hand, the alternative views might suggest that the ANS should work relatively well only for the nonsymbolic notation, but it should work relatively poorly for symbolic notation, because symbolic precise numbers are processed by other mechanisms. Finally, from a methodological point of view, it is also possible that the difference between the ANS and the alternative models is much smaller than the typical noise in the measured data, thus, even if there are differences between the symbolic and nonsymbolic comparisons, the signal-to-noise ratio is not high enough to reveal the difference. For this reason only different behavioral patterns of symbolic and nonsymbolic comparisons can be conclusive, supporting the alternative accounts, while lack of difference between the symbolic and nonsymbolic comparisons could be either due to the correct ANS description or due to the lack of statistical power.

Methods

Participants compared Indo-Arabic numbers in one condition, and they compared dot arrays in another condition. In both conditions error rate and reaction time were measured.

Stimuli and procedure

In a trial two numbers were visible on the left and on the right sides of the screen, and participants had to choose the larger one by pressing one of the two response keys. The stimuli were visible until key press. The response was followed by an empty screen for 500 ms, then the next trial started.

In the Indo-Arabic condition the numbers were between 1 and 9, to avoid multi-digit numbers (see footnote 6 for more details). All possible pairings of those values were presented, except ties, resulting in 72 possible pairs. All pairs were presented 10 times, resulting 720 trials in the condition. The order of the trials was randomized.

In the dots condition it is not appropriate to use the same 1-9 range as in the Indo-Arabic condition, because sets with less than 5 objects can be enumerated fast, which fast enumeration is termed subitizing (Kaufman, Lord, Reese, & Volkmann, 1949). Subitizing is not an ANS directed process (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008), but it is most probably based on pattern detection (Krajcsi, Szabó, & Mórocz, 2013; Mandler & Shebo, 1982). Therefore, to measure the ANS based dot estimation, the 1-4 range should be avoided. One option could be to use only the numbers between 5 and 9, however, this solution would considerably decrease the stimulus space. Instead, another solution was applied: it was not the 1-9 range itself that was kept in the dot condition, but the ratios of the 1-9 range. Because according to the ANS model, it is the ratio of the numbers that determines performance, changing the values should not change the performance if the ratios of the values are kept. Therefore, to avoid the 1-4 range, and to keep the critical ratio-based feature at the same time, all numbers between 1 and 9 were multiplied by 5, resulting in a number range between

5 and 45.[7] In an array of dots, black and white dots in random positions were shown against a gray background (Dakin et al., 2011), thus, the luminance of the stimuli was not informative about the numerosity. Dots of an array were drawn randomly in a 2×2 degrees area, with a dot diameter of 0.2 degrees, therefore, density and convex hull correlated with the numerosity. Although our stimuli do not control all perceptual features that might influence the perceived numerosity, in the current test, nonnumerical influence of the decision process is less relevant, because the ANS model suggests that number comparison is handled by an analogue system that could be used in any continuous physical feature comparison (Dehaene, 2007; Moyer & Landauer, 1967), hence, in a general sense, any continuous physical feature comparison working according to the Weber's law could be an appropriate task in our test. Additionally, a mixture of visual ratio-based performance and numerosity ratio-based performance should also produce an approximately ratio-based performance, as reflected in the similar psychometric functions of visual comparison and numerical comparison tasks. Therefore, the simple and limited visual control of the stimuli is appropriate for the aim of the current test.[8] As in the Indo-Arabic condition, all possible pairs were presented 10 times, resulting in 720 trials in the condition. The order of the trials was randomized.

The order of the conditions was counterbalanced across participants.

Participants

Twenty-four university students gave informed consent and participated in the study for partial credit course.[9] Four participants were excluded, because their error rates were higher than 1.5 standard deviation + mean error rates at least in one of the conditions (6% in the Indo-Arabic condition and 15% in the dots condition). Among the remaining 20 participants there were 4 males, the age range was 19-24 years, with a mean of 21.0 years.

Analysis methods

Figures used in the results section

To explore the results in more detail, instead of showing the distance and size effects in the traditional way, the full stimulus space is displayed. The left of Figure 28 shows how an ANS predicted pattern would look like. Rows and columns denote the two numbers to be compared, and

7 One might raise that this way the two notations do not use the same number ranges, consequently, the two conditions are not comparable. However, it is important to highlight that the current work tests the ANS model, and in this specific test any modifications that are in line with the ANS model are appropriate. If the ratio-based transformations were not allowed, it would already mean that the ANS model is incorrect, therefore, no further test would be needed.

8 Similar to the reasoning in the previous footnote, we take advantage of the fact that this work is an ANS test, and any addition that is in line with the ANS model, is allowed. If one questions that the number-based comparison performance has different properties than physical feature comparison performance, then the ANS model itself is questioned, therefore, no further test would be needed.

9 Because it is impossible to tell what effects sizes can be expected, or even what properties could differ between the two notations, it is not possible to specify an appropriate sample size in advance. Approximately 25 participants were set as a convenient sample size where the most important effects are firmly observable, but no reasonable prediction could be made regarding the reliability of yet unknown differences between the notations.

the cells include the performance for a specific number pair. In this figure larger values (on an arbitrary scale) and darker colors denote worse performance.

To relate the current figures to the more widely known effects, in Figure 29 some “pure” components of the typical patterns can be seen. Distance effect is displayed as the distance from the top-left and bottom-right diagonal, and size effect is displayed as the distance from the top-left corner along a top-left and bottom-right diagonal. Both effects can also be seen in Figure 28, because the task is harder close to the top-left and bottom-right diagonal (distance effect) and because the task is harder toward the bottom-right corner (size effect). Traditionally, distance and size effects are computed as calculating the mean performance of the cells with the same distance or size values. Sometimes the end effect is also observable (Figure 29), when performance is better with the largest or smallest numbers of the range used in the task (Balakrishnan & Ashby, 1991; Piazza, Mechielli, Butterworth, & Price, 2002; Sathian et al., 1999; Scholz & Potts, 1974).

Distance effect

	1	2	3	4	5	6	7	8	9
1		1	2	3	4	5	6	7	8
2	1		1	2	3	4	5	6	7
3	2	1		1	2	3	4	5	6
4	3	2	1		1	2	3	4	5
5	4	3	2	1		1	2	3	4
6	5	4	3	2	1		1	2	3
7	6	5	4	3	2	1		1	2
8	7	6	5	4	3	2	1		1
9	8	7	6	5	4	3	2	1	

Size effect

	1	2	3	4	5	6	7	8	9
1		3	4	5	6	7	8	9	10
2	3		5	6	7	8	9	10	11
3	4	5		7	8	9	10	11	12
4	5	6	7		9	10	11	12	13
5	6	7	8	9		11	12	13	14
6	7	8	9	10	11		13	14	15
7	8	9	10	11	12	13		15	16
8	9	10	11	12	13	14	15		17
9	10	11	12	13	14	15	16	17	

End effect

	1	2	3	4	5	6	7	8	9
1		0	0	0	0	0	0	0	0
2	0		1	1	1	1	1	1	0
3	0	1		1	1	1	1	1	0
4	0	1	1		1	1	1	1	0
5	0	1	1	1		1	1	1	0
6	0	1	1	1	1		1	1	0
7	0	1	1	1	1	1		1	0
8	0	1	1	1	1	1	1		0
9	0	0	0	0	0	0	0	0	

Figure 29 Distance, size and end effects displayed in the whole stimulus space.

These more detailed figures are more appropriate to explore the performance, because (1) any effects that are slightly deviating from the traditional distance and size effects are more visible, and (2) due to the large number of cells systematic patterns can be identified as reliable effects instead of being a random noise, thus, a continuous change in the pattern might signal a specific effect even without statistical hypothesis tests, and random irregularities can be identified as noise.

Error rate, reaction time and drift rate analysis

Error rate. In psychophysics, specific functions can be found that describe the error rates in a comparison task based on the stimulus intensities and the Weber ratio (Kingdom & Prins, 2010). These functions are also used in the numerical literature (Dehaene, 2007), serving as a firm base to characterize the ANS model prediction. The functions stem from the model summarized in the Introduction, suggesting that error rate is proportional to the overlap of Gaussian noisy representations. In our analysis we used the function described in Dehaene (2007 equation 10), which supposes a linear scaling in the ANS,

$$p_{correct}(n_{1,}n_2)=\int_0^{+\infty}\frac{e^{-\frac{1}{2}\left(\frac{x-(r-1)}{w\sqrt{1+r^2}}\right)^2}}{\sqrt{2\pi}w\sqrt{1+r^2}}dx$$

where n_1 and n_2 are the two numbers to be compared, r is the ratio of the larger and the smaller number, and w is the Weber ratio. According to the model this function should work with both symbolic and nonsymbolic comparison, although the Weber fraction could be different (Dehaene, 2007). In our analysis the error rates predicted by the specified function above were fit to the group mean of the error rates for both symbolic and nonsymbolic comparison for the whole stimulus space.

Reaction time. Current models are not straightforward about the reaction time prediction, and former descriptions (such as used in Moyer & Landauer, 1967) are incorrect from the viewpoint of the current models. Still, to test whether former pieces of evidence were used correctly to support the ANS model, we analyzed the reaction time data.

In the last decades the diffusion model (see the Drift rate section in the Analysis methods for details) became a successful and an increasingly popular tool to describe the reaction time of simple decision processes, including psychophysics comparison tasks. However, earlier works used some simpler models to describe the comparison tasks (Crossman, 1955; Moyer & Landauer, 1967; Welford, 1960). From the perspective of the diffusion models these early descriptions are incorrect, because, for example, they did not consider the Weber ratio of the processing system. Still, because evidence using these methods was considered to support the ANS model, in this detailed exploration we also investigate whether these historical tools can support the idea that the ANS processes both symbolic and nonsymbolic numbers.

In these early models, there was no clear consensus about the exact function that could describe the reaction time pattern. Psychophysics was more interested in error rates close to the threshold, and much less work investigated the reaction time far from the threshold (Crossman, 1955). For example, the seminal work by Moyer and Landauer (1967) used the *K × log (large_number/distance)* function[10], referring to the Welford (1960) paper, which in turn relied on Crossman (1955), however no straightforward solution was proposed then.

Although it is not easy to specify the function that was thought to describe correctly the reaction time pattern of comparison tasks, we can avoid this problem. First, as all models agree that dot comparison is handled by the ANS, dot comparison can be considered as the empirical specification of the required function. Second, in the early models, the specific functions could be fitted linearly to the reaction time: the model can be multiplied by a parameter to fit to the time scale of the comparison process, and a parameter can be added to account for the non-decision time. Moyer and Landauer (1967) also used this method implicitly: they reported Pearson product-moment correlation coefficient between the model and the data, which relies on simple linear regression. The linear transformation between the functions and the data means that the measured patterns should be linear transformations of other measured patterns, too. To summarize, according to the analysis methods of early works, the reaction time patterns of different notations are linear transformations of each other. To test this supposition, we fit the dot comparison reaction time pattern to the Indo-Arabic reaction time pattern. Because both dot and Indo-Arabic comparison data include noise, R^2 is not a suitable index to evaluate the similarity of the patterns. However, looking at the residuals can be more informative: if the two patterns readily fit, then only random noise is

10 In the Moyer and Landauer (1967) paper the *K log (large_number / large_number - small_number)* function can be found, without the necessary brackets around the large-small term, but most probably the calculation was performed with the correct function.

expected in the residuals. If, on the other hand, the two patterns differ in shape, then the residuals should show a systematic pattern.

It could be possible to have a more appropriate reaction time pattern with applying the diffusion models (see the next part for details), however, to our knowledge there is no clear consensus among others about the functional relationship between the drift rate and the representational overlap, consequently, the reaction time performance could not be specified easily.

Because the reaction time analysis applied here follows the reasoning of the early analysis, the current results cannot be considered as a reliable test of the ANS model, but we examine whether evidence offered formerly really support the common mechanism for symbolic and nonsymbolic number processing.

Drift rate. In the recent decades, the diffusion model and related models became increasingly popular to describe simple decision processes (Ratcliff & McKoon, 2008; Smith & Ratcliff, 2004). These models can recover background parameters directing both error rates and reaction times more sensitively. In the diffusion model, decision is based on a gradual accumulation of evidence offered by perceptual and other systems. Decision is made when appropriate amount of evidence is accumulated. Reaction time and error rates partly depend on the quality of the information (termed the drift rate) upon which the evidence is built. Larger drift rate usually results in faster and less erroneous responses. Drift rates are more informative than the error rate or reaction time in themselves, because drift rates reveal the sensitivity of the background mechanisms more directly (Wagenmakers, van der Maas, & Grasman, 2007). Importantly for our analysis, observed reaction time and error rate parameters can be used to recover the drift rates (Ratcliff & Tuerlinckx, 2002; Wagenmakers et al., 2007). The drift rates recovered from the behavioral data then can be used to investigate whether they are in line with the prediction of the ANS model.

In the ANS model, like in the case of the error rates, difficulty of the comparison of two properties might depend on the overlap of the two Gaussian random variables: larger overlap leads to worse performance (see the detailed mathematical description in Dehaene, 2007). In the diffusion model framework it is supposed that in a comparison task the drift rate depends purely on the overlap of the two random variables (Dehaene, 2007; Palmer, Huk, & Shadlen, 2005)[11].

To recover the drift rates for all number pairs in the two notations, the EZ diffusion model was applied, which can be used when the number of trials per cells is relatively small (Wagenmakers et al., 2007). Although this method has several limitations compared to more complex methods (Ratcliff & Tuerlinckx, 2002), (a) all other methods have different limitations, (b) according to

11 According to the current models, it is only the drift rate that is relevant in comparison performance (Dehaene, 2007; Palmer, Huk, & Shadlen, 2005). For example, non-decision time is not relevant in the distance effect, because it is not related to the comparison phase (Dehaene, 1996). Similarly, decision threshold is believed to be mainly modulated by the speed-precision instruction (Smith & Ratcliff, 2004) and not by the properties of the stimuli of specific trials. Although it is rare that other than the mean of the performance is investigated in a study, Rouder and his colleagues (Rouder, Lu, Speckman, Sun, & Jiang, 2005) measured reaction time properties as parameters of a 3-parameter Weibull distribution. They found that distance effect modified the scale parameter, but not the shape or location parameters. While the relation of the diffusion model generated performance distribution and the Weibull distribution is not fully understood (Rouder et al., 2005), drift rate change of the diffusion model can result in scale parameter change but not in shape or location parameter changes in the 3-parameter Weibull distribution (Rouder et al., 2005), in line with the idea that it is the drift rate that is related to the numerical distance effect.

current models, the constrains applied in the EZ-diffusion model might not influence the recovered drift rates essentially (although many aspects of the diffusion models are not known yet), and (c) in another numerical task analysis it was found that other tested diffusion models reveal the same pattern as the EZ diffusion model analysis (Kamienkowski, Pashler, Dehaene, & Sigman, 2011). For edge correction we used the half trial solution (see the exact details about edge correction in Wagenmakers et al., 2007). The scaling within-trials variability of drift rate was set to 0.1 in line with the tradition of the diffusion analysis literature.

In the analysis we investigated (a) whether the recovered drift rates are proportional to task difficulty and whether drift rates tend to 0 as the task difficulty increases, and (b) whether drift rates depend purely on the supposed representational overlap, as supposed by the ANS model. As in the case of the error rates, according to the ANS model, these properties should be present in both symbolic and nonsymbolic comparisons (Dehaene, 2007).

Results and discussion

Mean error rates and mean reaction times

Mean error rates and mean reaction times for correct responses were calculated for all number pairs for all participants in the two notations, then mean values across participants were computed (Figure 30). In both notations distance and size effects are visible, the patterns of the two notations seem similar, and based on first visual inspection the patterns could be in line with both the ANS model and the DSS model predictions.

Error rates — Dots

	5	10	15	20	25	30	35	40	45
5		2%	0%	0%	1%	0%	0%	0%	0%
10	3%		5%	1%	2%	1%	1%	1%	0%
15	2%	8%		14%	4%	5%	1%	2%	1%
20	1%	1%	18%		18%	11%	2%	2%	3%
25	1%	1%	2%	19%		28%	13%	4%	4%
30	0%	0%	0%	8%	22%		28%	17%	10%
35	0%	1%	2%	4%	9%	25%		32%	18%
40	1%	1%	1%	2%	4%	15%	26%		35%
45	0%	1%	2%	1%	3%	7%	14%	25%	

Reaction time (ms) — Dots

	5	10	15	20	25	30	35	40	45
5		751	586	529	505	548	513	492	504
10	748		891	710	577	583	557	557	520
15	623	940		1161	768	659	677	600	598
20	555	680	969		1110	912	746	702	668
25	526	624	842	1007		1462	1065	803	750
30	527	594	677	852	1006		1214	1056	898
35	574	570	637	750	937	1209		1304	1063
40	514	547	621	676	774	925	1170		1357
45	503	546	558	674	717	904	1163	1425	

Error rates — Indo-Arabic

	1	2	3	4	5	6	7	8	9
1		3%	2%	0%	1%	1%	1%	0%	0%
2	1%		3%	0%	1%	0%	1%	2%	1%
3	3%	1%		3%	2%	2%	3%	1%	1%
4	0%	1%	2%		4%	2%	4%	2%	0%
5	0%	1%	2%	5%		3%	5%	0%	1%
6	0%	1%	2%	2%	2%		8%	4%	1%
7	1%	2%	1%	2%	4%	10%		5%	9%
8	1%	1%	0%	0%	2%	2%	6%		11%
9	2%	1%	1%	0%	2%	2%	5%	2%	

Reaction time (ms) — Indo-Arabic

	1	2	3	4	5	6	7	8	9
1		576	576	567	547	533	533	527	536
2	598		641	601	597	563	590	552	602
3	569	633		665	632	609	608	596	635
4	586	598	737		671	626	608	573	559
5	524	600	668	680		662	640	624	631
6	556	599	604	612	662		721	654	627
7	537	546	570	599	678	730		660	625
8	521	547	603	579	617	668	671		687
9	532	548	546	569	585	616	632	652	

Figure 30 Error rates (left side) and reaction time (right side) in the whole stimulus space in dots (top) and Indo-Arabic (bottom) notations.

Two Weber ratios

The error rate results also revealed that the dot comparison is more erroneous than the Indo-Arabic comparison (the mean of the cells are 6.7% for dot notation and 2.0% for Indo-Arabic notation). On one hand, this result is hardly surprising: even common sense would suggest that the exact symbolic comparison is more precise than an imprecise dot array estimation. On the other hand it raises some nontrivial questions. If both types of comparisons are supported by the same representation, how is it possible that the two types of comparisons show radically different error rates and reaction times?

Because the ANS model suggests that the underlying representation works according to Weber's law, a reasonable idea is that the two notations are supported by different Weber ratios: for the Indo-Arabic comparison a more precise, low value is used, while for the dot array comparison a more imprecise, high value is applied. Dehaene (2007) also suggests that the different Weber ratios can be implemented in different neural cells, similar to the simulation in a connectionist model (Verguts & Fias, 2004). In this connectionist model an ANS-like layer represents the values, which layer works according to Weber's law, and after introducing symbolic notation to the network, the nodes of the number layer become more precise. While this explanation about the two Weber ratios seems compelling, there are some problems that are not trivial to solve. (1) Even if the Weber ratio is relatively small, soon it will reach a ratio in which the noise and the error rates will be too high to complete precise comparisons successfully. However, humans can compare numbers with any precision, which would require an unreasonably small Weber ratio. If one argues that there should be a supplementary mechanism that could help with the very small ratio number pairs, then why is its contribution practically invisible as suggested by the ANS model implicitly (i.e., if the Indo-Arabic comparison performance can be predicted precisely by the ANS model, then no other mechanism should have a major contribution to the measured performance)? (2) Actually, as already discussed in the Introduction, the Verguts model cannot be considered as an ANS model, because after introducing the symbolic numbers, the number layer cannot produce the size effect, violating the ratio-based performance which is a defining feature of the ANS model (Verguts & Fias, 2004), and only the addition of number frequency could restore the size effect in the model (Verguts et al., 2005), thus, the model cannot work according to Weber's law after the introduction of symbolic notation. Although none of those problems state that the ANS is incorrect, they indicate that some non-trivial problems should be solved to maintain its coherence.

Although we have not been able to find convincing answers to the questions mentioned so far, in the rest of our analysis we still suppose that the two Weber ratios model is correct, and investigate whether the ANS model with two ratios can explain the Indo-Arabic and dots comparisons equally well. This supposition is in line with the different mean error rate of the two notations, and it reflects the views of the proposers of the ANS model (e.g., Dehaene, 2007; Piazza et al., 2004).

ANS predictions for the error rates

In the present section we investigate whether the ANS model predicts the error rate patterns in both notations equally well. We calculated the error rate prediction pattern in our stimulus space for several Weber ratios. Two examples can be seen in Figure 31. Weber ratios between 0.05 and 0.25 with a step size of 0.02 were calculated, and fit of the models were calculated for all Weber ratios and for both dot comparison and Indo-Arabic comparison. Figure 32 shows the R^2 values (right y axes) for the dot comparison and the Indo-Arabic comparison as a function of the Weber ratio (x axis).

First, it is important to clarify that the overall R^2 value difference between the two notations is not appropriate to evaluate the ANS model. While the dot comparison reaches its R^2 maximum at around 0.95, the Indo-Arabic comparison R^2 is not higher than 0.6. The different maximum R^2 values can not only be the result of worse overall fit of the ANS model to the Indo-Arabic comparison, but it can also be the result of the smaller error rate in Indo-Arabic comparison. It is reasonable to suppose that the amount of noise is the same in both notations. However, because of the smaller error rate in Indo-Arabic comparison, the number pairs related variability is also smaller. Thus, the Indo-Arabic comparison has a lower signal-to-noise ratio. R^2 shows the percentage of the variance the model can explain of the data, but because of the lower signal-to-noise ratio, the percentage of the variance a perfect model could explain is smaller, thus, the maximum R^2 a perfect model could reach is also lower. Although the R^2 should be lower for a less appropriate model, here the variance of the R^2 is directed more strongly by the signal-to-noise ratio. This is another reason why the overall R^2 cannot be used to contrast the model's prediction in the two notations, but a more indirect analysis is required.

ANS model error rate prediction, Weber ratio = 0.19

	1	2	3	4	5	6	7	8	9
1		1%	0%	0%	0%	0%	0%	0%	0%
2	1%		7%	1%	0%	0%	0%	0%	0%
3	0%	7%		15%	4%	1%	0%	0%	0%
4	0%	1%	15%		21%	7%	3%	1%	0%
5	0%	0%	4%	21%		25%	11%	5%	2%
6	0%	0%	1%	7%	25%		28%	15%	7%
7	0%	0%	0%	3%	11%	28%		31%	18%
8	0%	0%	0%	1%	5%	15%	31%		33%
9	0%	0%	0%	0%	2%	7%	18%	33%	

ANS model error rate prediction, Weber ratio = 0.09

	1	2	3	4	5	6	7	8	9
1		0%	0%	0%	0%	0%	0%	0%	0%
2	0%		0%	0%	0%	0%	0%	0%	0%
3	0%	0%		1%	0%	0%	0%	0%	0%
4	0%	0%	1%		4%	0%	0%	0%	0%
5	0%	0%	0%	4%		8%	0%	0%	0%
6	0%	0%	0%	0%	8%		11%	1%	0%
7	0%	0%	0%	0%	0%	11%		15%	3%
8	0%	0%	0%	0%	0%	1%	15%		18%
9	0%	0%	0%	0%	0%	0%	3%	18%	

Figure 31 Error rate predictions of the ANS model in our full stimulus space for two Weber ratios. The Weber ratios were determined based on the mean error rates, see Figure 32 and the text.

Several properties of the ANS model are important, which properties can be used to assess how correct the model is for the two notations. These properties can also show why a more traditional model comparison method is not sufficient.

(1) Consistent predicted mean error rates and predicted performance patterns (R^2 values). Because the ANS model predicts the mean error rate directly, a model with appropriate Weber fraction should find the mean error rate of the measured performance. Additionally, because according to the ANS model the exact shape of the predicted performance (performance pattern) depends on the Weber-fraction of the representation[12], it also means that a linear fit of that prediction to the measured data should show the highest goodness of fit, when the model uses the appropriate Weber-fraction. Combining these statements, when the appropriate Weber-fraction is found, (a) the model should show the error rate prediction, and at the same time (b) it should show

12 In other words, ANS predicted performance patterns with different Weber-fraction, e.g., the two error rates shown in Figure 31, cannot be fitted perfectly with a linear transformation.

the highest goodness of fit (e.g., highest R^2 value) reflecting that the model finds the shape of the performance across the stimulus space.

To determine the Weber ratios for the two notations, we looked for the mean error rates of Weber ratios that are equal with the measured mean error rates of the two notations. Figure 32 shows the predicted mean error rate (left y axis) as a function of Weber ratios (x axis), and the measured Indo-Arabic and dot mean error rates (dashed horizontal lines). Intersections of the prediction (solid line with squares) and the measured data (dashed horizontal lines) specify the Weber ratios of the two notations. According to this, the Weber ratio of the dot comparison should be around 0.19, and the Weber ratio of the Indo-Arabic comparison should be around 0.09. The 0.19 value for nonsymbolic stimuli is indeed a typical Weber ratio according to former studies (see for example the results of an extensive measurement in Halberda & Odic, 2014; or the summary of Piazza, 2010 for a review about the development of the Weber ratio). One can note that in the measured data the large ratio cells (e.g., 2 vs 8, or 10 vs. 45) sometimes show a larger than 0% error rate (Figure 30), which is not in line with the prediction of the model (Figure 31), reflecting a base error rate, which is independent of the specific number pairs. Because the model cannot account for this error rate which is independent of the comparison stage, it could be more appropriate to subtract this base error rate (around 1%) from the measured error rate (lowering the horizontal dashed line on Figure 32). This correction would decrease the Weber ratios by a value around 0.02. All the following results are presented with the 0.19 and 0.09 Weber ratio values, although the same result patterns could be seen with the corrected 0.17 and 0.07 values, too.

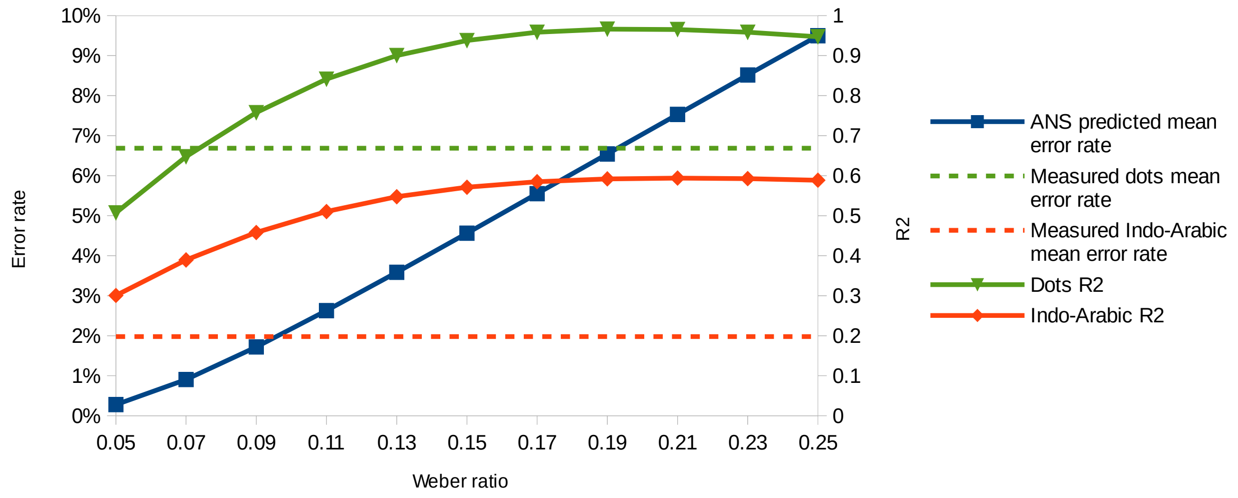


Figure 32 Predicted mean error rates (left y axis) as a function of Weber ratio, and measured mean error rates (left y axis) of the two notations. Goodness of fit (right y axis) as a function of Weber ratio for the dot comparison and the Indo-Arabic comparison.

After specifying the Weber ratios of the comparisons for the two notations, one can check if those Weber ratios also show the highest R^2 values. As discussed above, because the goodness of fit should be highest when the Weber ratio is specified correctly (i.e., the model should produce exactly the shape that was measured), the model predicts that the best fit (e.g., the highest R^2) can be obtained with the Weber ratio that is in line with the mean error rate of the notation. With all other Weber ratios the goodness of the fit should be worse. In the dot comparison task the R^2 indeed reaches its maximum around 0.19 Weber ratio, which Weber ratio was predicted based on the measured mean error rate. Thus, the ANS model predicts correctly that the Weber ratio of the best fitting pattern and the Weber ratio based on the mean error rates are approximately the same values. However, in the Indo-Arabic comparison the best R^2 value is around 0.2 Weber ratio, which is much larger than the 0.09 ratio specified with the mean error rate. This suggests that the ANS model cannot predict correctly the shape of the error rate pattern and the mean error rate at the same time in this symbolic comparison.

(2) Predicted error rate patterns. Based on the specified Weber ratios we can compare the predicted and the measured error rate patterns for the whole stimulus space, which can reveal further details how the ANS model prediction deviates from the measured symbolic comparison data. Figure 31 actually shows the predictions of the model for the Weber ratios with the identified dot and Indo-Arabic Weber ratios, thus, these patterns can be directly compared with the measured data (Figure 30). The difference of the measured and the predicted data can be seen in Figure 33. Because the model predicts directly the error rates, Figure 33 can be considered almost as the residuals after fitting the model to the measured data. Positive values show that the model underestimates the measured error rate, while negative values show that the model overestimates the actual error rate. In both notations the model and the actual data show systematic biases, however, they are qualitatively different in nature. (2a) In the dot comparison the misfit of the model is present because the measured data show an asymmetry related to the order of the stimuli, and the model cannot handle this asymmetry. In small ratio pairs large-small number pairs are responded to with smaller error rates (and faster, see Figure 30) than small-large number pairs. This effect can be the temporal congruity effect, in which large-small order pairs are handled faster than the small-large order pairs when the instruction is to choose the larger value (Schwarz & Stein, 1998). The effect may appear in our data if participants process the left stimulus first, which is consistent with the Western reading direction. The size of the temporal congruity effect is proportional to the difference of the onset of the two values, and disappears when the two stimuli are presented simultaneously (Schwarz & Stein, 1998). This latter property might explain why in our data the effect is only visible when the processing time is slow. It was proposed that the statistical feature of the data could be used to produce the effect: large numbers have higher probability to be the higher number in a pair, and according to this property, the decision criteria may be modified (Schwarz & Stein, 1998). Otherwise the prediction of the ANS model is relatively correct. (2b) On the other hand, residuals in the Indo-Arabic comparison show a completely different misfit. The model supposes that the error rate is very low for most of the number pairs, and error rate increases steeply for small ratio numbers. Instead of this pattern, measured error rates show that the small ratio number pairs do not show such a high error rate, and error rate starts to increase with larger distance in contrast with the model's prediction. These differences can be seen on the residuals as large overestimation for small ratios, and medium underestimation for medium ratios by the model. (These patterns remain if one would use the base error rate corrected 0.17 and 0.07 Weber ratios, although overall the models would underestimate the measured errors.) These observations suggest that while the ANS model predicts the ratio-based comparison error rates relatively correctly (except the order-based preference for the large-small stimuli in low ratio pairs, which asymmetric effect could be an additional effect), the model cannot describe appropriately the Indo-Arabic comparison error rate pattern.

Dot error rate – ANS with 0.19 Weber ratio

	1	2	3	4	5	6	7	8	9
1		1%	0%	0%	0%	0%	0%	0%	0%
2	2%		-2%	0%	2%	0%	0%	0%	0%
3	2%	1%		-1%	0%	4%	1%	2%	0%
4	0%	0%	3%		-3%	3%	-1%	1%	2%
5	0%	1%	-2%	-2%		2%	1%	-1%	2%
6	0%	0%	-1%	0%	-4%		-1%	2%	3%
7	0%	0%	1%	1%	-2%	-3%		1%	0%
8	1%	0%	1%	1%	-1%	0%	-5%		2%
9	0%	1%	1%	1%	1%	0%	-4%	-8%	

Indo-Arabic error rate – ANS with 0.09 Weber ratio

	1	2	3	4	5	6	7	8	9
1		3%	1%	0%	1%	1%	1%	0%	0%
2	0%		3%	0%	0%	0%	0%	1%	0%
3	2%	0%		1%	2%	2%	3%	1%	1%
4	0%	1%	0%		-1%	2%	3%	1%	0%
5	0%	0%	2%	1%		-5%	4%	0%	0%
6	0%	0%	2%	2%	-6%		-3%	3%	1%
7	1%	1%	1%	2%	4%	-1%		-10%	6%
8	1%	0%	0%	0%	2%	1%	-9%		-7%
9	1%	1%	0%	0%	1%	2%	2%	-16%	

Figure 33 Difference of the measured and predicted error rates for dot comparison (left) and Indo-Arabic comparison (right). Positive values show underestimation of the error rates by the model, negative values show overestimation.

(3) Linear regression parameters of the model. The found parameters of the fitting procedure shed additional light on how the ANS model fails to explain symbolic comparison data. The ANS error function predicts the error rate directly, therefore, with the appropriate Weber ratio the equation of the fit should be *measured_error = 1 × predicted_error + 0*. How do the parameters change across different Weber values? In the dot comparison task, for example for an incorrectly small 0.07 Weber ratio the fitted function is *2.83 × model + 0.04*. This high slope is reasonable, because the small Weber ratio predicts too small error rates that should be increased to fit the measured data. For larger Weber ratio the slope gradually decreases, and with the 0.19 Weber ratio (that was specified with the mean error rate) the function is *0.91 × model + 0.01*, in which the slope is rather close to the expected 1 value that the ANS predicts. In the Indo-Arabic comparison for a 0.07 Weber ratio the estimated function is *0.56 × model + 0.01*, which is decreasing further as the Weber increases, and for 0.09 Weber ratio the function is *0.37 × model + 0.01*. These much lower than 1 slopes reflect that the model predicts too sudden increase with small ratios (as observed in the direct comparison of the measured data and the model), and the fit is better when the model is flattened. Again, linear fit of the different Weber ratio models shows that while the ANS predicts correctly the dot comparison error rates, the model cannot predict the Indo-Arabic comparison.

To summarize, in a more extensive analysis, we found that on one hand the ANS model's prediction is coherent in the dot condition: a 0.19 Weber ratio correctly predicts the mean error rate, the relative shape of the error rates and the specific error rates for the number pairs. On the other hand, in the Indo-Arabic comparison the ANS model predicts a too steeply increasing error rate for small ratios, reflected in incoherent fit results. Again, the ANS model proposes that beyond the Weber fraction differences between the two notations, the same error function should hold for both notations (Dehaene, 2007), therefore, the lack of the precise ANS model description of the symbolic comparison is not the consequence of the notations specific processes. Thus, these results contradict the ANS model in its current form that suggests that both symbolic and nonsymbolic comparisons are handled by the same type of representations.

Linear similarity of the reaction time patterns

Group mean of dot comparison time for the whole stimulus space was fit to the group mean of Indo-Arabic comparison time for the whole stimulus space (right of Figure 30) According to the result, *Indo-Arabic_RT* = 0.17 × *dot_RT* + 474.8, $R^2 = 0.684$. Residuals of the fit (Figure 34) show an observable systematic pattern. The fitted dot data underestimate Indo-Arabic reaction time for small distance pairs, and overestimates it for large distance pairs. Additionally, the fitted dot data overestimate the cells with 1 and 9 values, similar to an end effect (see Figure 29). To test the presence of these effects in the residuals, multiple linear regression was used with linear distance effect and end effect regressors (see Figure 29), and the residual pattern was used as the dependent variable. Only the end effect regressor was significant (slope is 22.3, $p = 0.002$), while the distance effect was not (slope is 1.3, $p = 0.452$). The statistical lack of the distance effect contradicts the observable pattern, although visual inspection could be unreliable. One source of this contradiction could be the insufficient signal-to-noise ratio, and outliers might decrease the statistical power. After excluding two outlier cells (4-3 and 5-6) the correlation between the linear distance effect and the residuals when both numbers are in the 2-8 range (i.e., without the end effect cells) becomes significant, $r(38) = 0.28$, $p = 0.015$.[13] Thus, because of the observed systematic patterns in the residuals, the reaction time pattern of the dot and Indo-Arabic comparisons cannot be transformed to the other linearly, contrary to the former descriptions.

	1	2	3	4	5	6	7	8	9
1		25	-3	-3	13	35	28	31	24
2	3		-16	-6	-25	10	-21	17	-39
3	11	0		5	-27	-23	-19	-20	-59
4	-18	-8	-99		-9	3	-7	20	29
5	39	-20	-51	-36		59	15	-14	-30
6	8	-24	-15	7	-17		-41	-1	-1
7	35	24	12	2	-45	-51		35	29
8	40	20	-23	10	-12	-37	1		17
9	27	19	23	20	11	12	39	63	

Figure 34 Residuals after fitting dot comparison reaction time to the Indo-Arabic reaction time. Positive values denote higher fitted dot reaction time, negative values denote higher Indo-Arabic reaction time.

Although, as we have discussed, this analysis cannot be considered as a sufficiently precise method, it can be used to judge whether this type of reasoning has been cited correctly to support the common mechanism behind symbolic and nonsymbolic number processing. Our results suggest again that this test cannot confirm that nonsymbolic and symbolic numbers are processed by the same system.

13 One might suggest that the apparent distance effect in the residuals could be the artifact of fitting the dot data to the Indo-Arabic data with the end effect in the Indo-Arabic notation, and with the lack of the effect in the dot notation: because there is a stepwise change at the edge of the Indo-Arabic stimulus space, the “outer end” of the fitted distance effect will be lowered, creating a higher slope in the fitted line and a gradually increasing effect in the residuals (a distance like effect). However, such an artifact should underestimate large distance cells, while our data show an overestimation for those cells. Therefore, the distance effect in the residuals cannot be the artifact of the end effect in the Indo-Arabic notation.

Diffusion model analysis

The diffusion model analysis can be more sensitive than the error rate analysis, and more appropriate than the reaction time analysis by present-day standards. Drift rates for all number pairs and participants were calculated in both notations. The mean drift rates of the participants for the full stimulus space in the two notations are displayed in Figure 35. At first sight it is observable that drift rates show the distance and the size effects in both notations, and the dot comparison is harder than the Indo-Arabic comparison (dot drift rates are smaller), in line with the error rate and the reaction time data.

Dots

	5	10	15	20	25	30	35	40	45
5		0.2119	0.2761	0.3205	0.336	0.3366	0.3245	0.3477	0.3381
10	0.1859		0.1403	0.2143	0.2837	0.2843	0.2952	0.2997	0.3337
15	0.2375	0.1364		0.1068	0.1901	0.2229	0.225	0.2769	0.2563
20	0.2888	0.2192	0.1027		0.0972	0.1324	0.195	0.2082	0.2318
25	0.2998	0.2558	0.1715	0.1065		0.0488	0.1207	0.1699	0.1904
30	0.3063	0.259	0.2266	0.1592	0.0941		0.0354	0.113	0.1354
35	0.2706	0.2767	0.2291	0.1949	0.1426	0.0582		0.0354	0.0944
40	0.3154	0.2966	0.2338	0.2102	0.1749	0.1299	0.0552		0.0071
45	0.3434	0.2897	0.3024	0.2305	0.197	0.1409	0.1105	0.0632	

Indo-Arabic

	1	2	3	4	5	6	7	8	9
1		0.2851	0.2837	0.2929	0.3218	0.3535	0.3583	0.3662	0.3587
2	0.3005		0.2718	0.2976	0.3247	0.337	0.3246	0.3245	0.3206
3	0.2862	0.2866		0.2849	0.2648	0.2803	0.2957	0.31	0.2967
4	0.2965	0.2734	0.253		0.2422	0.2811	0.2753	0.3287	0.3507
5	0.3494	0.2964	0.2588	0.2296		0.2618	0.2419	0.3009	0.2479
6	0.3185	0.2788	0.2672	0.2767	0.254		0.1971	0.2585	0.2814
7	0.3349	0.3533	0.3051	0.2868	0.2281	0.2096		0.2256	0.2367
8	0.3517	0.3235	0.2764	0.3159	0.2566	0.2492	0.2228		0.2189
9	0.3684	0.315	0.3354	0.3151	0.308	0.2933	0.2507	0.2931	

Figure 35 Drift rates in the full stimulus space in dot comparison (left) and in Indo-Arabic comparison (right). Smaller values mean more difficult task.

Drift rate and task difficulty

The values shown in Figure 35 are displayed in a different way in Figure 36. In Figure 36 drift rates are displayed as the function of the difficulty of the task for the two notations. According to the current theories, the observable function in Figure 36 could be proportional, *drift_rate* = *k* × *task_difficulty*, (Dehaene, 2007; Palmer et al., 2005), or it could also include a power term as a generalization, *drift_rate* = *k* × *task_difficulty*$^{\beta}$, although the exponent is often close to 1, thus the first, proportional model approximates the second, power model. In the ANS model, task difficulty is measured as stimulus strength, which is calculated with the *distance/large_number* function as suggested by Palmer et al. (2005) for psychophysics comparison.[14] There are different properties that should be seen on this figure for any tasks or for tasks solved by an analogue system. (1) Easier tasks should show higher drift rates, i.e., in Figure 36 larger values on the x axis should go with larger values on the y axis, showing a positive slope for the curves. This is the case in both notations. However, while in the dot comparison the task difficulty and the drift rate are related more strictly (showing relatively small variance or error around a presumed regression curve), the same relation in the Indo-Arabic notation is much more noisy. (This is not caused by the cells involved in the end effect in Indo-Arabic comparison: after removing those cells, the difference is still visible.) This result is in line with a former study, finding that reaction time is better explained

14 Dehaene (2007) suggests that the difficulty of the task could be expressed as the logarithm of the ratios of the numbers, although that description is not entirely explicit how this function was found. One possibility is that this function was the one that could offer a linear relation between the difficulty of the task and the drift rates presented in that description. We also tested our data with the log(ratio) task difficulty scale, and the results could not be described neither with the proportional model (the curve is clearly non-linear), nor with the power model (the model strongly overestimates the drift rates for the easy tasks). However, Dehaene (2007) (a) used a more restricted diffusion model parameter recovery method, than the EZ diffusion model (although in the same paper EZ diffusion model was also used, its detailed results were not reported), and (b) he analyzed multi-digit number comparison. These differences can explain why a different expression was found as the measure of the task difficulty.

by the ratio in dot comparison task than in Indo-Arabic comparison task (Lyons, Nuerk, & Ansari, 2015 p. 1027). This might reflect that while the *distance/large_number* expression suggested by the ANS model might describe the difficulty of the dot comparison relatively well, it might not be applied readily for the Indo-Arabic notation. (2) In an analogue representation when the two signals almost completely overlap (i.e., two almost equal properties are shown) the system is hardly able to compare the two properties, which should result in a close to 0 drift rate in the diffusion model (i.e., no evidence is offered for the decision). On Figure 36 the difficulty is measured as *distance/large_number*, and an indistinguishable pair has a *0/large_number* value, which is 0. Thus, when difficulty tends to zero, drift rate should tend to zero, too, therefore, the intercept of the curves should be zero (Dehaene, 2007; Palmer et al., 2005). This is the case in the dot comparison condition, but Indo-Arabic comparison clearly shows a much higher intercept, somewhere around the 0.2 drift rate. This 0.2 intercept is in line with another single digit Indo-Arabic comparison task (Krajcsi et al., 2016), and with the non-zero intercept in multi-digit Indo-Arabic comparison (Dehaene, 2007). Again, these results show that while the dot comparison works according to the ANS model, the Indo-Arabic comparison follows other rules.

The 0 intercept of the dot comparison task also confirms that the use of the EZ diffusion model is at least partly appropriate, because its result correctly reflects an important property of an analogue mechanism, therefore validating the EZ method.

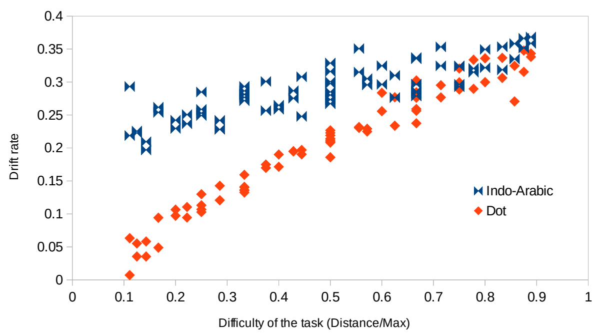


Figure 36 Drift rates of the number pairs as a function of the task difficulty in the two notations

Dehaene (2007) analyzed a similar data of an Indo-Arabic multi-digit comparison task, and he also found that the intercept of the drift rate function is larger than zero. We note that a multi-digit symbolic comparison might be a multi-step processing (Hinrichs, Berie, & Mosell, 1982; Krajcsi & Szabó, 2012; Poltrock & Schwartz, 1984), while diffusion model analysis is appropriate only for short, one cycle processing tasks (Wagenmakers et al., 2007), thus, the diffusion model analysis of multi-digit symbolic numbers should be handled cautiously. Still, independent of this problem, it is important to see how these results, which seemingly contradict the ANS model, could be interpreted to support the classic view. To explain the results in the ANS framework, Dehaene (2007) suggested that there could be two subsystems with two different Weber ratios working in a parallel way, and the interaction of these two subsystems could form the higher than zero intercept and the low slope for the Indo-Arabic number comparison. No further explanation was offered how the two subsystems could form this curve. We think that this two subsystems explanation raises some critical issues. First, it is hard to find why the interaction of two systems will produce high drift rate (and high intercept), when both systems can offer only low drift rates, if the stimuli are almost the same. One reasonable combination of the two drift rates could be the addition of the two values, but adding two small values, that are close to zero (as supposed by the ANS model), cannot result in a relatively high 0.2 value. As a more conceptual phrasing, if none of the two subsystems can differentiate between very small differences, why should any combinations of those analogue systems perform much better? Another reasonable combination of the two drift rates is that the higher drift rate should be applied, because the less precise subsystem cannot add any extra information to the already more precise subsystem. Again, it is still not clear how the intercept could increase radically. Another problem with this ANS explanation comes from the low slope of the Indo-Arabic drift rate curve. Dehaene (2007) suggests that in the linear model (*drift_rate* = *k* × *task_difficulty*) *k* is related to the Weber ratio: smaller Weber ratio (higher sensitivity) causes higher slope. Indeed, in the linear model the Weber ratio can be present only in that parameter. Now if we have a k_{dot} slope observed in the dot comparison task, the $k_{Indo\text{-}Arabic}$ slope in the more sensitive Indo-Arabic subsystem should be higher. If those parameters are combined, then again one option is to add the slopes, or another option is to use the larger slope. Both options predict a slope that is larger than the k_{dot} , however, the result shows a smaller value. In a more conceptual rephrase of this problem, the lower slope of the Indo-Arabic comparison suggests a higher (less sensitive) Weber ratio, which contradict the idea that the Indo-Arabic comparison must be more sensitive than the dot comparison. Overall, we cannot see how the ANS model could explain a drift rate curve with high intercept and low slope, and we propose that the analysis of the Indo-Arabic comparison drift rate data as a function of task difficulty is not in line with the ANS or any other representation working according to Weber's law.

Drift rate and representational overlap

While in the previous analysis the task difficulty was expressed by the relation of the two numbers, one can also incorporate the Weber ratio. The overlap of the representations of the two numbers can be calculated, that depends on the two values and the Weber ratio. The ANS model has another prediction that can be tested here: according to the model, the representational overlap predicts the drift rates in a comparison task. In contrast with the previous task difficulty vs. drift rate analysis, this relation of the drift rates and representational overlap is independent of the notation, because the different Weber ratios of the two notations are already incorporated in the overlap values.

To test whether drift rates depend purely on the representational overlap we calculated the representational overlap for all number pairs in our stimulus space for the two Weber ratios specified earlier. To calculate the overlap of two numbers, two Gaussian distributions were created on a linear scale, with the mean of the two numbers to be compared, and standard deviation was the product of the numbers and the Weber ratio (Halberda & Odic, 2014). Representational overlap values can be seen in Figure 37.

Representational overlap, Weber ratio is 0.19

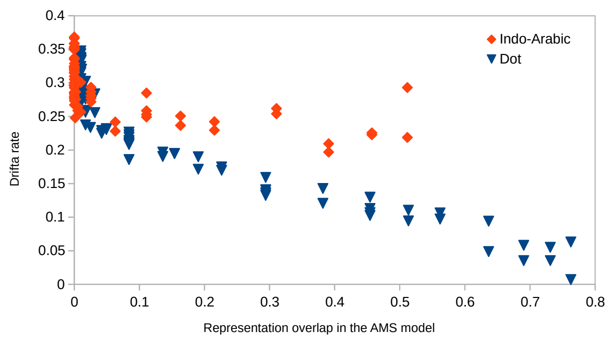
	1	2	3	4	5	6	7	8	9
1		0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.04		0.14	0.04	0.01	0.00	0.00	0.00	0.00
3	0.00	0.14		0.22	0.09	0.04	0.02	0.01	0.00
4	0.00	0.04	0.22		0.28	0.14	0.07	0.04	0.02
5	0.00	0.01	0.09	0.28		0.31	0.19	0.11	0.06
6	0.00	0.00	0.04	0.14	0.31		0.34	0.22	0.14
7	0.00	0.00	0.02	0.07	0.19	0.34		0.36	0.25
8	0.00	0.00	0.01	0.04	0.11	0.22	0.36		0.38
9	0.00	0.00	0.00	0.02	0.06	0.14	0.25	0.38	

Representational overlap, Weber ratio is 0.09

	1	2	3	4	5	6	7	8	9
1		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00		0.01	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.01		0.06	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.06		0.11	0.01	0.00	0.00	0.00
5	0.00	0.00	0.00	0.11		0.16	0.03	0.01	0.00
6	0.00	0.00	0.00	0.01	0.16		0.20	0.06	0.01
7	0.00	0.00	0.00	0.00	0.03	0.20		0.23	0.08
8	0.00	0.00	0.00	0.00	0.01	0.06	0.23		0.26
9	0.00	0.00	0.00	0.00	0.00	0.01	0.08	0.26	

Figure 37 Representational overlap in our stimulus space predicted by the ANS model for Weber ratios 0.19 and 0.9.

Left side of Figure 38 shows the drift rates as a function of representational overlap in the two notations. In the data for small overlaps the signs of the two notations largely overlap, and to show the potentially hidden dot data, dot data are shifted to the right by 0.01. Also, because the data are hard to explore for small overlap values, the same plot is displayed on a log overlap scale on the right of Figure 38. The dot data are not shifted on the latter plot.



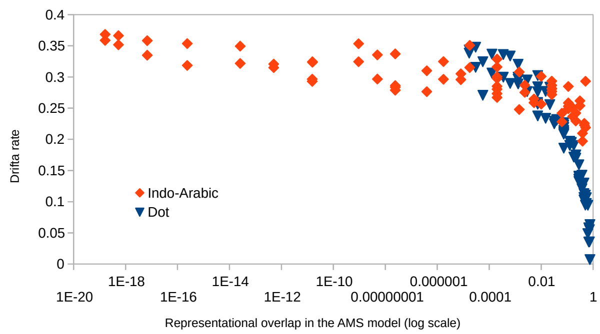


Figure 38 Drift rates as a function of representational overlap in the ANS model in the two notations. Overlap is displayed on linear (left) and logarithmic (right) scale. On the left plot, dot data are shifted by 0.01 to the right not to be covered by the Indo-Arabic data.

According to the ANS model same representational overlap values should result in same drift rate values, independent of the Weber ratio. While for small overlap values the drift rates of the two notations vary in the same range in line with the ANS prediction, for large overlap values Indo-Arabic drift rates are higher than the appropriate dot drift rates, contradicting the ANS model. (This is not caused by the end effect in Indo-Arabic notation: most of the high drift rate values in the large overlap range are not involved in the end effect. Additionally, the same pattern can be seen with the 0.17 and 0.07 Weber rates which are based on the corrected base error rate.) These data, again, show that the ANS model cannot describe the appropriate representations for both notations.

We also note that while there could be uncertainties whether EZ-diffusion model works correctly, in the current analysis all predictions of the ANS model in the dot comparison task proved to be correct, validating the EZ-diffusion model at the same time. This validation confirms that this simple to use diffusion parameter recovery method can be applied appropriately in the current comparison task.

General discussion

The present work investigated whether symbolic Indo-Arabic number comparison and nonsymbolic dot comparison can be described by the same model, as predicted by the widely accepted ANS model, or whether the two notations show systematic differences as suggested by the increasing body of evidence and some alternative accounts of symbolic number processing. Although formerly the ANS description for different notation comparisons has been tested, and the fit was found to be satisfactory, the similarity between the ANS and the recently proposed DSS model predictions required a more rigorous and extensive test.

Our results investigating several properties of the ANS model consistently showed that while the ANS model describe several behavioral aspects of the nonsymbolic dot comparison relatively well, the symbolic Indo-Arabic comparison deviated from the ANS description at several points. More specifically, (1) while the ANS model predicts the error rate pattern correctly and consistently for nonsymbolic dot comparison, it predicts too high error rates in Indo-Arabic comparison for the small ratio pairs, and too low error rates for medium ratio pairs. (2) The reaction time patterns of the two notations have different shapes which cannot be fitted linearly without systematic residuals, although early description of the comparison task reaction time would suggest a stricter similarity between the two patterns. (3a) In the diffusion model framework, while the dot drift rates are more clearly proportional to the difficulty of the task as defined in the ANS model, the relation between the Indo-Arabic drift rates and the ANS derived task difficulty is noisier. (3b) While the dot drift rates tend to zero when the number pairs become indistinguishable, the Indo-Arabic drift rates remain relatively high, contradicting the supposed functioning of a noisy analogue representation. (3c) Across the notations, the drift rates do not show the same values depending on the representational overlap as suggested by the ANS model, showing that the two notation comparisons cannot be described by the same mechanism. All of these results show that (a) nonsymbolic dot comparison and symbolic Indo-Arabic comparison do not rely purely on the same type of mechanism, and (b) while the ANS model can describe the nonsymbolic dot comparison, it cannot describe the symbolic Indo-Arabic notation.

One might wonder whether alternative forms of the ANS model could give an account for our findings, either by modifying the specific functions utilized in the present analyses or by conceptually modifying the model. At least one aspect of our results questions whether this is possible. In Indo-Arabic number comparison the drift rate does not tend to zero when the stimuli become almost indistinguishable, which result cannot be explained by any analogue representation working according to the Weber's law. This is an analogous form of the problem that it is difficult to explain how the imprecise ANS could be responsible for precise number processing. If the EZ diffusion model recovered appropriately the drift rates (we indeed found that many properties of the nonsymbolic drift rates are in line with the psychophysics model, which validates the EZ model), then the symbolic number comparison cannot be processed by any analogue representation working according to the Weber's law, which is a defining feature of the ANS model. Thus, we argue that the ANS model cannot be modified to account for the present findings.

One might also wonder whether shorter presentation of the dot stimuli could modify the results, because that could ensure that the diffusion model analysis handles a single step decision process instead of a multi-step counting process. However, the relatively precise prediction of the ANS model in dot comparison reflects that the current stimuli are successful enough to show the appropriateness of the ANS model, and further refinements can only improve this appropriateness. More generally, because the current design and stimuli were already appropriate to show that the ANS model describes nonsymbolic comparison correctly, there is no need to further improve the current methods using the nonsymbolic stimuli.

Beyond the current empirical results, suggesting that only nonsymbolic comparison seems to be supported by an analogue representation, but not symbolic comparison, we briefly summarize some non-trivial key problems of the ANS model explaining symbolic number processing. (1) As we have mentioned, how could an imprecise system, as the ANS, solve precise symbolic comparison? Even a smaller Weber ratio (more sensitive system) is inappropriate to solve this issue. (2) If a supplementary precise system helps to solve precise symbolic comparison, why is this system invisible in a sense that dominant part of the variance in the comparison performance is purely influenced by the ANS? Additionally, why is the ANS thought to dominantly influence performance in cases when it cannot solve the problem at all? (3) If the supplementary precise system has an effect on the performance, how do we know by looking at the performance that the ANS is also activated in a comparison task? If performance is partly comprised of a hypothetical precise system, then without specifying that precise component, one can not find the rest of the performance that could support the ANS processing either.

To summarize, all of our results show that symbolic and nonsymbolic comparisons show several critical differences, and while the ANS model can successfully describe the nonsymbolic dot comparison, it cannot account for many features of the symbolic Indo-Arabic comparison. Therefore, we argue that while nonsymbolic comparison is supported by the ANS, symbolic comparison and number processing is supported by an alternative system. Further research can confirm whether the increasing amount of data suggest correctly that symbolic and nonsymbolic numbers are processed by different types of systems, and if so, what representation is utilized to process symbolic numbers.

Acknowledgments

We thank Krisztián Kasos, Ákos Laczkó, and Katalin Oláh for their comments on an earlier version of the manuscript.

References

Balakrishnan, J. D., & Ashby, F. G. (1991). Is subitizing a unique numerical ability? *Perception and Psychophysics, 50,* 555–564.

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses. *NeuroImage*, *87*, 311–322. https://doi.org/10.1016/j.neuroimage.2013.10.049

Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2015). Visual Number Beats Abstract Numerical Magnitude: Format-dependent Representation of Arabic Digits and Dot Patterns in Human Parietal Cortex. *Journal of Cognitive Neuroscience*, *27*(7), 1376–1387. https://doi.org/10.1162/jocn_a_00787

Crossman, E. R. F. W. (1955). The measurement of discriminability. *Quarterly Journal of Experimental Psychology*, *7*(4), 176–195. https://doi.org/10.1080/17470215508416692

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557. https://doi.org/10.1073/pnas.1113195108

Damarla, S. R., & Just, M. A. (2013). Decoding the representation of numerical values from brain activation patterns. *Human Brain Mapping*, *34*(10), 2624–2634. https://doi.org/10.1002/hbm.22087

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (1996). The Organization of Brain Activations in Number Comparison: Event-Related Potentials and the Additive-Factors Method. *Journal of Cognitive Neuroscience*, *8*(1), 47–68. https://doi.org/10.1162/jocn.1996.8.1.47

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition* (Vol. XXII, pp. 527–574). Cambridge: Harvard University Press.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal representation in human intraparietal cortex. *Neuron*, *37*, 719–725.

Gebuis, T., & Reynvoet, B. (2012). Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology*, *49*(11), 1649–1659. https://doi.org/10.1111/j.1469-8986.2012.01461.x

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. In D. C. Geary, D. B. Berch, & K. Mann Koepke (Eds.), *Evolutionary Origins and Early Development of Number Processing* (pp. 305–333). Waltham, MA: Academic Press.

Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory & Cognition*, *10*(5), 487–495.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Huber, S., Nuerk, H.-C., Willmes, K., & Moeller, K. (2016). A general model framework for multisymbol number comparison. *Psychological Review*, *123*(6), 667–695. https://doi.org/10.1037/rev0000040

Kallai, A. Y., & Tzelgov, J. (2009). A generalized fraction: An entity smaller than one on the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1845–1864. https://doi.org/10.1037/a0016892

Kamienkowski, J. E., Pashler, H., Dehaene, S., & Sigman, M. (2011). Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making. *Cognition*, *119*(1), 81–95. https://doi.org/10.1016/j.cognition.2010.12.010

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*(4), 498–525.

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: a practical introduction*. London: Academic.

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., & Szabó, E. (2012). The role of number notation: sign-value notation number processing is easier than place-value. *Frontiers in Psychology*, *3*(463). https://doi.org/10.3389/fpsyg.2012.00463

Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology*, *60*(4), 227–234. https://doi.org/10.1027/1618-3169/a000191

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *70*(1), 12–23. https://doi.org/10.1037/cep0000070

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics, Herzliya, Israel*, 357–362.

Lyons, I. M., Ansari, D., & Beilock, S. L. (2015). Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain. *Human Brain Mapping*, *36*(2), 475–488. https://doi.org/10.1002/hbm.22641

Lyons, I. M., Nuerk, H.-C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General*, *144*(5), 1021–1035. https://doi.org/10.1037/xge0000094

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*, 177–190. https://doi.org/10.1038/nrn1626

Noël, M.-P., & Rousselle, L. (2011). Developmental Changes in the Profiles of Dyscalculia: An Explanation Based on a Double Exact-and-Approximate Number Representation Model. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00165

Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 1–1. https://doi.org/10.1167/5.5.1

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Patro, K., Nuerk, H.-C., Cress, U., & Haman, M. (2014). How number-space relationships are assessed before formal schooling: A taxonomy proposal. *Developmental Psychology*, *5*, 419. https://doi.org/10.3389/fpsyg.2014.00419

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Piazza, M., Mechielli, A., Butterworth, B., & Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *NeuroImage*, *15*, 435–446.

Pinhas, M., & Tzelgov, J. (2012). Expanding on the Mental Number Line: Zero Is Perceived as the “Smallest.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1187–1205. https://doi.org/10.1037/a0027390

Poltrock, S. E., & Schwartz, D. R. (1984). Comparative judgments of multidigit numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 32–45. https://doi.org/10.1037/0278-7393.10.1.32

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation*, *20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. https://doi.org/10.3758/BF03196302

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614.

Rouder, J. N., Lu, J., Speckman, P., Sun, D., & Jiang, Y. (2005). A hierarchical model for estimating response time distributions. *Psychonomic Bulletin & Review*, *12*(2), 195–223. https://doi.org/10.3758/BF03257252

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *The Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Sathian, K., Simon, T. J., Peterson, S., Patel, G. A., Hoffman, J. M., & Grafton, S. T. (1999). Neural evidence linking visual object enumeration and attention. *Journal of Cognitive Neuroscience*, *11*(1), 36–51.

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12372

Scholz, K. W., & Potts, G. R. (1974). Cognitive processing of linear orderings. *Journal of Experimental Psychology*, *102*(2), 323–326. https://doi.org/10.1037/h0035873

Schwarz, W., & Stein, F. (1998). On the temporal dynamics of digit comparison processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1275–1293. https://doi.org/10.1037/0278-7393.24.5.1275

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006

Stoianov, I., & Zorzi, M. (2012). Emergence of a “visual number sense” in hierarchical generative models. *Nature Neuroscience*, *15*(2), 194–196. https://doi.org/10.1038/nn.2996

Tzelgov, J., Ganor-Stern, D., & Maymon-Schreiber, K. (2009). The representation of negative numbers: exploring the effects of mode of processing and notation. *Quarterly Journal of Experimental Psychology*, *62*(3), 605–624. https://doi.org/10.1080/17470210802034751

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. https://doi.org/10.3758/BF03194023

Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years’ progress. *Ergonomics*, *3*, 189–230. https://doi.org/10.1080/00140136008930484

Chapter 6: Numerical distance and size effects dissociate in Indo-Arabic number comparison

The present chapter has been published:

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. Psychonomic Bulletin & Review. 24(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Numerical distance and size effects dissociate in Indo-Arabic number comparison

Attila Krajcsi
Eötvös Loránd University, Institute of Psychology, Department of Cognitive Psychology

Numerical distance and size effects (easier number comparison with large distance or small size) are mostly supposed to reflect a single effect, the ratio effect, which is the consequence of the analogue number system (ANS) activation, working according to Weber's law. In an alternative model, symbolic numbers can be processed by a discrete semantic system (DSS), in which the distance and the size effects could originate in two independent factors: the distance effect depends on the semantic distance of the units, and the size effect depends on the frequency of the symbols. While in the classic view both symbolic and nonsymbolic numbers are processed by the ANS, in the alternative view only nonsymbolic numbers are processed by the ANS, but symbolic numbers are handled by the DSS. The current work contrasts the two views, investigating whether the size of the distance and the size effects correlate in nonsymbolic dot comparison and in symbolic Indo-Arabic comparison tasks. If a comparison is backed by the ANS, the distance and the size effects should correlate, because the two effects are merely two ways to measure the same ratio effect, however, if a comparison is supported by other system, for example the DSS, the two effects might dissociate. In the current measurements the distance and the size effects correlated very strongly in the dot comparison task, but they did not correlate in the Indo-Arabic comparison task. Additionally, the effects did not correlate between the Indo-Arabic and the dot comparison tasks. These results suggest that symbolic number comparison is not handled by the ANS, but by an alternative representation, such as the DSS.

Keywords: numerical distance effect; numerical size effect; analogue number system; discrete semantic system;

Introduction

In their seminal work Moyer and Landauer (1967) described that in an Indo-Arabic single digit number comparison task the performance is worse (i.e., reaction time is slower and error rate is higher) when the difference between the two numbers is relatively small (numerical distance effect) or when the numbers are relatively large (numerical size effect). They proposed that these two effects are the reflection of a single effect based on the ratio of the numbers: number pairs with smaller ratio are harder to process (Figure 39). This ratio-based performance was thought to be the result of a simple representation working according to Weber's law, similar to the representations working behind simple physical feature comparison tasks. This Analogue Number System (ANS) is supposed to operate behind any number comparison, independent of the notation of the numbers (Dehaene, 1992; Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003). The numerical cognition literature dominantly accepts the ANS interpretation of the number comparison (Dehaene, 1992; Nieder, 2005; Piazza, 2010).

However, an alternative account can also explain the numerical distance and size effects in symbolic number processing. It might be rather intuitive to imagine that symbolic numbers are stored in a Discrete Semantic System (DSS), similar to the mental lexicon or a semantic network. In this system, numbers are represented by nodes, and the connections of the nodes reflect the semantic relations of the nodes. These relations might mainly be directed by the values of the numbers, but other properties might also have an effect, such as parity or primeness. The distance effect might be originated in the semantic relation of the nodes (Figure 39). As a similar distance based non-numerical example, in a picture naming task it was found that naming time was influenced by the previous picture, and the influence was proportional to the semantic distance of the priming and the target pictures (Vigliocco, Vinson, Damian, & Levelt, 2002). This semantic distance effect is similar to the numerical distance effect: in both tasks the performance was influenced by the semantic distance of the items. Although the cited distance effect in the picture naming task is a priming distance effect which is proposed to be different than a comparison distance effect in the number comparison task (Reynvoet, De Smedt, & Van den Bussche, 2009), the current example points out that a discrete representation can also produce a distance based gradual effect. The numerical size effect can also be explained by the DSS view. It is well known that the frequency of numbers is not uniform, but smaller numbers are more frequent than larger numbers (Dehaene & Mehler, 1992). It is also known that stimulus processing is influenced by the stimulus frequency. Based on these starting points one can imagine that larger numbers are harder to process, because they are less frequent (Figure 39). It was also shown that in an artificial number symbol comparison, in which the frequency of the digits can arbitrarily be manipulated, the size effect followed the frequency of the numbers (Krajcsi, Lengyel, & Kojouharova, submittedc), reflecting that the numerical size effect is indeed a frequency effect. Combination of the semantic based distance effect and the frequency based size effect can predict the performance seen in comparison tasks, and this prediction correlates strongly with the ANS model prediction, revealing similar descriptions of the comparison performance by the two models (Krajcsi et al., submittedc). To summarize, the DSS model can give simple explanations for the symbolic numerical distance and size effects (see additional examples how further symbolic numerical effects can be explained in Krajcsi, Lengyel, & Kojouharova, submittedb). It is important to note that the DSS only accounts for the symbolic number processing, while nonsymbolic number processing could still be supported by the ANS.

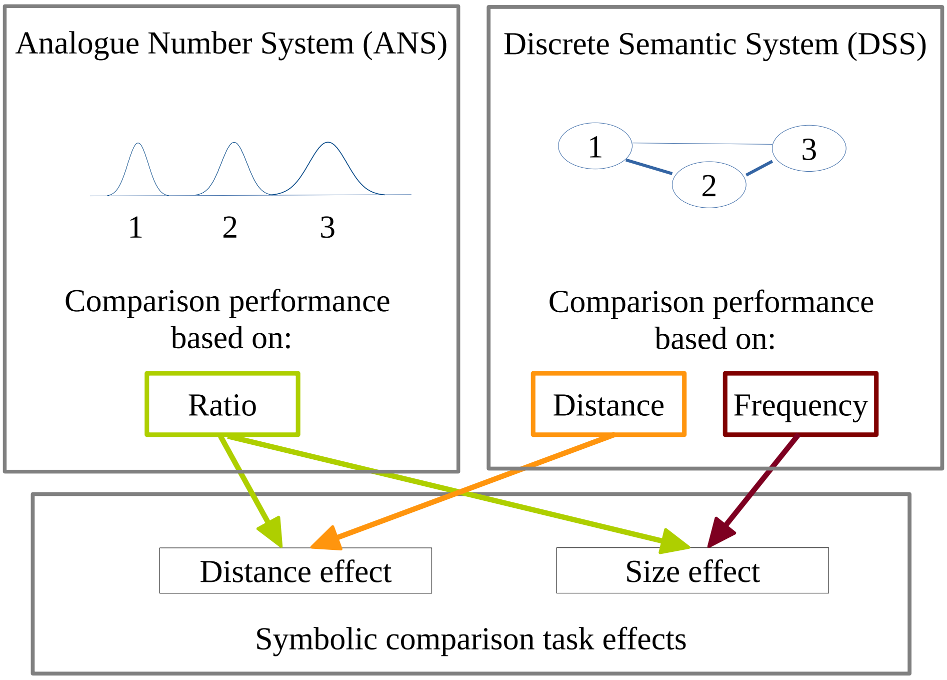


Figure 39. The sources of the symbolic distance and size effects according to the two models.

In the current study the two models are contrasted by correlating the size of the numerical distance and size effects. According to the ANS model, the distance and size effects are merely two ways to measure the same ratio effect, therefore the two effects should be closely related. To discuss this prediction in more detail, in the ANS model, the size of the distance and size effects are influenced by scaling parameters as in Moyer & Landauer (1967) and by the Weber ratio (Dehaene, 2007). Importantly, these parameters modify both the distance and size effects at the same time, because the equations describing the performance handle only the ratio, and the distance and size effects are the consequences of the ratio effect. Because of the way the size of the distance and size effects is calculated it is not trivial to specify whether the relation between the effects is linear, however, the relation should be at least monotonic. It can also be important to note that distance and size effects cannot correlate as an artifact of the way the two effects are calculated, because the distance effect relies on the difference of two numbers to be compared, while the size effect relies on the sum of the two numbers (see also the Results part), therefore the two effects form orthogonal dimensions in the stimulus space, resulting in 0 correlation between the distance and the size of number pairs. According to the DSS model, in symbolic comparison, the distance and the size effects derive from different mechanisms and potentially different parameters, consequently, the two effects could be partly unrelated. At the same time, according to the DSS model it is possible that some parameters are shared between the distance-based and the frequency-based mechanisms (either as DSS specific parameters, or as some general states as recently discussed in Cantlon, 2015), thus, some correlation could be observed. Therefore, according to the DSS model, correlation value might be found in a wide range, but because of the differing mechanisms behind distance and size effects, if the different mechanisms include different parameters that influence performance, then the correlation should be lower than the ANS predicted high correlation.

The main question of the current study is whether the distance and size effects strongly correlate in symbolic notation as predicted by the ANS model, or whether the correlation is smaller, as allowed by the DSS model. Correlation is also measured in nonsymbolic notation, because both the ANS and the DSS view suggest that nonsymbolic comparison is backed by the ANS, this nonsymbolic comparison correlation serves as a baseline. Statistically, the main question is whether the correlation of the distance and the size effects is smaller in symbolic Indo-Arabic comparison than in nonsymbolic dot comparison.

Methods

The current study reanalyzes the data of two former studies. One of them is a control group data of a neuropsychology study in preparation, and the other study investigates psychophysics properties of nonsymbolic and symbolic comparisons (Krajcsi, Lengyel, & Kojouharova, submitteda). Both the aim of those studies and the analysis are different from the current one. Two sets of data are used to ensure the reliability of the results with their replicability. The two studies mostly used the same methods, and only a few differences can be found.

Participants compared Indo-Arabic numbers in one condition, and they compared dot arrays in another condition.

In a trial two numbers were visible on the left and on the right sides of the screen, and participants had to choose the larger one by pressing one of the two response keys. The stimuli were visible until key press. The response was followed by an empty screen for 500 ms, then the next trial started.

In the Indo-Arabic condition the numbers were between 1 and 9, i.e., all single digit numbers. Processing multi-power numbers include additional mechanisms handling the powers (Hinrichs, Berie, & Mosell, 1982; Poltrock & Schwartz, 1984), therefore, it is more appropriate to use only single digit numbers. All possible pairings of those values were presented, except ties, resulting in 72 possible pairs. All pairs were presented 3 times (study 1) or 10 times (study 2), resulting in 216 or 720 trials, respectively. The order of the trials was randomized.

In the dots condition it is not appropriate to use the same 1-9 range as in the Indo-Arabic condition, because sets with less than 5 objects can be enumerated fast, which fast enumeration is termed subitizing (Kaufman, Lord, Reese, & Volkmann, 1949). Subitizing is not an ANS directed process (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008), but it is most probably based on pattern detection (Krajcsi, Szabó, & Mórocz, 2013; Mandler & Shebo, 1982). Therefore, to measure the ANS based dot estimation, the 1-4 range should be avoided. In order to avoid the 1-4 range, and to keep the critical ratio-based feature at the same time, all numbers between 1 and 9 were multiplied by 5, resulting in a number range between 5 and 45. According to the ANS model, because ratio is the only main source of the performance (Dehaene, 2007), transformations keeping the ratios should not change neither the performance in general, nor the correlations specifically. In an array of dots, black and white dots in random positions were shown against a gray background (Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011), thus, the luminance of the stimuli was not informative about the numerosity. Dots of an array were drawn randomly in a 2×2 degrees area, with a dot diameter of 0.2 degrees, therefore, density and convex hull correlated with the numerosity.

Although our stimuli do not control all perceptual features that might influence the perceived numerosity, perfect visual control might be impossible in some simple methods usually utilized in the literature (Gebuis & Reynvoet, 2012). More importantly, in the current test, nonnumerical influence of the decision process is less relevant, because the ANS model suggests that number comparison is handled by an analogue system that could be used in any continuous physical feature comparison (Dehaene, 2007; Moyer & Landauer, 1967), hence, in a general sense, any continuous physical feature comparison working according to the Weber's law could be an appropriate task in our test. Additionally, a mixture of visual ratio-based performance and numerosity ratio-based performance should also produce an approximately ratio-based performance, as reflected in the similar psychometric functions of visual comparison and numerical comparison tasks. Therefore, the simple and limited visual control of the stimuli should be appropriate for the aim of the current test. As in the Indo-Arabic condition, all possible pairs were presented 3 times (study 1) or 10 times (study 2), resulting in 216 or 720 trials, respectively. The order of the trials was randomized.

In the first study all sessions started with the Indo-Arabic condition, and finished with the dot condition. In the second study, the order of the conditions was counterbalanced across participants.

Presentation of the stimuli and the measurement of the responses were managed by the PsychoPy software (Peirce, 2007).

In the first study 19 university students participated for partial credit course. No participants were excluded from the analysis. In the sample there were 2 males, the age range was 18-24 years, with a mean of 20.1 years. In the second study 24 university students participated for partial credit course. Four participants were excluded, because their error rates were higher than 1.5 standard deviation + mean error rates at least in one of the conditions (6% in the Indo-Arabic condition and 15% in the dots condition). Among the remaining 20 participants there were 4 males, the age range was 19-24 years, with a mean of 21.0 years.

Analysis

The slopes of the distance and size effects based on reaction times were calculated for all participants and for both notations. For the slopes reported here mean reaction times were calculated for all distance values (absolute value of the difference of the two numbers) or for all size values (sum of the numbers) for all participants and notations, and linear regression slopes were calculated on these mean values. The effects were alternatively calculated (a) with only the correct responses, (b) using the median instead of the mean, or (c) with the slope divided by the mean of the comparison time to handle the slope change caused by general speed differences. These alternative calculation methods revealed the very same results as the first one (i.e., significant correlation in the nonsymbolic comparison, nonsignificant correlation in the symbolic comparison and significant difference between the symbolic and nonsymbolic correlations), thus, their results are not reported here. In the dot comparison task, for the distance effect the slope was calculated as if the number of dots were between 1 and 9, although the dots were between 5 and 45, but this method reflected more appropriately the ANS model driven consideration, that the two notations cover the same ratio range. Importantly, this linear transformation of the slopes do not change the correlational coefficients.

For the correlational analysis, (1) Pearson's product-moment coefficient was calculated, which makes the current result comparable with previous correlational studies. Additionally, (2) Spearman's rank correlation coefficient was also calculated, because in contrast with the Pearson correlation, (a) it is not sensitive to outliers, (b) to violation of normality, (c) to violation of homoscedasticity, (d) and it can measure nonlinear monotonic relation more sensitively. Finally, (3) reliability of the variables in the correlation was handled. In any correlation the variables include both the signals to be measured and the noise. The noise decreases the reliability of the variables, which in turn constrains the maximum correlation one might see. Because the ANS model predicts a high correlation between distance and size effects, it is essential to handle the potential unreliability of the variables to see if the two variables measure the same mechanism. Also, it is possible that there could be smaller distance effect-size effect correlation in the Indo-Arabic comparison, because it has a smaller signal-to-noise ratio (resulting in smaller reliability, and consequently lower correlation), and not because symbolic comparison is processed by the DSS, therefore, the critical difference of the correlations between the notations might be a bias of different signal-to-noise ratio in those notations. Spearman's method was applied to estimate the "real" correlation of the variables, removing the role of the unreliability (Spearman, 1904). Distance and size effects were calculated again, as for the main analysis, but similar to an even-odd split-half method, even and odd trials were handled separately, thus, for all effects even and odd versions were calculated. Corrected correlation was calculated as

$$\frac{r\left(distance_{even},\mathrm{size}_{even}\right)+r\left(distance_{even},\mathrm{size}_{odd}\right)+r\left(distance_{odd},\mathrm{size}_{even}\right)+r\left(distance_{odd},\mathrm{size}_{odd}\right)}{4\times\sqrt{r\left(distence_{even},distance_{odd}\right)\times r\left(\mathrm{size}_{even},\mathrm{size}_{odd}\right)}},$$

where r is the Pearson correlation coefficient.

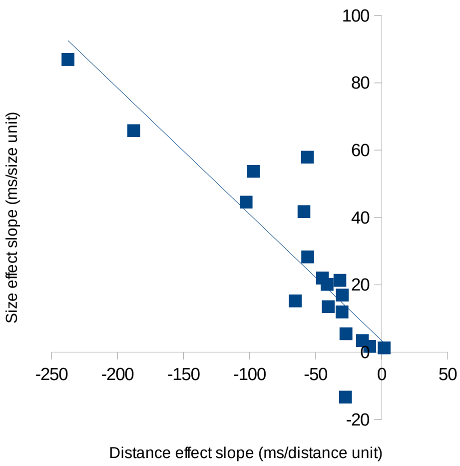
Results

In the main analysis the correlation of the slopes of the two effects were investigated in both notations and in both studies. In the dot comparison task the distance and size effects strongly correlated in both studies (Figure 40, left side). Critically, the estimated correlation coefficients removing the unreliabilities of the variables show values very close to the value one[15], reflecting a perfect correlation between distance and size effects in nonsymbolic dot comparison. In the Indo-Arabic comparison task the correlation is weak and not significant in any of the studies (Figure 40, right side). These results are not the artifact of outliers, non-normality of the variables or heteroscedasticity, because the very same results can be observed with both the Pearson's and Spearman's correlation coefficient. The difference between the correlations is significant in both studies.

15 It is possible to have a larger than 1 value for the corrected correlation coefficient. Still, the corrected correlation coefficient of the dot comparison effects in the second study (-1.06) is very close to the value 1.

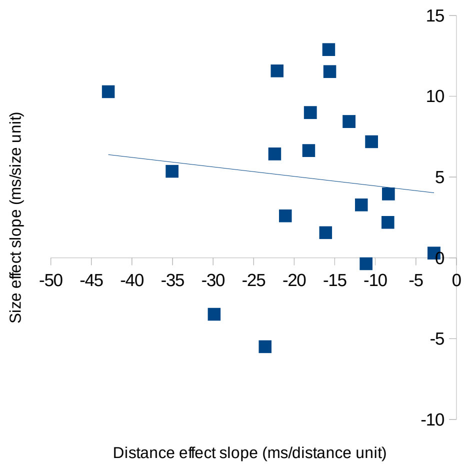
Study 1

Dot notation



(1) Pearson $r(17) = -0.88$, $p < 0.001$,
95% CI of r [-0.95, -0.70]
(2) Spearman $r_s(17) = -0.9$, $p < 0.001$,
95% CI of r_s [-0.96, -0.75]
(3) Corrected with reliability $r = -0.99$

Indo-Arabic notation

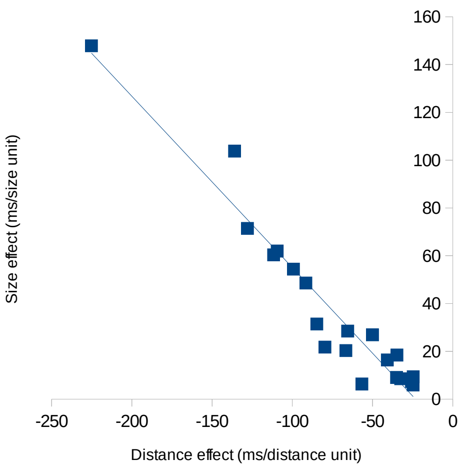


(1) Pearson $r(17) = -0.11$, $p = 0.65$,
95% CI of r [-0.54, 0.36]
(2) Spearman $r_s(17) = -0.11$, $p = 0.66$,
95% CI of r_s [-0.54, 0.36]
(3) Corrected with reliability $r = -0.22$

Difference between the Pearson correlations: $Z = 3.5$, $p < 0.001$

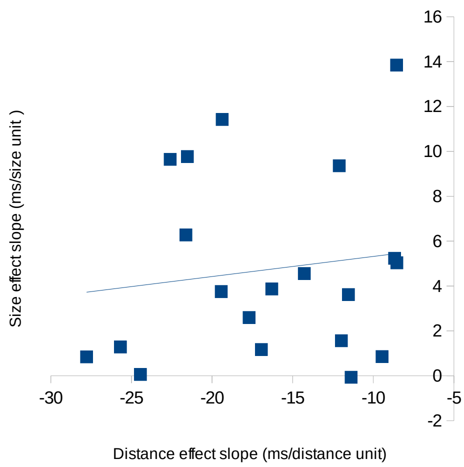
Study 2

Dot notation



(1) Pearson $r(18) = -0.96$, $p < 0.001$,
95% CI of r [-0.99, -0.91]
(2) Spearman $r_s(18) = -0.92$, $p < 0.001$,
95% CI of r_s [-0.97, -0.8]
(3) Corrected with reliability $r = -1.06$

Indo-Arabic notation



(1) Pearson $r(18) = -0.13$, $p = 0.57$,
95% CI of r [-0.33, 0.54]
(2) Spearman $r_s(18) = 0.14$, $p = 0.56$,
95% CI of r_s [-0.32, 0.55]
(3) Corrected with reliability $r = 0.26$

Difference between the Pearson correlations: Z = 6.53, p < 0.001

Figure 40. Relation of the distance and size effect slopes displayed on scatter plots and measured with correlation coefficients in dot comparison (left) and Indo-Arabic comparison (right)

According to some former reports, nonsymbolic dot comparison performance and the mainly symbolic mathematical performance are related (e.g., Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Halberda, Mazzocco, & Feigenson, 2008; Lourenco, Bonny, Fernandez, & Rao, 2012), therefore, it could be of interest how the symbolic and nonsymbolic comparisons are related in the current data. The correlations of the effects across the notations were calculated, e.g., whether Indo-Arabic distance effect and dot distance-effect correlate. None of the effects correlated across the notations: in the first study, for the distance effects $r(17) = 0.00$, $p = 1.00$, 95% CI of r [-0.45, 0.46], $r_s = 0.08$, for the size effects $r(17) = 0.13$, $p = 0.59$, 95% CI of r [-0.34, 0.55], $r_s = 0.11$, in the second study, for the distance effects $r(18) = 0.15$, $p = 0.52$, 95% CI of r [-0.31, 0.56], $r_s = 0.12$, for the size effects $r(18) = 0.32$, $p < 0.17$, 95% CI of r [-0.15, 0.67], $r_s = 0.24$.

Discussion

In the present work it was investigated how strongly the numerical distance and size effect slopes correlate in Indo-Arabic and dot comparison tasks. (1) It was found that distance and size effect slopes strongly correlate in the dot comparison task (Figure 40, left side), and after correcting for the reliabilities of the variables, the correlation is very close to the value of 1, reflecting a perfect connection between the two effects. This result is in line with the classic ANS model that suggests that in a comparison task the distance and the size effects are the direct consequences of the ratio effect (Moyer & Landauer, 1967), and since both distance and size effects are modified by the same parameters, the two effects should strongly correlate. Additionally, the very strong correlation demonstrates that the method used in the present study is appropriate to reveal this strong relation between the distance and the size effects. (2) It was also found that the distance and the size effects are barely related in the Indo-Arabic comparison task (Figure 40, right side), and the correlation is clearly smaller than the correlation found in the dot comparison task. This result is in conflict with the classic ANS model, because according to the ANS model the same strong correlation should have been found as in the dot comparison task, because the distance and size effects are directed by a single ratio-based effect. However, the result is in line with the DSS model, which suggests that the distance and the size effects rely on different mechanisms and probably partly on different parameters: the distance effect could be rooted in a mechanism based on the semantic relation of the units (as in Vigliocco et al., 2002), while the size effect might be related to the frequency of the symbols (Dehaene & Mehler, 1992; Krajcsi et al., submittedc). (3) Finally, it was found that the effects do not correlate between the notations: it seems that the size of the distance effect in the Indo-Arabic comparison is independent of the size of the distance effect in the dot comparison, and the same holds for the size effect. Although it was suggested that nonsymbolic performance correlates with mathematical achievement, suggesting that the ANS is one of the main root of mathematical knowledge (e.g., Halberda et al., 2012, 2008; Lourenco et al., 2012), the current data are not in line with these findings, but reflect the finding of a more systematic review, revealing that only symbolic comparison correlates with children's mathematical achievement, but nonsymbolic comparison and mathematical achievement mostly do not correlate (De Smedt, Noël, Gilmore, & Ansari, 2013). The current results also replicate the findings of studies which demonstrate that symbolic and nonsymbolic comparisons do not correlate in children (Holloway & Ansari, 2009; Sasanguie, Defever, Maertens, & Reynvoet, 2014), although in those reports in the nonsymbolic comparisons set sizes from the subitizing range (i.e., 1-4) were also applied, thus, the validity of some of those data could be questioned.

The current results are not in line with the ANS model, which model suggests that the same type of process handles both the symbolic and nonsymbolic numbers (Dehaene, 1992; Eger et al., 2003). Although there could be differences between the symbolic and nonsymbolic number processing (e.g., different Weber fractions across notations, see Dehaene, 2007), the ANS model unequivocally states that the performance is based on the ratio of the values, coming from a representation working according to Weber’s law, and this feature is independent of the notation of the values, let it be symbolic or nonsymbolic. Therefore, the current results cannot be explained by the fact that there are differences between symbolic and nonsymbolic comparisons, and the current data simply could be another example of those differences, because according to the ANS model the distance-size effects correlation should be observable in any number comparisons.

While the current results are in line with the DSS model, and give some support to this alternative view, this single study clearly does not test the DSS model extensively. It is also possible that another alternative could be found. Still, the DSS seems a reasonable alternative at the moment, and it can explain not only former data about symbolic number comparison, or number processing in more general, but it can also explain the current new findings. Further research should reveal whether the DSS model is an appropriate explanation for symbolic number processing or other alternative should be found.

Based on the current data, it can be possible that while comparing symbolic numbers, both the ANS and the DSS are activated, and both of them influence the observed performance. It may be possible that the correlation of the effects in symbolic comparison is smaller than in nonsymbolic comparison, because the contribution of the DSS decreases the coefficient, yet, part of the correlation still might be originated in the ANS processing. Although the very low coefficients seen in symbolic comparison hint a pure DSS processing, still, because of the uncertainties of the analysis methods, it is hard to precisely quantify the different predictions of a pure DSS activation and a mixed DSS and ANS activation explanations. Still, it is clear that the symbolic comparison results are not in line with a pure ANS explanation.

These results and interpretations are in accord with some former results of the literature. A few studies have shown that effects that should be related according to the ANS model are actually independent. In an artificial new number symbol system it was found that the size effect follows the frequencies of the symbols, independent of the ratio of the values (Krajcsi et al., submittedc). In a connectionist simulation, various symbolic numerical effects could be modeled coherently, and importantly, the size effect could be modeled with the introduction of the frequencies of the values, independent of the distance effect (Verguts, Fias, & Stevens, 2005). Similarly, in another connectionist model of number comparison, the distance and the size effects emerge from independent components of the network (Zorzi & Butterworth, 1999). Finally, the comparison distance effect and the priming distance effect (in a priming task the size of the priming effect depends on the distance of the prime and the target numbers) were found to be independent (Reynvoet et al., 2009), although they should be related according to the ANS model. In a different group of works it was found that the symbolic and nonsymbolic number processing have differences where the ANS model would predict similarities. In a detailed analysis it was demonstrated that contrary to former analyses while the ANS model can describe dot comparison performance relatively well, the model has systematic biases in describing Indo-Arabic comparison, therefore, symbolic and nonsymbolic number comparison might rely on two different representations (Krajcsi et al., submittedb). Also, as it was mentioned above, according to several findings symbolic and nonsymbolic comparisons do not correlate in children (Holloway & Ansari, 2009; Sasanguie et al., 2014), and nonsymbolic comparison mostly does not correlate with typically symbolic mathematical achievement (De Smedt et al., 2013).

The current work investigated a fundamental prediction of the ANS model, whether distance and size effects correlate in any number notations, and the results do not support the model in symbolic comparison. While it is clear that the ANS model is supported by many empirical results and by many theoretical considerations, the DSS can explain many phenomena that was formerly attributed to the ANS model, and in some contrasts the DSS can offer better account for former and new phenomena. To explore the status of the ANS and the DSS models, further effort should be taken to clarify the scope of the former data, and new tests and considerations are required to evaluate the competing models.

Acknowledgments

I thank Krisztián Kasos, Petia Kojouharova, Ákos Laczkó and Gábor Lengyel for their comments on an earlier draft of the manuscript.

References

Cantlon, J. F. (2015). Analog origins of numerical concepts. In D. C. Geary, D. B. Berch, & K. Mann Koepke (Eds.), *Evolutionary Origins and Early Development of Number Processing* (Vol. 1, pp. 225–251). London: Academic Press.

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557. https://doi.org/10.1073/pnas.1113195108

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition* (Vol. XXII, pp. 527–574). Cambridge: Harvard University Press.

Dehaene, S., & Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition*, *43*, 1–29. https://doi.org/10.1016/0010-0277(92)90030-L

Eger, E., Sterzer, P., Russ, M. O., Giraud, A.-L., & Kleinschmidt, A. (2003). A supramodal representation in human intraparietal cortex. *Neuron*, *37*, 719–725.

Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General*, *141*(4), 642–648. https://doi.org/10.1037/a0026218

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Hinrichs, J. V., Berie, J. L., & Mosell, M. K. (1982). Place information in multidigit number comparison. *Memory & Cognition*, *10*(5), 487–495.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children’s mathematics achievement. *Journal of Experimental Child Psychology*, *103*(1), 17–29. https://doi.org/10.1016/j.jecp.2008.04.001

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*(4), 498–525.

Krajcsi, A., Lengyel, G., & Kojouharova, P. (submitteda). Symbolic number comparison is not processed by the analogue magnitude system: different symbolic and nonsymbolic numerical distance and size effects.

Krajcsi, A., Lengyel, G., & Kojouharova, P. (submittedb). Symbolic number processing: Analogue Number System or Discrete Semantic System?

Krajcsi, A., Lengyel, G., & Kojouharova, P. (submittedc). The source of the symbolic numerical distance and size effects.

Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology*, *60*(4), 227–234. https://doi.org/10.1027/1618-3169/a000191

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences*, *109*(46), 18737–18742. https://doi.org/10.1073/pnas.1207212109

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*, 177–190. https://doi.org/10.1038/nrn1626

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Poltrock, S. E., & Schwartz, D. R. (1984). Comparative judgments of multidigit numbers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 32–45. https://doi.org/10.1037/0278-7393.10.1.32

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*(6), 607–614.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology, 103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Sasanguie, D., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *The Quarterly Journal of Experimental Psychology*, *67*(2), 271–280. https://doi.org/10.1080/17470218.2013.803581

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. https://doi.org/10.2307/1412159

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review, 12*(1), 66–80. https://doi.org/10.3758/BF03196349

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Zorzi, M., & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the twenty first annual conference of the Cognitive Science Society* (pp. 778–783). Mahwah: Erlbaum. Retrieved from http://eprints.ucl.ac.uk/91024/

Chapter 7: Interference between number magnitude and parity: Discrete representation in number processing

The present chapter has been published:

Attila Krajcsi, Gábor Lengyel, Ákos Laczkó (2018). Interference between number magnitude and parity: Discrete representation in number processing. Experimental Psychology. 65(2), 71–83. https://doi.org/10.1027/1618-3169/a000394

Interference between number magnitude and parity: Discrete representation in number processing

Attila Krajcsi[1], Gábor Lengyel[2], Ákos Laczkó[1]

[1]Eötvös Loránd University, Department of Cognitive Psychology

[2]Central European University, Department of Cognitive Science

Interference between number magnitude and other properties can be explained by either an analogue magnitude system interfering with a continuous representation of the other properties or by discrete, categorical representations in which the corresponding number and property categories interfere. In this study, we investigated whether parity, a discrete property which supposedly cannot be stored on an analogue representation, could interfere with number magnitude. We found that in a parity decision task the magnitude interfered with the parity, highlighting the role of discrete representations in numerical interference. Additionally, some participants associated evenness with large values, while others associated evenness with small values, therefore, a new interference index, the dual index was introduced to detect this heterogeneous interference. The dual index can be used to reveal any heterogeneous interference that were missed in previous studies. Finally, the magnitude-parity interference did not correlate with the magnitude-response side interference (SNARC effect) or with the parity-response side interference (MARC effect), suggesting that at least some of the interference effects are not the result of the stimulus property markedness.

Keywords: analogue magnitude system; discrete number representation; numerical interference; homogeneous interference; heterogeneous interference; SNARC effect; MARC effect; PNARC effect

Highlights:

- Numbers interfere with discrete parity property, supporting discrete number representation models in numerical interference effects
- Numerical interference effects do not correlate, contradicting the polarity and the markedness interference models

- Heterogeneous interference can be revealed with the new dual index method

Number magnitude interference effects

Number magnitude can interfere with other properties. A salient example is the SNARC effect. In the SNARC (Spatial-Numerical Association of Response Codes) effect, typically in a parity task, participants respond faster for small values with the left response button than with the right response button, and they respond faster for large values with the right response button than with the left response button (Dehaene, Bossini, & Giraux, 1993). Thus, small numbers are associated with the left side, and large numbers with the right side, although the association depends in part on cultural background, e.g., Iranian participants associate small numbers with the right side, and large numbers with the left side (Dehaene et al., 1993). Another example of numerical interference is the size congruity effect, in which the physical size of the stimuli interfere with the numerical value (Henik & Tzelgov, 1982). As another example, number magnitude also interferes with duration (Oliveri et al., 2008).

There are various explanations for such interference effects. In the present introduction we take the SNARC effect as an example, but the explanations could be generalized to several other symbolic numerical interference. To explain the SNARC effect some of the explanations have presupposed analogue (i.e., continuous) representations in the background (top of Figure 41), while other accounts have supposed discrete (i.e., categorical) representations (bottom of Figure 41).

Analogue explanations

As a first explanation, a continuous, noisy representation, the Analogue Number System (ANS) is proposed, which system works according to the Weber’s law, similar to many other representations processing simple perceptual properties (Moyer & Landauer, 1967). Dehaene and his colleagues propose that this ANS is what causes the interference: it might have a spatial property (values are connected to spatial locations), and this spatial property interferes with the spatial representation of the response locations (Dehaene et al., 1993).

According to a related, second explanation, many non-numerical properties and the numerical values can be stored in similar analogue systems. While these systems process different input, all of them adhere to Weber’s law (Figure 41). This Generalized Magnitude System (GMS) model supposes that partly because of the similar mechanisms, these representations might interfere with each other (Cantlon, Platt, & Brannon, 2009). In line with this reasoning, many interference effects were described between continuous properties, e.g., pitch interfering with response location (Rusconi, Kwan, Giordano, Umiltà, & Butterworth, 2006), luminance interfering with response location (Fumarola et al., 2014), or duration interfering with number magnitude (Oliveri et al., 2008). Although based on different motivations, similar additional models predicting interference of continuous representations have been proposed (Bueti & Walsh, 2009; Cohen Kadosh, Lammertyn, & Izard, 2008; Henik, Leibovich, Naparstek, Diesendruck, & Rubinsten, 2012; Walsh, 2003).

Discrete explanations

A next group of models explaining the SNARC effect suppose discrete representation in the background (bottom of Figure 41). According to a third model, the SNARC effect is rooted in the interference of discrete concept pairs (Proctor & Cho, 2006). In this polarity model, antonyms, like small-large, left-right or true-false, have polarity properties: one of the items is positive, the other one is negative. (*Positive* and *negative* are arbitrary labels of the items, and they are not directly related to the sign of the numbers.) Concepts with the same polarity can enhance the processing, while concepts with opposing polarity might inhibit it, resulting in interference. In the case of the SNARC effect the small-large concepts interfere with the left-right concepts, supposing that in Western culture *small* and *left* have the same polarity. The polarity model was successfully applied to simulate the SNARC effect (Leth-Steensen, Lucas, & Petrusic, 2011). (Although see a study where an extension of the polarity model in interference effects was not confirmed: Santiago & Lakens, 2013.)

The markedness model offers a similar explanation as the polarity model (Hines, 1990). The markedness model proposes that one member of an opposing pair is marked, for example, in the case of parity, odd is marked. As originally proposed in linguistics, markedness of a word has many related properties, e.g., marked members of the pairs are usually slower to process, children learn them later, and sometimes they are linguistically formed from the unmarked member, etc. Critically, markedness can explain interference as well: items with same markedness can be processed faster, while items with opposing markedness are processed slower (Nuerk, Iversen, & Willmes, 2004). The markedness model is similar to the polarity model with a few important differences: while markedness is related to language and is a relatively stable property of a pair, polarity is more flexible and can be task dependent (Cho & Proctor, 2007). Also, in the polarity model it is not always clear why positive or negative values are set for a specific property (Huber et al., 2015). Because polarity and markedness models have very similar predictions about the potential interference effects, we handle these two models together in several sections of the paper (but see the discussion of the critical differences below in the discussion of the PNARC interference and in the General discussion). Although in the numerical interference context the markedness model was proposed only to explain the parity-response interference (MARC effect[16]), but not the SNARC effect (Nuerk et al., 2004), it is reasonable to extend to several other interference effects, as discussed in the work of Proctor and Cho (2006), or as it was noted later (Patro, Nuerk, Cress, & Haman, 2014).

A fourth model also proposes a discrete explanation for the SNARC effect. A formerly introduced delta-rule connectionist model of numerical effects (Leth-Steensen & Marley, 2000; Verguts & Fias, 2004, 2008; Verguts, Fias, & Stevens, 2005; Verguts & Van Opstal, 2014) could explain and model the SNARC effect (Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006). In this model the number layer (number field) is connected to additional layers that are representing parity or categorical

16 The term MARC is used inconsistently in the literature. Nuerk et al. (2004) originally used the term MARC for both the parity-response interference, and for the markedness model, and the literature uses the term in both meanings. Here, we use MARC to denote the interference effect, in line with other similarly named interference effects, and we term the model as markedness, as the term has already been used in the linguistic and numerical cognition literature.

magnitude (large or small), finally, a response layer is used to represent left and right responses of a task (Figure 41). Based on cultural experiences, association is formed between the magnitude and the response layers. In a task, the magnitude is automatically activated, independent of whether the layer is relevant (e.g., in a comparison task) or not (e.g., in a parity task), then this categorical magnitude layer automatically activates the response layer, which will interfere with the task-relevant response layer activation, resulting in the SNARC effect. The model successfully explains some additional effects regarding the SNARC interference, for example, why slower responses show stronger effect, or why the SNARC effect shows a categorical pattern instead of a continuous pattern in a comparison task (Gevers et al., 2006). Because the model supposes that the key for the explanation is the categorization of the values (large vs. small in the magnitude layer), the model can be considered as a discrete model.[17]

17 Although the number field of the model was originally interpreted as an implementation of the ANS (Verguts & Fias, 2004; Verguts, Fias, & Stevens, 2005), the model is unable to produce Weber's law, since the size effect can be seen only after introducing the frequency of the numbers to the model (Verguts et al., 2005). Instead of the original interpretation, it might be more appropriate to consider the number field as an implementation of a discrete model, in which the "noise" of the number field is not a noise based on Weber's law, but the spreading activation of the discrete units (Krajcsi, Lengyel, & Kojouharova, 2016). Thus, the delta-rule connectionist model is a discrete model not only because of the magnitude layer, but also because of the number field.

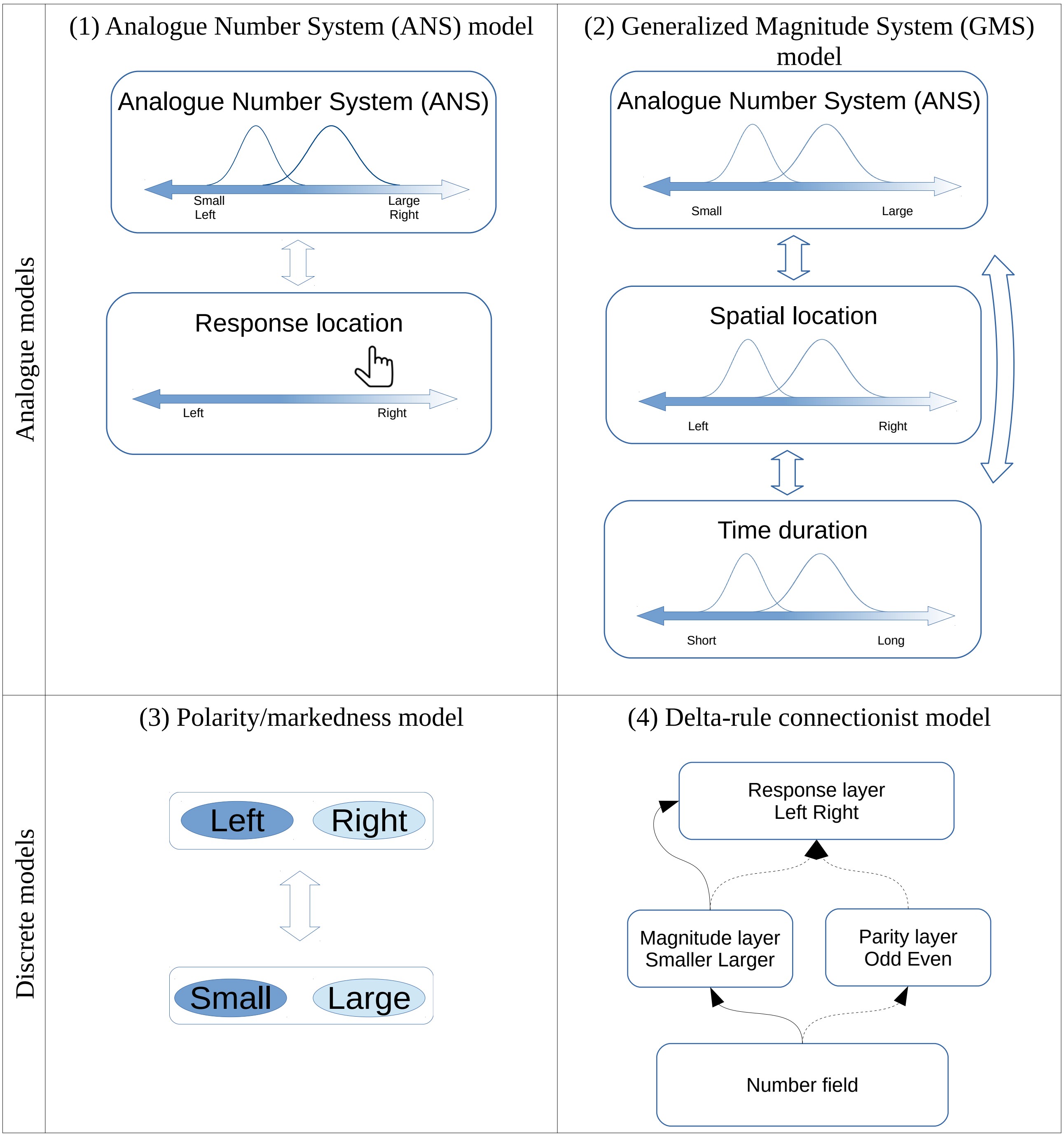


Figure 41 Four explanations of the SNARC effect. Light and dark arrows and light and dark nodes represent the same measurement direction and the same polarity/markedness. Solid lines in the delta-rule model represent automatic connections.

Some of the models introduced above propose an analogue explanation for the SNARC effect, while others propose discrete representations. One cannot contrast the two model types based solely on the presence of the SNARC effect, because all of the models give satisfactory explanations for it. In the present work we contrast the current analogue and the current discrete explanations with a new type of numerical interference.

Does numerical information interfere with a discrete property? The number magnitude-parity interference

One reason why the presence of the SNARC effect cannot contrast the analogue and the discrete models is that both properties of the interference (number and space) could be represented both continuously or discretely: either on analogue representations as proposed by the ANS or the GMS models, or on discrete representations as suggested by the polarity/markedness and the delta-rule connectionist models. This problem also applies to other known numerical interference effects, e.g., in the size congruity effect (Henik & Tzelgov, 1982), the physical size could also be stored on both continuous and discrete representations.

The present work investigates whether number magnitude interferes with a discrete property that is supposedly not stored on an analogue representation. Here, we investigate the presence of the magnitude-parity interference. Following the tradition of the literature we term it the PNARC (Parity-Numerical Association of Representational Codes) effect.[18]

Existence of the PNARC effect would be in line with the discrete models above, but cannot be explained by the presented analogue models (Figure 42). More specifically, (1) extending the ANS model could imply that the ANS has a parity feature (similar to the spatial feature), which does not seem to be a feasible assumption. Consequently, the ANS model does not suppose the presence of the PNARC effect. (2) Because most probably parity is not stored in an analogue representation, the GMS model does not presuppose the PNARC effect either. While in a strict sense neither the ANS nor the GMS model exclude the presence of the PNARC effect (they might interpret the potential PNARC effect as the result of an unspecified additional numerical representation), the PNARC effect would question the general statement that the ANS is the only source of the magnitude-related symbolic numerical effects. (3) According to the polarity/markedness models, the PNARC effect can be present, because any property pairs that have polarity/markedness can interfere. Actually, the models have an even stronger prediction. In our parity task there are three properties that are relevant in the interference effects we investigate (Figure 43): parity, number and response location, and all of them can interfere with each other. It is known that numbers and response location interfere (the SNARC effect: Dehaene et al., 1993), and response location and parity also interfere (the MARC effect: Nuerk et al., 2004). Based on these results the models suppose that all three relevant properties have polarity, consequently the models also predict that the interference of parity and magnitude should also exist. Thus, the polarity/markedness models not only allow the appearance of the PNARC effect, but expect it. (4) In the case of the SNARC effect the delta-rule connectionist model relies on the dual-route model of interference, in which both the relevant and the irrelevant information is processed simultaneously, and the identical or differing results of the two routes generate the interference (Gevers et al., 2006). The architecture of the model displayed above is deliberately an initial version, and additional details, e.g., connections can be added (Verguts & Fias, 2008). It is easy to imagine that an automatic connection between the magnitude and the parity layers can produce the PNARC effect. Although it is not trivial what environmental

18 Note that unlike in the term SNARC (Spatial-Numerical Association of Response Codes) effect, in the term PNARC effect, R stands for Representational and not Response, because in a parity-number association response is probably not a key component.

experiences would form such a connection, the possibility cannot be excluded. Overall, based on our current knowledge of the model, the connectionist model does not predict the appearance of the PNARC effect, although the model allows for its presence. (See the summary of the four models about the appearance of the PNARC effect in Table 14.)

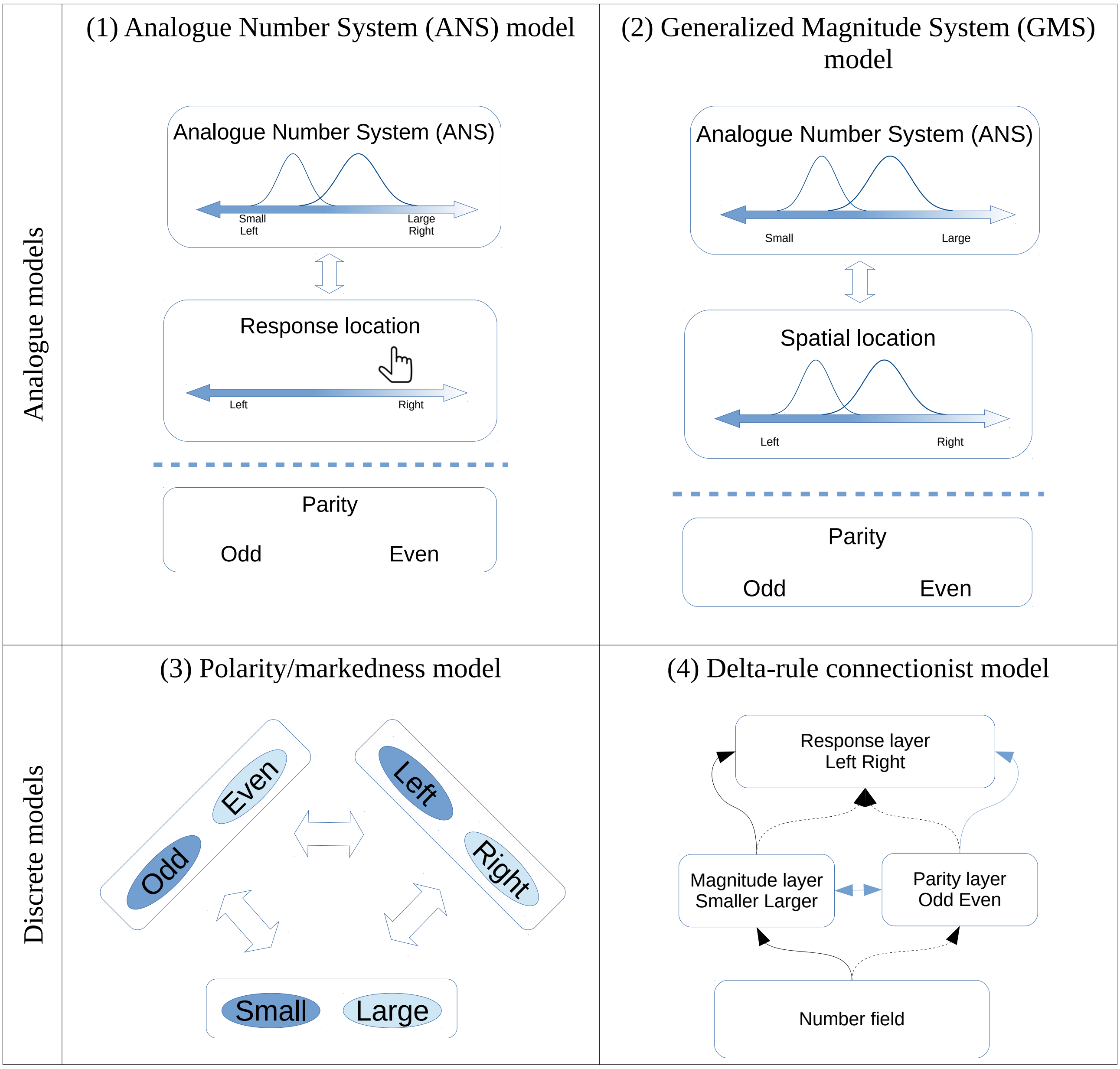


Figure 42 Prediction of the four models for the MARC and PNARC effects. In the delta-rule connection model the new connections are denoted with blue arrows.

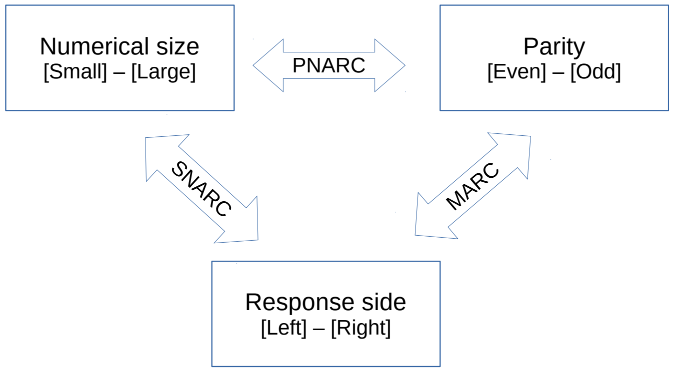


Figure 43 Three critical properties in our parity task, and the interference effects between them

Interestingly, the presence of the PNARC effect has already been reported in the literature, although its importance or even its presence was not discussed. While investigating the MARC and SNARC effect in various conditions (positive and negative Indo-Arabic numbers and number words) in a parity decision task, Nuerk et al. (2004) analyzed the data with ANOVA, where the number magnitude-parity interaction was also reported. They found that in positive Indo-Arabic numbers both error rate and reaction time showed a number magnitude-parity interaction. In number words, only the error rate showed this interaction, while in negative Indo-Arabic numbers no interaction was observed. While the significant results were reported, the presence or the importance of these interactions as a PNARC effect was not discussed either in the results part or in the discussion. Here, we try to replicate those former findings, and in contrast with the Nuerk et al. (2004) study we discuss the theoretical consequences of the presence of that interference, and we also investigate the relations of the interference effects measured in parity task (see the Possible correlation section below).

To summarize, the suggested PNARC effect cannot be explained by the current analogue models, while it would be in line with the current discrete models. The main question of the present study is whether the PNARC effect exists, supporting the discrete models, or does not exist, supporting the analogue models.

Possible correlation of the interference effects

As already noted above, the planned parity task includes three properties that might interfere with each other: parity, number and response location (Figure 43). The three properties might form three interference effects: (a) the SNARC effect between the number magnitude and response code (Dehaene et al., 1993), (b) the MARC effect between the response code and the parity (Nuerk et al., 2004), and (c) the potential PNARC effect between the parity and number magnitude.

Beyond the presence of the PNARC effect, it might be of interest whether the interference effects correlate (Table 14). A precondition of such analysis is the appropriate variability of the interference effects. Accordingly, several SNARC studies revealed individual differences in the interference indices (see a summary in Hoffmann, Mussolin, Martin, & Schiltz, 2014). Regarding the predictions of the models (1-2) both ANS and GMS models argue that the SNARC effect originates in the ANS, while the MARC and supposedly the potential PNARC effects are independent of the ANS. Thus, the SNARC effect probably does not correlate with the other two interference effects. The models do not have any predictions about the correlation of MARC and PNARC effects. (3) The polarity/markedness models predict that interference effects that contain overlapping property might correlate (e.g., SNARC and PNARC effects might correlate, because both contain the magnitude), supposing that individual differences of the indices originate in the individual differences of the overlapping property. Thus, the models predict correlation, if sufficient statistical power is given. (4) According to the delta-rule connectionist model the interference effects root in the connections between specific layers, and because technically all connections between layers are independent of other connections, in a strict sense the interference effects should not correlate. However, considering the environmental sources of those connections, some correlation might be expected. Due to the flexibility of this model and because only a few relevant details are known about it, it is not trivial to propose a straightforward prediction about the correlation of the interference effects.

There are two general methodological constrains in this test. First, it is possible that there are general components (e.g., the conflict resolution), that are independent of the property-specific processing, and which might generate correlation. Second, because the sizes of the potential correlations are not known, it is possible that our test will lack statistical power. Thus, either we can observe correlations even if the overlapping representations would not predict them, or we might miss observing the correlations even if they are present. Therefore, it is possible that the presence or the absence of the correlations cannot contrast the models. However, a further aspect of the results will help us to clarify these issues, which we are going to discuss in the relevant parts of the Results section.

To summarize, in this study we investigate (a) whether parity interferes with number magnitude, and (b) whether the three interference effects of the three relevant properties (SNARC, MARC and PNARC) correlate. The predictions of the four models are summarized in Table 14.

	Analogue models	Discrete models	
	(1) ANS and (2) GMS models	(3) Polarity/markedness models	(4) Delta-rule connectionist model
Appearance of the PNARC effect	Strictly, independent of the ANS/GMS models. More generally, PNARC should not exist.	PNARC should exist	PNARC might exist
Correlation of the interference effects	SNARC does not correlate with MARC or PNARC. Correlation between MARC and PNARC is unspecified	The three interference effects should correlate	Either lack or presence of correlation is possible

Table 14 The prediction of the four models for the two critical results.

Methods

Participants. Fifty-seven university students took part in the experiment for partial course credit. Participants with higher error rate than the mean + 2 standard deviation of the group, which was

14.9%, were excluded (3 participants). Data of 54 participants were analyzed: 39 women, with mean age of 21.4 years, 2.4 years standard deviation. Six of the participants were left-handed, and because this is only a very small part of the sample, right- and left-handed participants were not analyzed separately.

Stimuli and procedure. Participants made parity decisions. Numbers were between 1 and 9, and the instruction was explicit about this range.[19] In a trial, a single Indo-Arabic digit appeared on the screen, and the participants decided whether the digit was even or odd, by pressing one of the two response buttons. The digit was visible until response. In case of erroneous response an auditory feedback was given. After the response, a blank screen was shown for 700 ms. Two hand conditions were used to ensure that all numbers and parities can be assigned to both response sides. In one of the hand conditions even numbers were responded to with the left Control key with the left hand, and odd numbers were responded to with the decimal sign on the numeric keypad with the right hand (the leftmost and the next to rightmost keys on the keyboard), while in the other hand condition the stimulus-response association was reversed. In both conditions all numbers between 1 and 9 were presented 40 times. The order of the stimuli was randomized. All participants took part in both hand conditions, and the order of the conditions was counterbalanced across participants. The instruction stressed both speed and accuracy. Running the whole experiment required approximately 30 minutes for a participant. Stimulus presentation and data collection was performed with PsychoPy software (Peirce, 2009).

Analysis methods

Unified index calculation. First, the interference effects were computed with the regression analysis method that is frequently applied for testing the SNARC effects (Fias, Brysbaert, Geypens, & d’Ydewalle, 1996; Lorch & Myers, 1990). For example, to compute the PNARC effect, median reaction time of the correct responses was calculated for all digits for all participants. Then, the slope of the reaction time change across digits were calculated for both the odd and the even numbers. Finally, the difference of the two slopes was calculated (even slope subtracted from odd slope), and the deviation of the PNARC index from zero was tested.[20] Similar methods were applied to the other interference effects.

Dual index calculation. However, the interference effects can be investigated with an alternative method. While analyzing the data, we observed that in the case of the PNARC effect, the slopes

19 Unlike in some similar studies, number 5 was also used here, resulting in more odd than even numbers, therefore more ‘odd’ than ‘even’ responses in the parity decision task. However, this inequality in the responses does not cause any bias in the interference analyses applied in the present study.

20 Although this procedure is slightly different from the classic calculation procedure of regression analysis (Fias, Brysbaert, Geypens, & d’Ydewalle, 1996; Lorch & Myers, 1990), it gives the exact same result. In the case of the SNARC effect, first, the reaction times for all digits and separately for both hands are calculated, then the differences of the two hands for all digits are computed, and finally, the slopes of the reaction time change across digits are calculated. However, in the case of the PNARC effect, one cannot use a similar procedure, because analogously one should subtract the data of the even digits from the data of the odd digits, but a digit cannot be both even and odd. Consequently, the classic SNARC-like calculation procedure is impossible for the PNARC effect. Nevertheless, in the SNARC procedure, the slope of the hand differences gives mathematically the same result as the difference of the single hand slopes. Importantly, the latter version is the calculation procedure that we applied here for the PNARC effect. Thus, the procedure applied here gives the very same result as if the classic SNARC computation procedure were used.

showed a systematic relation: most of the participants, who had positive slope for the odd numbers (i.e., larger odd numbers are responded slower than smaller odd numbers), had negative slope for even numbers (i.e., larger even numbers are responded faster than smaller even numbers). In the same time, most of the participants showing negative slope for the odd numbers, has shown positive slope for even numbers. This means that some participants associated oddness with small numbers and evenness with large numbers, and some other participants associated these properties in a reversed way. We term this type of interference a *heterogeneous interference*, because although there are associations between the properties, the direction of the association is not uniform within the group. To analyze this relation in a more statistical way, one can plot the even and odd slopes on a scatter plot (as in Figure 44 later in the Results section) and look for the correlation. If there is a systematic relation between the two slopes as described above, one should observe a negative correlation. This negative correlation means that the magnitude information influences the parity decision time, but the direction of this influence shows individual differences, forming a heterogeneous interference.[21]

This heterogeneous interference in which different part of the sample shows different direction of the association is a known phenomenon in the literature. A similar heterogeneous interference was observed in Dehaene et al. (1993) where the direction of the numerical-spatial association depended on reading habits: while Iranian participants, moving recently to France, associated large numbers with the left side (in line with the right-to-left Persian writing system), Iranian participants, living in France for a longer time, associated large numbers with the right side (in line with the left-to-right Western writing system). (See various additional factors that may change the direction of the SNARC effect in Shaki & Gevers, 2011) It is not the presence of the heterogeneous interference that is new in our analysis, but the method with which the interference can be detected: while in the cited work the time of movement from Iran to France is necessary to reveal the interference, with the current method this additional information is not necessary to discover a heterogeneous interference. (We discuss this example in more details in the next subsection.)

While we have discussed the method in the context of the PNARC effect, this correlational method can be applied to any interference effects investigated in the present study, or to any other interference effects.

See more details about the relation of the unified index and the dual index, and what information they are sensitive to in the Electronic Supplemental Material.

Results and discussion

The raw data are available at https://osf.io/g7t2q/.

21 Importantly, the correlation reflecting a heterogeneous interference can not be a mathematical artifact, because the slopes of the even and the odd numbers are independent, they are calculated based on different trials, thus, they could have been uncorrelated, i.e., participants might show increasing or decreasing reaction time for larger numbers independent of the parity of the values.

Interference effects

PNARC effect. PNARC effect was calculated as the reaction time change (slope) across the even or the odd numbers. For the unified index the slopes of the two parity were subtracted (even slope subtracted from odd slope). The PNARC index did not differ from zero with the unified index, mean slope = 1.33, 95% CI [-1.03, 3.69], $t(53) = 1.13$, $p = 0.265$, showing that homogeneous PNARC effect is not observable. However, measured with the dual index, the two slopes show a clear negative correlation, $r(52) = -0.461$, $p < 0.001$ (Figure 44).

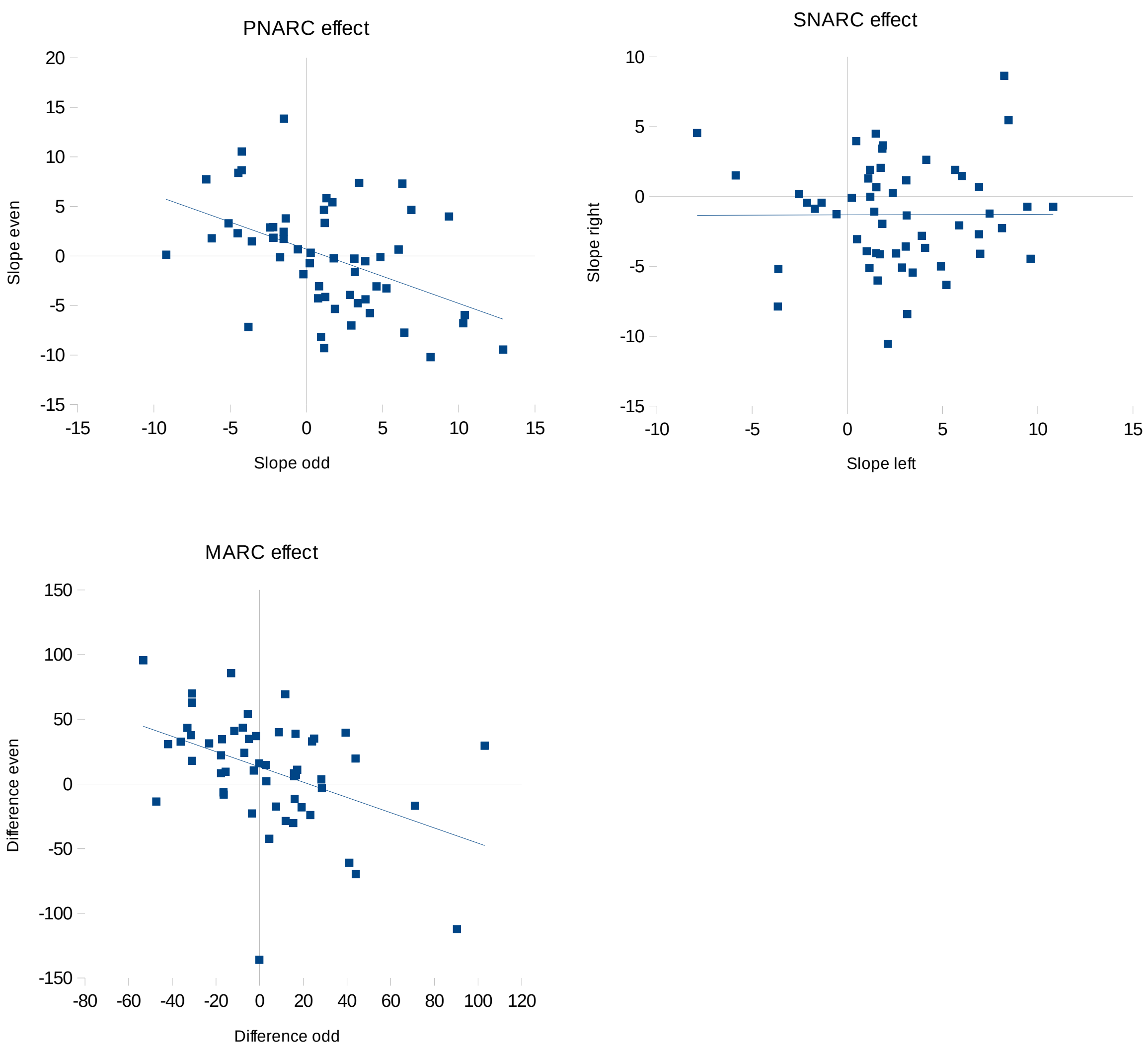


Figure 44 Scatterplot of (a) the slopes of even and odd numbers (PNARC), (b) the slopes of the right and left hands (SNARC), and (c) the hand differences of the even and odd numbers (MARC).

The present results reveal the PNARC effect, although the direction of parity and number magnitude association is not homogeneous in our sample, similar to the already mentioned SNARC effect of Experiment 7 in Dehaene et al. (1993). On one hand, the present results replicated the findings of Nuerk et al. (2004), because both studies found PNARC effect. On the other hand, while in the present study the PNARC effect is heterogeneous, in the Nuerk et al. (2004) study the interference is homogeneous. The difference most probably cannot be explained by different analysis methods, because both ANOVA and the slope measurement methods rely on similar information, and they should give similar results (Pinhas, Tzelgov, & Ganor-Stern, 2011), At the moment it is not clear what could cause the difference, if this difference is reliable. Additionally, it is not possible to reveal the cause of the different parity-number associations based on the present results. Different associations can be formed by different individual numerical experiences (find some influencing factors in the case of the SNARC effect in Shaki & Gevers, 2011), or the differences can be ad hoc associations applied only in a single session (as similarly observed in the case of the SNARC effect in Fischer, Mills, & Shaki, 2010). Still, for the aim of the present work, it is not an important issue to know why the interference is heterogeneous, or what the sources of these individual differences are, but the important result is that the interference between parity and number magnitude is clearly observable. The appearance of the PNARC effect is not in line with the current analogue models, while it is more coherent with the presented discrete models.

SNARC effect. SNARC effect was calculated as the reaction time change (slope) across the values for both hands. For the unified index the slopes of the two hands were subtracted (left slope subtracted from right slope). SNARC effect was observable with the unified index, mean slope = -3.94, 95% CI [-2.48, -5.41], $t(53) = -5.4$, $p < 0.001$. Measured with the dual index, there was no correlation (Figure 44), $r(52) = 0.004$, $p = 0.975$. These results replicated many studies, confirming that the SNARC effect is predominantly homogeneous in the published studies.

MARC effect. MARC effect was calculated as the reaction time difference between left-hand and right-hand responses (left hand subtracted from right hand) for both even and odd numbers. For the unified index the two hand-differences values were subtracted (even subtracted from odd). MARC effect was not shown with the unified index, mean slope = 6.54, 95% CI [23.72, -10.63], $t(53) = 0.764$, $p = 0.448$, while it was significant with the dual index, $r(52) = -0.432$, $p = 0.001$ (Figure 44). These results show that MARC effect is mainly heterogeneous in our sample.

Lack of the MARC effect with the unified index is not surprising, since previous studies reported mixed results. Some studies found significant MARC effect with typically 20-30 participants (Berch, Foley, Hill, & Ryan, 1999; Cho & Proctor, 2007; Landy, Jones, & Hummel, 2008; Nuerk et al., 2004), while some others did not find the effect (Nuerk, Bauer, Krummenacher, Heller, & Willmes, 2005; Roettger & Domahs, 2015), and some additional studies found the effect in some tasks and/or conditions, and they could not find it in some other tasks or conditions, while the critical variable that makes the effect appear and disappear could not be identified (Berch et al., 1999; Cho & Proctor, 2007; Nuerk et al., 2004).

These previous studies might suggest that the size of the MARC effect is much smaller than the size of the SNARC effect, which is easier to replicate. However, our result also indicates that the MARC effect might be stronger than it has been shown in previous studies, because the heterogeneous nature of the effect in these samples could not be unveiled formerly.

To summarize, all three interference effects which were investigated here are observable, and in our sample the SNARC effect is homogeneous, while the PNARC and MARC effects are heterogeneous.

The present sample included 6 left-handed participants. It has been demonstrated that the MARC effect might depend on the handedness of the participant, while in the SNARC effect no handedness effect was observed (Huber et al., 2015). To investigate whether the present results are stable independent of the handedness of the participants, the data were reanalyzed without the 6 left-handed participants. (Left-handed participants were not analyzed separately or in contrast with the right-handed participants because of the very small sample size.) Removing the six participants hardly changed the statistical results, and the very same significance pattern was found as with the whole sample. See the detailed results and the scatterplots as a function of handedness in the Electronic Supplemental Material.

See a replication study that ensures the reliability of these findings in the Electronic Supplemental Material.

While setting the aims of this study, we did not consider the possibility of a heterogeneous interference. Does the existence of the heterogeneous interference influence the predictions of the models? While in several models it is possible that the direction of the association is flexible, the sources of the directions may set some constrains. First, the markedness model supposed that the direction of the association is rooted in language use. However, language use in a linguistically and culturally homogeneous sample, as in our study, would predict a homogeneous interference. Therefore, markedness model seems to be an improbable explanation for the observed heterogeneous interference. Similarly, because the SNARC effect has been demonstrated to be dependent on the writing system (Dehaene et al., 1993), it is supposed that the ANS and the GMS models suppose homogeneous interference effects, unless the non-numerical component of the interference could explain the heterogeneous direction. Importantly, as discussed above, ANS/GMS models cannot explain either the MARC or the PNARC effects, because the MARC effect does not rely on the number magnitude information, and the PNARC effect relies on a discrete property. Finally, the more flexible polarity and the delta-rule models can be in line with heterogeneous effects.

Critically, our results show the presence of the PNARC effect. This result can be explained by two of the presented discrete models, while the introduced analogue models can not account for it. Additionally, the PNARC effect is an interference between number magnitude and another property, as in the case of the SNARC effect, and as in the cases of other symbolic numerical interference effects. These similarities raise the possibility that if the PNARC effect cannot be explained by an analogue model, then maybe the analogue explanations are also incorrect in the cases of other similar symbolic numerical interference effects.

Correlation of the three interference effects

To repeat the predictions of the models, according to the analogue models the SNARC effect does not correlate with the other two interference effects, according to the polarity/markedness models the three interference effects correlate, while the delta-rule connectionist model is underspecified to have straightforward predictions about the correlations (Table 14). Note that unified indices are used for these correlations.

However, appearance of the heterogeneous interference modified the prediction of the polarity/markedness model. Note that in the case of heterogeneous interference, the correlation of the interference effects means that a specific direction in one of the interference effects is related to another specific direction in the other interference. E.g., correlation between the MARC and the PNARC effect means that if, for example, someone associates evenness with large numbers, then the same person associates evenness with the right side (direction of the associations depend on the sign of the correlation). Regarding the modified predictions, first, heterogeneous interference ensures the statistical power in the correlation of heterogeneous interference effects if the polarity/markedness model is correct. In the polarity/markedness model, a heterogeneous interference signals that one of the properties shows variance in the polarity (i.e., it changes across participants). Because SNARC is homogeneous, it means that in our sample the polarity of the number magnitude and the response side is fixed (i.e., these properties do not change across participants), therefore it is the parity that should vary across participants. Additionally, a significant correlation of the dual index shows that the data have sufficient variance for the given sample size, and a similar variance will be used in the interference correlation study: instead of the negative correlation seen in the scatterplot of the dual index, the variation along the y = -x axis of the unified index will be used (Figure 2 in ESM). Thus, if the polarity/markedness model is correct, in the present sample the PNARC and the MARC effects should correlate significantly. Second, according to the polarity/markedness model the heterogeneous interference effects will not correlate strongly with the homogeneous interference (i.e., the SNARC effect), because in the heterogeneous polarity the unified index changes its sign, and while on one side (i.e., with one sign) it creates positive correlation with the homogeneous interference, it creates negative correlation with the other sign, and the two correlations extinguish each other in the whole sample. Finally, the appearance of the heterogeneous interference does not change the predictions of the other models about the correlations.

Results did not show a significant correlation between the MARC and the PNARC effects, $r(52) = -0.101$, $p = 0.469$. This result is at odds with the polarity/markedness model. As it was discussed, the lack of the correlation cannot be caused by the lack of the statistical power, because the dual index correlation reflected sufficient statistical power in an analysis based on similar variance. This single result can be in line with the analogue models and the delta-rule connectionist model.

Neither the SNARC and the PNARC effects correlated, $r(52) = -0.099$, $p = 0.476$, nor the SNARC and the MARC effects, $r(52) = 0.251$, $p = 0.067$. (Correlation of the SNARC and MARC effects were also investigated in Huber et al., 2015, and in line with our result, no significant correlation was found.) The lack of these correlations can be in line with all models, thus these correlations cannot contrast the models. Specifically, (1-2) according to the analogue models an analogue system caused SNARC effect should not correlate with the other systems related MARC or PNARC effects. (3) According to the polarity/markedness model, because SNARC was a homogeneous interference, and PNARC and MARC were heterogeneous interference effects, these interference effects should not correlate. (4) The delta-rule connectionist model is flexible and for our purposes it is underspecified at the moment, thus any result can be consistent with it, although it is not trivial to specify what environmental experiences can form connection weights that could cause uncorrelated interference effects.

Among the three correlations, the PNARC-MARC correlation was critical from the viewpoint of model testing, because polarity/markedness models clearly predicted a correlation. Nonetheless, the correlation was not observed. This result does not entirely invalidate the polarity/markedness model, however. It is possible that while not all interference effects are caused by the mechanisms described by the model, some of them are. Although in this case the polarity/markedness model should specify why some of the interference effects can be explained by the model and why some of them cannot.

General discussion

The main aim of the present study was to contrast the current models explaining numerical interference, in the frame of the analogue and discrete representation types. Our first specific question was whether number magnitude can interfere with a discrete property (in this case, parity), which is not likely to be handled by an analogue representation. Our results show that parity interferes with number magnitude, in this case supporting the introduced discrete representation models, and opposing the current analogue models.

After running the present study we found a conference publication describing a study with a different motivation, but with results relevant to our inquiry (Landy et al., 2008). In this work, the spatial feature of the ANS was contrasted with the non-spatial polarity model. To contrast whether a spatial or a non-spatial model can explain the SNARC effect, the parity task was modified to avoid the two-side spatial responses, and participants used verbal yes-no answers to decide whether a number was even or not. If in this paradigm there is still a number-response interference, then it cannot be explained by a spatial account. On the other hand if the interference cannot be observed in this paradigm, it means that the spatial response (and a spatial representation) was responsible for the interference. The SNARC-like number-response effect was observed even with these yes-no responses (i.e., faster “yes” response with larger numbers than with smaller numbers, and faster “no” response with smaller numbers than with larger numbers), showing that it is not the spatial feature of the ANS that causes the interference. From our viewpoint this study has another relevant aspect: the utilized yes-no responses are discrete categories, similar to the parity, unlikely to be stored in an analogue representation. This serves as another example that discrete and continuous properties might interfere, again supporting the discrete explanations against the current analogue models. While our work shows an interference between two properties of the stimuli, the verbal response parity task shows an interference between a property of the stimuli and the responses. The parity-number magnitude interference demonstrated in our study and the yes-no – number magnitude interference shown by Landy et al. (2008) converge to the same direction, supporting the role of the discrete representation in numerical interference effects.

Additionally, we found that the PNARC interference was heterogeneous. This property cannot be explained by the markedness model, because in the present sample with homogeneous linguistic and cultural background one would have expected a homogeneous interference. Markedness model seems to be an improbable explanation for the numerical interference effects for other reasons, as well. For example, Cho and Proctor (2007) has demonstrated that response side interferes not only with the parity of the number, but also with the property whether a number can be divided by 3. Because it is unlikely that the category whether a number can be divided by 3 is already stored as a linguistic category, linguistic markedness does not seem to be a likely explanation for that interference. More generally, flexibility of these interferences (Shaki & Gevers, 2011) makes the markedness model a less probable account for the numerical interference effects.

Interpreting these results more generally, different options are possible. (1) On one hand, we can keep the suppositions of the current models, suggesting that (a) all number magnitude related interference effects are handled by the same number representation and (b) the interference effects are rooted in the same types of representations (i.e., both representations are either continuous or discrete). If these suppositions are correct, then number representation should be a discrete one, because while some numerical interference effects can be explained by both types of representations, the PNARC effect and the yes-no – number magnitude interference can be supported only by a discrete system. This supposed discrete number representation would be in line with other approaches proposing that other symbolic number processing effects can also be explained with discrete number representations (Krajcsi, Lengyel, & Kojouharova, 2016). (2) However, alternatively, it is possible that different numerical interference effects are supported by different number representations (Patro et al., 2014), and therefore, in some interference effects potentially an analogue number representation is involved, while in some other interference effects a discrete system is utilized. (Various homogeneous and heterogeneous numerical interference effects found in the present study may also be in line with this supposition.) If this is the case, further studies are needed to investigate what type of representations are involved in the specific numerical interference effects, and it cannot be taken for granted that interference effects of continuous properties necessarily mean analogue representations in the background. (3) Finally, one might also question the second supposition of the current models, and imagine that an analogue representation could interfere with a discrete representation, which possibility cannot be excluded at the moment. Importantly, this hypothesis should introduce an entirely new type of model to account for the numerical interference effects.

The second main question of the present study was whether the interference effects correlate, since this result may test the polarity and markedness models. We found no correlation between the interference effects (in line with the finding of Huber et al., 2015). These results may be consistent with other findings showing differing properties of different interference effects (e.g., Huber et al., 2015 found that while handedness modulated the MARC effect, it did not modulate the SNARC effect), suggesting that different interference effects may rely on different processes. This result shows that the interference effects cannot be rooted exclusively in the polarity or markedness of the categories. Finally, the architecture of the delta-rule connectionist model could explain for non-correlating interference effects, because the connections between the properties can be changed independently between different property pairs. However, at the moment it is not clear what environmental input would set the connections independently for all property pairs. Overall, in its current form none of the current models can readily explain the present findings. Additional convergent data and further modifications of the model seem to be necessary to ensure that the models are appropriate to describe the symbolic numerical effects.

More generally, it is possible that the lack of the correlations between the interference effects reflects that the interference effects do not root in the labeling of the properties (e.g. markedness or polarity of the number magnitude, parity, etc.) themselves, but in the relation or connection of those properties, as for example, in a delta-rule model. Alternatively, it is also possible that different interference effects may be supported by different number representations, and that is why the interference effects do not correlate.

A third, methodological result of the present work is the introduction of the method for revealing heterogeneous interference. The former, unified index was only appropriate to show homogeneous interference, and occasional methodological extensions were required to unveil heterogeneous interference (as in Dehaene et al., 1993). With the newly presented dual index heterogeneous interference can also be shown. We argue that the two methods mostly complement each other, and interference studies would benefit from using both indices to discover both homogeneous and heterogeneous interference effects.

Acknowledgments

We thank Noémi Éltető, Krisztián Kasos, Petia Kojouharova, Craig Leth-Steensen, Péter Levente Soós and Eszter Szabó for their comments on a previous version of the manuscript.

References

Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. (1999). Extracting Parity and Magnitude from Arabic Numerals: Developmental Changes in Number Processing and Mental Representation. *Journal of Experimental Child Psychology*, *74*(4), 286–308. https://doi.org/10.1006/jecp.1999.2518

Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1525), 1831–1840. https://doi.org/10.1098/rstb.2009.0028

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91. https://doi.org/10.1016/j.tics.2008.11.007

Cho, Y. S., & Proctor, R. W. (2007). When is an odd number not odd? Influence of task rule on the MARC effect for numeric classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(5), 832–842. https://doi.org/10.1037/0278-7393.33.5.832

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*(2), 132–147. https://doi.org/10.1016/j.pneurobio.2007.11.001

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and mental number magnitude. *Journal of Experimental Psychology: General*, *122*, 371–396. https://doi.org/10.1037/0096-3445.122.3.371

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. https://doi.org/10.1080/135467996387552

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial–numerical associations. *Brain and Cognition*, *72*(3), 333–336. https://doi.org/10.1016/j.bandc.2009.10.010

Fumarola, A., Prpic, V., Pos, O. D., Murgia, M., Umiltà, C., & Agostini, T. (2014). Automatic spatial association for luminance. *Attention, Perception, & Psychophysics*, *76*(3), 759–765. https://doi.org/10.3758/s13414-013-0614-y

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 32–44. https://doi.org/10.1037/0096-1523.32.1.32

Henik, A., Leibovich, T., Naparstek, S., Diesendruck, L., & Rubinsten, O. (2012). Quantities, amounts, and the numerical core system. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00186

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395. https://doi.org/10.3758/BF03202431

Hines, T. M. (1990). An odd effect: lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*(1), 40–46. https://doi.org/10.3758/BF03202644

Hoffmann, D., Mussolin, C., Martin, R., & Schiltz, C. (2014). The Impact of Mathematical Proficiency on the Number-Space Association. *PLoS ONE*, *9*(1), e85048. https://doi.org/10.1371/journal.pone.0085048

Huber, S., Klein, E., Graf, M., Nuerk, H.-C., Moeller, K., & Willmes, K. (2015). Embodied markedness of parity? Examining handedness effects on parity judgments. *Psychological Research*, *79*(6), 963–977. https://doi.org/10.1007/s00426-014-0626-9

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01795

Landy, D. H., Jones, E. L., & Hummel, J. E. (2008). Why Spatial-Numeric Associations Aren't Evidence for a Mental Number Line. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 357–362). Austin, TX: Cognitive Science Society.

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics, Herzliya, Israel*, 357–362.

Leth-Steensen, C., & Marley, A. A. J. (2000). A model of response time effects in symbolic comparison. *Psychological Review, 107*(1), 62–100. https://doi.org/10.1037/0033-295X.107.1.162

Lorch, R. F., & Myers, J. L. (1990). Regression-analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *16*, 149–157. https://doi.org/10.1037/0278-7393.16.1.149

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature, 215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Nuerk, H.-C., Bauer, F., Krummenacher, J., Heller, D., & Willmes, K. (2005). The power of the mental number line: How the magnitude of unattended numbers affects performance in an Eriksen task. *Psychology Science*, *47*(1), 34–50.

Nuerk, H.-C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC - and the MARC - (linguistic markedness of response codes) effect. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(5), 835–863. https://doi.org/10.1080/02724980343000512

Oliveri, M., Vicario, C. M., Salerno, S., Koch, G., Turriziani, P., Mangano, R., … Caltagirone, C. (2008). Perceiving numbers alters time perception. *Neuroscience Letters*, *438*(3), 308–311. https://doi.org/10.1016/j.neulet.2008.04.051

Patro, K., Nuerk, H.-C., Cress, U., & Haman, M. (2014). How number-space relationships are assessed before formal schooling: A taxonomy proposal. *Developmental Psychology*, *5*, 419. https://doi.org/10.3389/fpsyg.2014.00419

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Pinhas, M., Tzelgov, J., & Ganor-Stern, D. (2011). Estimating linear effects in ANOVA designs: The easy way. *Behavior Research Methods*, *44*(3), 788–794. https://doi.org/10.3758/s13428-011-0172-y

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Roettger, T. B., & Domahs, F. (2015). Grammatical number elicits SNARC and MARC effects as a function of task demands. *The Quarterly Journal of Experimental Psychology*, *68*(6), 1231–1248. https://doi.org/10.1080/17470218.2014.979843

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: the SMARC effect. *Cognition*, *99*(2), 113–129. https://doi.org/10.1016/j.cognition.2005.01.004

Santiago, J., & Lakens, D. (2013). Polarity correspondence does not explain the SNARC effect. In *Proceedings of the 36th Annual Conference of the Cognitive Science Society*. Austin, Texas. Retrieved from http://www.ugr.es/~santiago/SantiagoLakens-CogSci2013.pdf

Shaki, S., & Gevers, W. (2011). Cultural Characteristics Dissociate Magnitude and Ordinal Information Processing. *Journal of Cross-Cultural Psychology*, *42*(4), 639–650. https://doi.org/10.1177/0022022111406100

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Verguts, T., & Fias, W. (2008). Symbolic and Nonsymbolic Pathways of Number Processing. *Philosophical Psychology*, *21*(4), 539–554. https://doi.org/10.1080/09515080802285545

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

Chapter 8: Symbolic number comparison and number priming do not rely on the same mechanism

The present chapter has been published:

Krajcsi, A., & Szűcs, T. (2022). Symbolic number comparison and number priming do not rely on the same mechanism. Psychonomic Bulletin & Review. https://doi.org/10.3758/s13423-022-02108-x

Symbolic number comparison and number priming do not rely on the same mechanism

Attila Krajcsi, Tamás Szűcs

Department of Cognitive Psychology, Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

In elementary symbolic number processing, the comparison distance effect (in a comparison task, the task is more difficult with smaller numerical distance between the values) and the priming distance effect (in a number processing task, actual number is easier to process with a numerically close previous number) are two essential phenomena. While a dominant model, the approximate number system model, assumes that the two effects rely on the same mechanism, some other models, such as the discrete semantic system model, assume that the two effects are rooted in different generators. In a correlational study, here we investigate the relation of the two effects. Critically, the reliability of the effects are considered, therefore, a possible null result cannot be attributed to the attenuation of low reliability. The results showed no strong correlation between the two effects, even though appropriate reliabilities were provided. These results confirm the models of elementary number processing that assume distinct mechanisms behind number comparison and number priming.

Keywords: Approximate Number System; Discrete Semantic System; comparison distance effect; priming distance effect;

Introduction

In the field of numerical cognition, the approximate number system (ANS) is believed to be a fundamental representation to understand and solve mathematical problems. For example, it was proposed that higher sensitivity of the ANS is related to better math performance in school (Halberda et al., 2008; Schneider et al., 2017), the representation may play a role in the initial acquisition of the symbolic numbers (Mussolin et al., 2012; Piazza, 2010; Rousselle et al., 2004; Wagner & Johnson, 2011), or the impairment of this system may lead to developmental dyscalculia, a math-specific learning deficit (Molko et al., 2003; Piazza et al., 2010; Price et al., 2007).

The ANS is a number representation that obeys Weber's law. For example, in a number comparison task (larger of two values should be chosen), number pairs with a ratio closer to 1 are harder to process, i.e., the responses are slower and more error-prone, an effect termed ratio effect (Dehaene, 2007). Two other comparison effects are believed to reflect this ratio effect as well. The comparison distance effect (CDE, worse performance when the two to-be-compared values are numerically closer – while the sizes of the pairs are the same) and the comparison size effect (worse performance when the values are larger – while the distances of the pairs are the same) are thought to be two different ways to measure the ratio effect (Dehaene, 2007; Moyer & Landauer, 1967). In other words, these two effects can also be considered as the artifacts of the ratio effect.

An often cited possible implementation of the ANS is a representation where the stored values are noisy, and the noise is proportional to the to-be-stored value, i.e., larger numbers are noisier (Figure 45). In this model, the difficulty of a comparison task is proportional to the overlap between two number representations: The smaller the overlap is, the more discriminable the two values are and the easier the comparison task is. Additionally, there could be individual differences in how precise the system is (Halberda et al., 2012). Mathematically, the standard deviation of the noisy representation for a specific value (described as the Weber fraction parameter) is the precision or sensitivity of the ANS. In the model, higher sensitivity (i.e., smaller Weber fraction) leads to smaller representational overlap between the representation of two values, and e.g., to more efficient comparison performance.

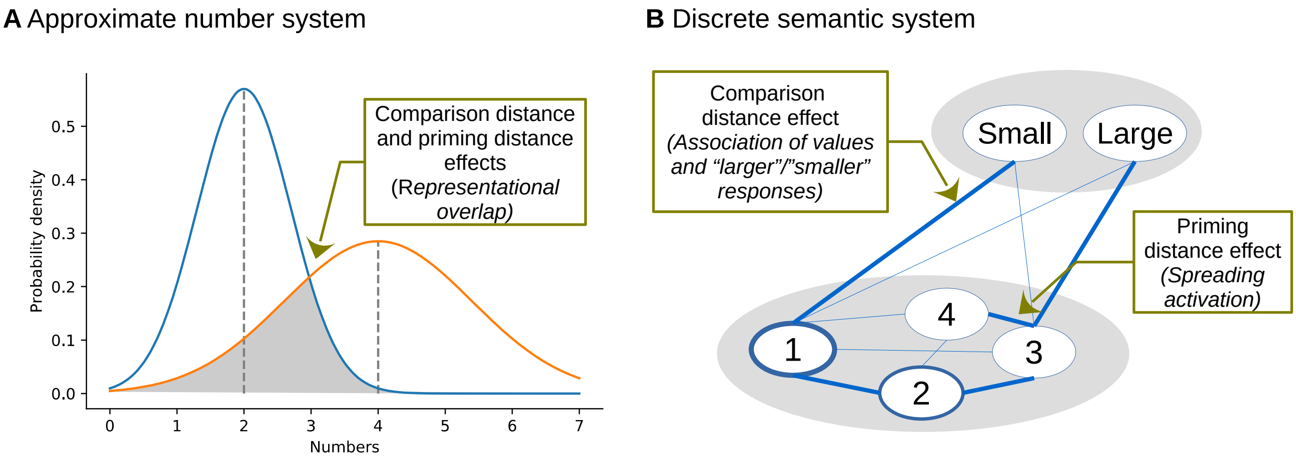


Figure 45 ***A*** *A possible implementation of the approximate number system (ANS) representation.* ***B*** *The discrete semantic system (DSS) representation can explain comparison distance and priming distance effects. Note that the connectionist model of Verguts provides a functionally similar solution (Verguts et al., 2005).*

This simple representation is believed to account for a series of other phenomena. Among others, the ANS model can explain the numerical priming distance effect (PDE). In the numerical PDE, the processing of a former value can enhance the processing of a later value, and the closer the two values are the stronger the priming effect is (Koechlin et al., 1999; Reynvoet & Brysbaert, 1999). According to the ANS account, the strength of the priming effect depends on the overlap of the prime and target values (Dehaene, 2004; Koechlin et al., 1999). In line with the ratio-based nature of the representation, the PDE is also an aspect of the ratio effect.

Note that while, in the CDE, the performance is worse when the distance is smaller, in the PDE, the performance is better for smaller distance (e.g., see Figure 46 in the present results). This difference in itself may hint that the two effects do not rely on the same mechanism. However, according to a possible explanation, the single ANS may account for the opposing effect directions: When numbers are presented simultaneously, the representations may interfere, which causes worse performance with larger representational overlap and when the numbers are presented consecutively, the prime representation may help the activation of the target representation (Koechlin et al., 1999). Note that this hypothesis should be confirmed: Consecutive presentation of the numbers of a pair should reverse the distance effect in other paradigms as well. Contrarily, in comparison tasks with serially presented values of the pairs, we found CDE-like distance effect both in Arabic number comparison and in nonsymbolic dot comparison (unpublished data). Overall, an ANS model should account for the opposing directions of the CDE and PDE slopes, which account may not be entirely satisfactory at the moment.

While the ANS model is a dominant model in the numerical cognition area, there are other explanations how comparison distance and priming distance effects are generated. In an alternative model, a connectionist network may account for a series of elementary number processing phenomena (Verguts et al., 2005). Importantly, it has been proposed that while the CDE is rooted in the connections between the number nodes and the “larger” responses, the PDE is caused by the spreading activation between the number nodes (Verguts et al., 2005). In another alternative explanation, similar to the connectionist network above, the Discrete Semantic System (DSS) model assumes that numbers and related concepts are stored in a network of nodes and relevant effects can be generated by this simple architecture (Figure 45, right) (Krajcsi et al., 2022). In line with the connectionist model, the DSS model proposes that the CDE may be rooted in the connection of the number nodes and the “small”–“large” nodes: Smaller numbers are more strongly associated with the “small” label than the larger numbers and, in a comparison task, numbers with larger distance are easier to process because their association with “large” and “small” labels is more dissimilar. This hypothesis has been confirmed empirically in studies where the association between numbers and the “larger” response was manipulated in a comparison task and the distance effect followed the association of the numbers and the “larger” response instead of the values of the numbers (Kojouharova & Krajcsi, 2018; Krajcsi & Kojouharova, 2017). On the other hand, in the DSS model, it may be reasonable to assume that the priming effect is generated by the spreading activation between the number nodes (Figure 45). The DSS is similar to the connectionist model by Verguts et al. (2005) in many aspects and the two models can also be considered as complementary descriptions of the same system.

To summarize, while the widely discussed ANS model assumes that the CDE and the PDE are rooted in the same mechanism (representational overlap in a noisy number representation), alternative models assume that CDE and PDE are generated by different mechanisms: CDE is the result of the connections between the values and “larger” and “smaller” labels and PDE is rooted in the spreading activation between the representations of the values.

Empirical results on the possible common sources of number comparison and number priming

Several studies investigated the possible common causes of the CDE and PDE. A first group of relevant empirical studies investigated the correlation of the CDE and the PDE. According to the ANS account, because both the CDE and the PDE rely on the same representation and the performance in both effects depends on the Weber fraction (i.e., sensitivity) of the ANS, the correlation coefficient of the two effects ideally should be 1 (see a similar approach in Krajcsi, 2017, where the expected perfect correlation between nonsymbolic numerical comparison effects has been observed). This prediction was not confirmed by empirical studies: Significant non-zero correlation was found neither in children with symbolic numbers (Reynvoet et al., 2009) nor in adults with nonsymbolic values (Sasanguie et al., 2011). Although these results may seem to confute the ANS account, an essential limitation of correlational studies should be considered. Observed correlations are attenuated by the reliability of the variables: The noisier the variables, the smaller the observed correlation can be (Spearman, 1910). This also means that the true correlation is equal to or larger than the observed correlation, and the difference between the true and measured values depends on the reliability of the variables. If the variables have low reliability, it may completely obscure the correlation. In the cited studies, if the reliability of either the comparison or the priming distance effect is low, it is possible that while the true correlation coefficient is 1, the observed correlation is as low as 0. Importantly, Sasanguie et al. (2011) reported the reliabilities of the effects, and while the reliability of the CDE was partially acceptable (0.40 correlation between the first and second halves of the comparison task), the reliability of the PDE was very low (a non-significant 0.21 coefficient). This confirms that when measuring the correlation between the CDE and PDE indexes, the reliability of these indexes can be low, which can entirely obscure the true correlation. To summarize, although empirical studies demonstrate that the comparison and priming distance effects do not correlate, their results cannot be conclusive because one may not know whether the observed low correlation is the result of the low true correlation or the result of the low reliability of the variables or both.

A second group of relevant studies investigates whether the priming effect shows the ratio effect. The ANS assumes an asymmetric priming effect around the target value, i.e., a prime smaller than the target should evoke a smaller priming effect compared to a larger prime with the same distance because the smaller prime has smaller noise than the larger prime, which leads to smaller representational overlap with the smaller prime than with the larger prime. In a review of several empirical works, Verguts et al. (2005) found that the PDE is symmetric in symbolic stimuli, which does not confirm the ANS account. In other words, the priming effect depends only on the distance of the prime and the target values but not on the ratio of the values, which means that while there is a PDE, it is not rooted in a possible priming ratio effect, but it is a separate effect.

A third line of evidence argues that the CDE and PDE do not rely on the same mechanism because the two effects dissociate in letter processing. It was found that while, in symbolic number comparison task, both comparison and priming distance effects can be observed, in letter comparison (i.e., is the presented letter before or after a reference letter in the alphabet), only the CDE can be observed but not the PDE (Opstal et al., 2008). Importantly, this argument assumes that number and letter comparisons rely on the same or same type of representations which assumption is backed by the similar nature of the distance effect in number and letter comparison. However, it is possible that similar effects have different generators. For example, Vigliocco et al. (2002) demonstrated a semantic distance effect in a picture naming task where a psychophysical representation (i.e., a representation obeying Weber's law) is unlikely, while nonsymbolic number CDE most likely relies on a psychophysical representation (Krajcsi et al., 2018).

A forth type of evidence consists of a dissociation in the neural background of the CDE and PDE (Zhang et al., 2016). In an fMRI study, it was demonstrated that while the CDE relies more heavily on right parietal areas, the PDE more heavily uses left parietal areas. However, critical aspects of the tasks were admittedly not aligned: While, in the comparison task, response selection was needed (participants responded whether the first or the second number presented consecutively was larger), in the priming task, no such selection was required (participants had to press a button only if second number matched the first number). For this reason these results may not be conclusive.

Overall, while there are several works investigating whether the PDE may be backed by the ANS or whether the CDE and the PDE are related or dissociated, there are methodological issues that question the conclusions of several of those works.

The aim of the present study is to contrast the theoretical accounts of the CDE and PDE by measuring the correlation of the CDE and the PDE slopes. To consider the possible attenuation effect, the reliability of the appropriate indexes are measured here and we intend to ensure that they are appropriate.

Methods

In two comparison tasks, the comparison and the priming distance effects were measured, and the correlation of these effects was calculated while the reliability of the effects was considered.

Participants

Pilot studies (see below in the Stimuli and procedure subsection) indicated that with the planned paradigms a reliability of at least 0.51 for the CDE can be reached. Note that the present paradigm includes 4 times the number of trials compared to the pilot paradigm, therefore, we expect considerably higher reliabilities in the present data. According to the pilot study, a reliability of 0.51 for the PDE can be reached. Reliability was measured with the Spearman–Brown corrected even-odd split-half reliability values. If the true correlation coefficient is 1 as predicted by the ANS account, the observed correlation coefficient with the given reliabilities should be 0.51.[22] To reach 95% power for this observed correlation, at least 44 participants are needed.

22 By rearranging the formula for attenuated correlation coefficients proposed by Spearman (1904), we can calculate the expected measured correlation coefficient, given the supposed true correlation and the reliabilites of the variables as $r_{observed} = r_{real} \cdot \sqrt{reliability_x \cdot reliability_y}$, where $r_{observed}$ is the observed correlation, r_{real} is the supposed real correlation, and $reliability_x$ and $reliability_y$ are the Spearman-Brown corrected split-half reliabilities of variable X and Y, respectively. Here, the expected observed correlation is $1 \cdot \sqrt{0.51 \cdot 0.51}$, which is 0.51.

Eighty-four university students from various majors completed all sessions for partial course credit, 64 of them females, with a mean age of 21.9 years, SD was 4.6 years. A few additional participants who made random guesses in any of the sessions were formerly excluded. The study was approved (201710) by the ethics committee of the Psychology Institute, ELTE Eötvös Loránd University, Hungary.

Stimuli and procedure

In the task that measured the CDE, two single-digit numbers appeared on the two sides of the screen and the participants had to choose the larger one by pressing the appropriate response button. In the task that measured the PDE (we term this task here the priming task, even if both tasks included comparisons as tasks), a single-digit number appeared in the middle of the screen and participants had to choose whether the number is smaller or larger than 5 by pressing the appropriate response button. In both tasks, the numbers were visible until response. After the response, a blank screen was visible for 700 ms and an auditory feedback was given during the blank screen.

In both the comparison and the priming tasks, numbers between 1 and 4, and between 6 and 9 were used. In the comparison task, all possible number pairs with different values were presented (56 possible number pairs) 40 times, resulting in 2240 trials. In the priming task, the same 56 number pairs were used as in the comparison task, but the numbers of a pair were presented in two trials, where the first value was later considered as the prime number and the second value was the target. In the priming task, the 112 (i.e., 56 number pairs where the two values of a pair were presented in separate trials) numbers were repeated 120 times, resulting in 13440 trials. For both tasks, the trials within a session (see below) were randomized (with the constraint that, in the priming task, the prime-target values were presented in consecutive trials).

The whole experiment was divided into 5 sessions, approximately 1 hour each. The first session included the comparison task, while the remaining 4 sessions included the priming task, where each priming task session included trials with all 56 number pairs presented 30 times.

Although this long procedure was demanding for the participants, we used this version because it could provide acceptable reliabilities for the CDE and PDE. While, in typical paradigms that are used in the literature, the reliability of the CDE is acceptable, the reliability of the PDE is rather low (Gilmore et al., 2011, 2014; Sasanguie et al., 2011). After a series of pilots, we found that it is only the number of trials that can improve reliability further. Our preparatory studies showed that for the CDE (with all number pairs repeated 10 times instead of 40 times as used here) reliability was 0.38 for the error rates and 0.51 for the reaction time data (Spearman–Brown corrected even-odd split-half reliability). We considered that multiplying the number of trials by four will lead to a satisfying reliability for the CDE slope index. For the PDE, with similar parameters as used in the present study, the reliability was 0.79 for the error rates and 0.51 for the reaction time. In former pilot studies, we could not find appropriate parameters for the paradigm that could have lead to acceptable reliability in a single session measurement of PDE, where the session could not be longer than 60 minutes to avoid fatigue.

The data were collected online on the Cognition platform (www.cognition.run), the script was written in jsPsych (de Leeuw, 2015) using the jspsych-psychophysics plugin (Kuroki, 2021). Participants received detailed instructions on how to provide optimal circumstances for the data collection. The time interval between two consecutive sessions was between 2 hours and 3 days.

Analysis

While the distance (or the ratio) effect slope is widely used in the literature as an approximation for the ANS sensitivity, its use is not recommended in correlational studies. A main problem is that the relationship of the distance effect slope and the Weber fraction is not only non-linear but it is non-monotonic: Depending on the specific ratios of the stimuli and the Weber fractions of the participants, in some cases, larger distance effect slope means larger Weber fraction, while in some other cases, smaller Weber fraction (Chesney, 2018). This bias can fundamentally reduce the observed correlations (Krajcsi, 2020). (For a detailed explanation of this non-trivial relation of the effect slopes and the Weber fraction, see Chesney (2018) and Krajcsi (2020).) One way to overcome this issue in the present study was to use the same value-pairs in both tasks. This way since the stimuli (and the relevant ratios or distances) are the same across the two tasks and the Weber fractions of the participants are also the same in the two tasks, the Weber fractions are transformed in the same way to the CDE and PDE. Therefore, if there is a correlation between the Weber fractions behind the CDE and PDE (i.e., because they are the same Weber fraction of a single ANS), then the CDE and the PDE should correlate as well.

In the priming task, only the congruent trials (i.e., either both the prime and target numbers are smaller than 5 or both of them are larger than 5) were used, and incongruent trials were excluded from the analysis because incongruent trials show an interference effect that masks the priming effect (Reynvoet et al., 2002). The same restrictions were applied to the comparison task to avoid the issue rooted in the non-monotonic relation between the distance effects and the Weber fraction (Chesney, 2018; Krajcsi, 2020). In the priming task, only the target trials were analyzed.

The CDE and the PDE were calculated for both the error rates and the reaction time data. Mean error rate and median reaction time for the correct responses were calculated for each distance and each participant. Distance effect slopes were calculated with linear regressions where the regressor was the distance and the outcome variable was the performance (error rate and reaction time for the CDE and PDE) for each participant. The output of these analyses were the CDE and PDE slopes for error rates and reaction times for each participant (i.e., four slope values per participant).

The reliability of the four indexes (i.e., CDE and PDE for error rates and reaction times) were investigated. It is important to highlight that even if the reliabilities of similar tasks have been reported in other works (e.g., Sasanguie et al., 2011), those results may not be relevant here because reliability depends on the specific parameters of a paradigm (e.g., the number of trials is a strong predictor) and on the population (since the commonly used test-retest correlation is a relative index of the true variance of the variable in the population’s total variance) (Lindskog et al., 2013). Therefore, reliability indexes should be reported in correlational studies unless the paradigm and the population are similar to other paradigms and populations for which the reliability is known. To calculate the reliability in the present study, even-odd split-half reliability together with the

Spearman–Brown prediction formula were applied (Spearman, 1910). Because the stimuli were randomized, when trials are split into even and odd trials, some conditions (in the present analysis, distances) may include more trials in the even half compared to the odd half or the other way around. Decreased number of trials in the even or the odd half of the condition can lead to lowered observed reliability. To overcome this problem, trials of a task were first sorted according to the relevant conditions (i.e., distances), and an even-odd split was applied on the ordered data, therefore, the size of the even and odd halves were equal (or the difference was only 1 if the total number of trials were odd) in all cells. Note that the split-half reliability of a multi-session data is conceptually a mixture of a single-session split-half index and a test-retest reliability index: Similar to the test-retest index, it includes the variability of the changes between sessions and similar to the single-session split-half method, it splits the data not following the sessions but in a more gradual way.

Correlation for the reliability and for investigating the CDE–PDE relationship was calculated not only with the Pearson correlation coefficient, but also with Spearman's rank correlation coefficient because the latter is not sensitive to outliers. Therefore, similar results with Pearson's r and Spearman's r_s denote that the results with Pearson's r are not the result of some of the statistical artifacts.

For the analysis custom Python scripts, LibreOffice v7.2 (2021) and CogStat v2.1 (Krajcsi, 2021) were used.

Results and discussion

The raw data of the experiment reported here is available at https://osf.io/bs94q/. The experiment was not preregistered.

CDE was observed in the comparison task, both for the error rates and reaction time (Figure 46; for the error rate the slope mean was -1.4% with 1.3 SD, for the reaction time, the mean slope was -36 ms with 19 SD; the slopes significantly differed from 0 both in error rates, Wilcoxon signed-rank test: $T = 23.5$, $p < .001$, and in reaction times, Wilcoxon signed-rank test: $T = 0.00$, $p < .001$). Similarly, PDE was present in the priming task, both for error rates and reaction times (Figure 46; for the error rate the slope was 1.1% with 0.9 SD, for the reaction time the slope was 12 ms with 8 SD; Wilcoxon signed-rank test: $T = 94.00$, $p < .001$, Wilcoxon signed-rank test: $T = 15.00$, $p < .001$, respectively).

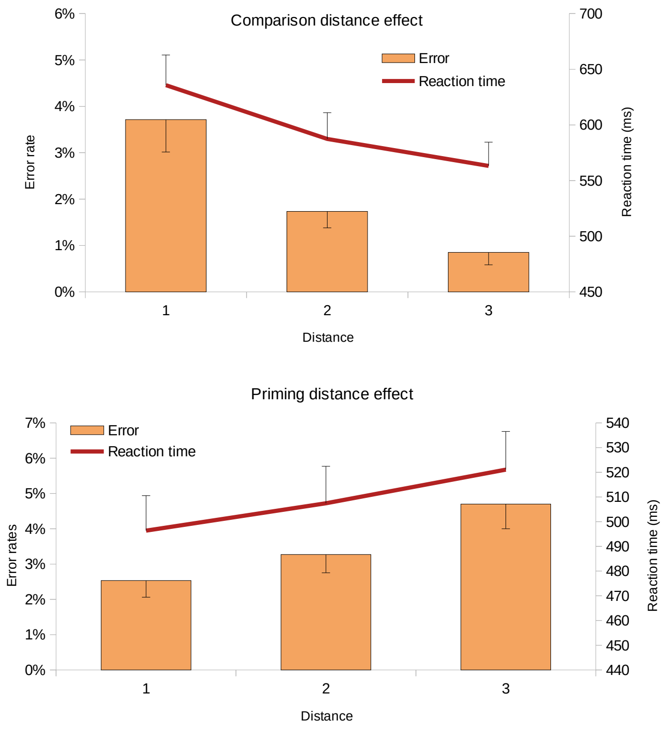


Figure 46 Comparison distance effect (left) and priming distance effect (right) for the error rates and reaction time. Error bars represent 95% confidence intervals.

The reliability of the 4 slope indexes were calculated (Table 45). All of the reliabilities were satisfying, the Spearman-Brown corrected split-half reliabilities were between 0.72 and 0.87. Similar Pearson and Spearman correlation values demonstrate that the relatively high correlations are not the result of outliers. Relying on these reliabilities, one can consider the attenuation in the prediction of the ANS model. While the ANS model predicts a true correlation value of 1 between the CDE and PDE slopes, the measured correlation should be lower because of the smaller than 1 reliabilities. According to the equation in footnote 22, the expected measured correlation is 0.8 for the CDE and PDE error rate slopes (i.e., $1\cdot\sqrt{0.868\cdot0.738}$) and 0.78 for the CDE and PDE reaction time slopes (i.e., $1\cdot\sqrt{0.841\cdot0.72}$). In the following analysis, it is investigated whether the observed correlation equals these predicted correlations. Statistically, it is investigated whether the confidence intervals of the measured correlations include these predicted values (Cumming, 2014).

	Reliability – Pearson correlation	Reliability – Spearman correlation
CDE error rates	r_{SB} = 0.868 r = 0.767, [0.661, 0.843]	r_s = 0.685, [0.551, 0.784]
CDE reaction times	r_{SB} = 0.841 r = 0.725, [0.604, 0.813]	r_s = 0.693, [0.562, 0.790]
PDE error rates	r_{SB} = 0.738 r = 0.585, [0.424, 0.710]	r_s = 0.527, [0.353, 0.666]
PDE reaction times	r_{SB} = 0.720 r = 0.563,[0.396, 0.694]	r_s = 0.462, [0.275, 0.616]

Table 15 Reliability of the CDE and PDE indexes. Cells include the Spearman-Brown prediction correlation values for the Pearson correlation and the Pearson and Spearman correlation coefficients with 95% confidence intervals. All correlations significantly differ from 0, $p < 0.001$.

The correlation values of the CDE and PDE slopes did not reach the prediction of the ANS model (see the scatter plots in Figure 47). Although the CDE and PDE slopes for the error rates correlated significantly (i.e., the correlation coefficient differed from 0), the confidence interval did not include the predicted -0.8 value ($r(82) = -0.31$, $p = .005$, [-0.488, -0.097]; $r_s(82) = -0.38$, $p < .001$, [-0.554, -0.186]). For the reaction time data, the CDE and PDE slopes were not different from zero significantly, and similar to the error rates data, the confidence interval did not include the predicted -0.78 value ($r(82) = -0.19$, $p = .092$ [-0.384, 0.031]; $r_s(82) = -0.15$, $p = .167$ [-0.355, 0.064]).

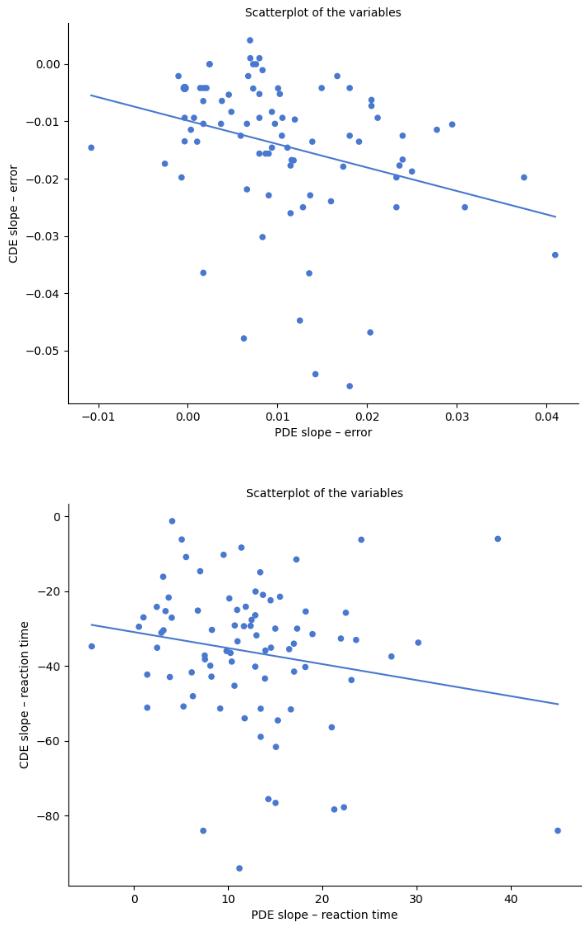


Figure 47 Scatter plots of the CDE and PDE in error rates (left) and reaction time (right)

The CDE was measured in a separate session, and the PDE was measured in four additional sessions. Can the fluctuation between the sessions be responsible for the observed low correlation? This possibility is not likely. The observed PDE reliability was based on a split-half index that aggregated the data of four sessions, and the relatively high PDE reliability suggests that the fluctuation between the sessions could not be high. If one assumes that the between-session fluctuation is similarly low in the CDE, then the low correlation between the CDE and PDE slopes cannot be attributed to a high fluctuation between the sessions. Alternatively, if one assumes that the fluctuation is much higher in the CDE than in the PDE, then this assumption contradicts the ANS model: The ANS account suggests that both the CDE and PDE effects rely only on the ANS sensitivity, therefore, their fluctuations cannot be different. To sum up, the results cannot be attributed to the fluctuation of the slopes between the sessions unless the ANS account is already incorrect.

While the present study carefully ensured that the reliabilities of the measured slope indexes are appropriate to observe a possible high correlation, the correlations were considerably lower than predicted by the ANS model. The present results are not in line with the ANS account of a common mechanism underlying symbolic comparison and priming.

General discussion

The main question of the present study was whether the CDE and PDE rely on the same mechanism as predicted by the ANS model, or whether they are independent as predicted by the connectionist model of Verguts et al. (2005) or by the DSS model (Krajcsi et al., 2022). Unlike previous similar correlational studies, the current work considered the reliability of the CDE and PDE. This is critical because low reliabilities may attenuate the observed correlation, and an observed correlation around zero cannot be conclusive. Here, the reliabilities were satisfying, still, the observed correlations were considerably lower than predicted by the ANS model. These results question the ANS account of a common mechanism behind the comparison distance effect and priming distance effect. On the other hand, the results are in line with models that assume distinct mechanisms behind number comparison and number priming, such as the connectionist model of Verguts et al. (2005) or the DSS model.

In the theoretical motivation of this work, we considered the most widely cited, coherent, and mathematically relatively detailed version (Dehaene, 2007) of the ANS model. However, there are several approaches in the literature that question various details of the ANS model and provide related modifications (e.g., see the reviews of Clarke & Beck, 2021; Kadosh & Walsh, 2009; Leibovich et al., 2017). Importantly, we are not aware of any specific ANS variant that may have alternative predictions about the correlation of the CDE and PDE.

The present results are in line with former findings. Our results are in line with the symmetric priming effect (Verguts et al., 2005), i.e., that the size of the priming effect depends on the distance of the prime and the target values but not on the ratio of them. Both findings question the ANS model and confirm the network-based models of elementary symbolic number processing. The present results may be in line with additional findings, such as the dissociation of CDE and PDE in letter processing (Opstal et al., 2008), or the neural dissociation of the CDE and PDE (Zhang et al., 2016), although these latter findings may need additional confirmation. Also, a systematic empirical test is needed for any explanation that tries to account for the opposing direction of the CDE and PDE slopes, such as whether simultaneous and consecutive presentation of the numbers of a pair can change the direction of the distance effect – otherwise the opposing direction of these effects strengthen the idea that the CDE and PDE have two different generators.

Finally, it is important to highlight the essential role of reliability in correlational studies in cognitive psychology. In many cognitive studies, reliability may be low (Hedge et al., 2018). These low reliabilities attenuate the observed correlation, thus, an observed low correlation in itself cannot be conclusive: Low correlation may mean either low true correlation or low reliability or both. Moreover, when the tasks are not standardized (which is typical in cognitive psychology, where, the specific stimuli or number of trials may vary between studies), the reliability may differ from what has been published in former works since reliability depends on the properties of the design. Therefore, it is essential that correlational studies should consider reliability, and if the relevant parameters of the task or the sample are unique, then reliabilities should also be reported.

Acknowledgments

We would like to thank Mateusz Hohol and Melody Wiseheart for their comments on the manuscript. The work of Attila Krajcsi was supported by the National Research, Development and Innovation Fund (NKFI 132165). The work of Tamás Szűcs was supported by the ÚNKP-20-2 New National Excellence Program of the Ministry of Innovation and Technology from the source of the National Research, Development and Innovation Fund (ÚNKP-20-2-II-ELTE-904).

References

Chesney, D. (2018). Numerical distance effect size is a poor metric of approximate number system acuity. *Attention, Perception, & Psychophysics*, *80*(5), 1057–1063. https://doi.org/10.3758/s13414-018-1515-x

Clarke, S., & Beck, J. (2021). The number sense represents (rational) numbers. *Behavioral and Brain Sciences*, *44*. https://doi.org/10.1017/S0140525X21000571

Cumming, G. (2014). The New Statistics Why and How. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dehaene, S. (2004). The neural bases of subliminal priming. In N. Kanwisher & J. Duncan (Eds.), *Functional Neuroimaging of visual cognition* (pp. 205–224).

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Vol. XXII* (pp. 527–574). Harvard University Press.

Gilmore, C., Attridge, N., De Smedt, B., & Inglis, M. (2014). Measuring the approximate number system in children: Exploring the relationships among different tasks. *Learning and Individual Differences*, *29*, 50–58. https://doi.org/10.1016/j.lindif.2013.10.004

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *The Quarterly Journal of Experimental Psychology*, *64*(11), 2099–2109. https://doi.org/10.1080/17470218.2011.574710

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences*, *109*(28), 11116–11120. https://doi.org/10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Kadosh, R. C., & Walsh, V. (2009). Numerical representation in the parietal lobes: Abstract or not abstract? *Behavioral and Brain Sciences*, *32*(3–4), 313–328. https://doi.org/10.1017/S0140525X09990938

Koechlin, E., Naccache, L., Block, E., & Dehaene, S. (1999). Primed numbers: Exploring the modularity of numerical representations with masked and unmasked semantic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1882–1905. https://doi.org/10.1037/0096-1523.25.6.1882

Kojouharova, P., & Krajcsi, A. (2018). The Indo-Arabic distance effect originates in the response statistics of the task. *Psychological Research*. https://doi.org/10.1007/s00426-018-1052-1

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A. (2020). Ratio effect slope can sometimes be an appropriate metric of the approximate number system sensitivity. *Attention, Perception, & Psychophysics*, *82*(4), 2165–2176. https://doi.org/10.3758/s13414-019-01939-6

Krajcsi, A. (2021). *CogStat – An automatic analysis statistical software* (2.1.0) [Computer software]. https://www.cogstat.org

Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02013

Krajcsi, A., Kojouharova, P., & Lengyel, G. (2022). Processing Symbolic Numbers: The Example of Distance and Size Effects. In J. Gervain, G. Csibra, & K. Kovács (Eds.), *A Life in Cognition: Studies in Cognitive Science in Honor of Csaba Pléh* (pp. 379–394). Springer International Publishing. https://doi.org/10.1007/978-3-030-66175-5_27

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analogue number system: Different symbolic and nonsymbolic numerical distance and size effects. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00124

Kuroki, D. (2021). A new jsPsych plugin for psychophysics, providing accurate display duration and stimulus onset asynchrony. *Behavior Research Methods*, *53*(1), 301–310. https://doi.org/10.3758/s13428-020-01445-w

Leibovich, T., Katzin, N., Harel, M., & Henik, A. (2017). From “sense of number” to “sense of magnitude”: The role of continuous magnitudes in numerical cognition. *Behavioral and Brain Sciences*, *40*. https://doi.org/10.1017/S0140525X16000960

Lindskog, M., Winman, A., Juslin, P., & Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00510

Molko, N., Cachia, A., Rivière, D., Mangin, J.-F., Bruandet, M., Le Bihan, D., Cohen, L., & Dehaene, S. (2003). Functional and Structural Alterations of the Intraparietal Sulcus in a Developmental Dyscalculia of Genetic Origin. *Neuron*, *40*(4), 847–858. https://doi.org/10.1016/S0896-6273(03)00670-6

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, *1*(1), 21–31. https://doi.org/10.1016/j.tine.2012.09.003

Opstal, F. V., Gevers, W., Moor, W. D., & Verguts, T. (2008). Dissecting the symbolic distance effect: Comparison and priming effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*(2), 419–425. https://doi.org/10.3758/PBR.15.2.419

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41. https://doi.org/10.1016/j.cognition.2010.03.012

Price, G. R., Holloway, I., Räsänen, P., Vesterinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, *17*, R1042–R1043.

Reynvoet, B., & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition*, *72*, 191–201.

Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus-response associations may be semantically mediated. *Psychonomic Bulletin & Review*, *9*(1), 107–112.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children’s representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*(4), 480–489. https://doi.org/10.1016/j.jecp.2009.01.007

Rousselle, L., Palmers, E., & Noël, M.-P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology*, *87*(1), 57–84. https://doi.org/10.1016/j.jecp.2003.10.005

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychologica*, *136*(1), 73–80. https://doi.org/10.1016/j.actpsy.2010.10.004

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3). https://doi.org/10.1111/desc.12372

Spearman, C. (1910). Correlation Calculated from Faulty Data. *British Journal of Psychology, 1904-1920*, *3*(3), 271–295. https://doi.org/10.1111/j.2044-8295.1910.tb00206.x

The Document Foundation. (2021). *LibreOffice* (7.2) [Computer software]. https://www.libreoffice.org/

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review, 12*(1), 66–80. https://doi.org/10.3758/BF03196349

Vigliocco, G., Vinson, D. P., Damian, M. F., & Levelt, W. (2002). Semantic distance effects on object and action naming. *Cognition*, *85*(3), B61–B69. https://doi.org/10.1016/S0010-0277(02)00107-5

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition*, *119*(1), 10–22. https://doi.org/10.1016/j.cognition.2010.11.014

Zhang, L., Cai, F., Chen, C., & He, Q. (2016). Different Neural Mechanisms for the Comparison and Priming Distance Effects: An fMRI Study. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01633

Chapter 9 Conclusions and outlook

Conclusions

The main question of the present project is whether the ANS can be a key representation supporting symbolic number handling and whether it may be responsible for effects observed in simple symbolic numerical tasks. We also provided an alternative account assuming that simple symbolic number processing is backed by the DSS.

A series of tests were provided to contrast the ANS and the DSS accounts, and all of these tests revealed results that cannot be explained by the ANS account but are consistent with the DSS account. None of our published or unpublished (pilot, work in progress, etc.) studies indicated that the ANS model could better account for the new empirical results than the DSS model.

Appropriate accounts for symbolic number processing

In our works, similar to the previous chapters, we usually argue that the results indicate that the ANS cannot be responsible for key phenomena of symbolic number processing but the DSS can. In personal communications, reviews, and other formats, some colleagues raise that if some details of the DSS model seem to be incorrect then that is an argument for the ANS model being correct. However, this reasoning is inaccurate.

The appropriateness of the ANS and the DSS models can be evaluated differently. Regarding the ANS model a defining feature of the ANS model is that it is a psychophysical model, obeying Weber’s principle (Algom, 2021; Dehaene, 2007; Moyer & Landauer, 1967). While there are various ideas about how these properties may emerge and work (Burr & Ross, 2008; Dehaene, 2007; Dehaene & Changeux, 1993; Stoianov & Zorzi, 2012; Testolin et al., 2020), one key statement about the ANS is that it is an evolutionarily old, culture-independent representation that is present in non-human animals and infants, and, critically, it is working according to Weber’s principle (Brannon & Merritt, 2011; Dehaene, 1997; Feigenson et al., 2004; Odic & Starr, 2018; Parrish & Beran, 2022). Some details of the ANS model can be modified, but psychophysics-like processing is a keystone of the model; if one changes the key details, then it is not the ANS model anymore. Regarding the empirical tests provided here, many of these results suggest that simple symbolic number comparison is not a psychophysical process. For example, the size and the distance effects can be manipulated independently of each other; they can substantially be influenced by statistical features of the stimuli; they depend on different statistical features of the stimuli (the comparison size effect dominantly depends on frequency of the stimuli, the comparison distance effect relies mainly on the conditional frequency of the number pairs); psychophysical model provides a biased prediction of small ratio comparisons, etc. (e.g., Hohol et al., 2020; Krajcsi, 2017; Krajcsi et al., 2016, 2018; Krajcsi & Kojouharova, 2017). These are robust, easy-to-replicate, simple phenomena that the ANS model cannot account for. Seeing the long list of phenomena that contradict the psychophysical account, it is difficult to maintain the dominant role of the ANS model. While some phenomena related to symbolic number comparison seemingly support the ANS model, there are too many disturbances around the model that cannot be resolved. Importantly, these issues are problematic independent of what alternative model is provided:

Independent of whether the DSS model is correct or not, the ANS model does not seem to be correct since it contradicts many simple, robust phenomena. To summarize, the ANS as a model of symbolic comparison and base of simple symbolic number processing seems to be incorrect, and it seems to be incorrect, independent of whether the DSS model is correct or not.

Regarding the DSS model, the results that we have considered so far suggest that the model is consistent with both previous results found by other labs and recent results found by our lab. In that sense, the DSS model seems to be correct. However, one may never know if new results or neglected previous results will falsify the model. This possibility cannot be excluded in the case of the DSS model either. Note that the DSS model is compatible with many previous models (e.g., Henik & Tzelgov, 1982; Hines, 1990; Leth-Steensen et al., 2011; Pinhas & Tzelgov, 2012; Proctor & Cho, 2006; Rinaldi & Marelli, 2020; Verguts et al., 2005; Verguts & Van Opstal, 2014) in the sense that those models and the DSS have several overlapping features and they have similar predictions. The present argument can be generalized to those models as well. Still, the DSS model seems to be consistent in itself, and it is supported by many empirical results; therefore, it can be considered correct. Let us highlight again that if future works find that the DSS model is incorrect, it does not mean that the ANS model can be correct again, but it means that another account should be found.

Outlook

The role of the DSS in higher-level mathematical thinking

In the ANS model, the ANS is often considered the foundation of mathematical thinking. Better ANS acuity can lead to better mathematics grades in school and better mathematical thinking in everyday life (Libertus et al., 2011; Szkudlarek & Brannon, 2017). Also, improving ANS acuity may in itself lead to better mathematical thinking (Park & Brannon, 2013, 2014). In sum, ANS is a keystone in understanding numbers and mathematical thinking in general.

If the ANS model of symbolic number processing seems to be incorrect, and the DSS or other model takes its place, one may think that the representation of the new model may support a similar fundamental role. However, this is not necessarily the case. As explained above, for example, in the symbolic number comparison, the distance and size effects are statistical effects. In other words, unlike in the ANS model, in the DSS model, the comparison distance and size effects do not reflect semantic processing, but rather, they reflect statistical learning. While, in general, statistical learning is fundamental in various levels of cognition, the conceptual understanding of numbers and math directly may rely on other mechanisms (see below some specific representations that may take this role). For this reason, we speculate that it is unlikely that the DSS is a fundamental part of higher-level mathematical thinking, although it may be related to some aspects of mathematical problem-solving, for example, because DSS-related phenomena may reflect statistical learning or because the connections of the nodes may hold part of the meaning (see a more detailed discussion of this latter example below).

The role of the ANS in higher-level mathematical thinking

In the previous subsection, the ANS account of symbolic number processing was considered. However, on a more general level, the hybrid ANS-DSS framework proposes that while the ANS model may be incorrect for simple symbolic number processing, it may account for nonsymbolic processing. While above, we argued that the ANS is unlikely to back simple symbolic number processing and, consequently, cannot be the base of more complex, higher-level mathematical thinking, one may ask whether the ANS as a supporter of the nonsymbolic number processing may influence higher mathematical thinking.

There are several arguments assuming that symbol use, in general, may be anchored in nonsymbolic representations. Similarly, it is possible that the same may apply to mathematical thinking as well. For example, it has been reported that while many calculation errors are related to performing the calculation “blindly” (i.e., without considering the meaning of the symbolic number manipulation, for example, in a multi-digit multiplication), preliminary estimation of the results may help to avoid such calculation errors (Sternberg & Ben-Zeev, 1996). One possible mechanism for such preliminary estimation is the nonsymbolic number processing-based ANS (Sella et al., 2021). Also, additional mediator mechanisms could be imagined that highlight the role of the nonsymbolic number processing-based ANS in higher-level symbolic mathematical thinking. Overall, it is possible that the ANS supports higher-level mathematical thinking, and it is an empirical question whether this relation can be observed.

One commonly used research method to investigate the role of the ANS on other mathematical cognition processes is to measure the ANS acuity (the Weber fraction or a proxy of it) and correlate it with higher-level mathematical tasks. One critical issue is that the Weber fraction measurement can often be unreliable or invalid (Dietrich et al., 2015; Krajcsi et al., 2024). There are many issues regarding the Weber fraction measurements, and most measurement methods cannot be correct (Dietrich et al., 2015; Krajcsi et al., 2024). Consequently, most reported results in the literature must be suboptimal, and those reports must be inconclusive as well. To overcome these shortcomings, more effort is needed to find the appropriate methods to measure the ANS acuity and then to investigate the potential role of the nonsymbolic number processing-based ANS in higher-level symbolic mathematical thinking.

The role of other representations in higher-level mathematical thinking

In recent decades, the ANS has been a key representation that has been investigated in the literature, and a huge number of studies have explored its role in higher-level mathematical thinking. Here, we argued that while it is possible that a nonsymbolic number processing-based ANS may have some role in mathematical problem-solving, its potential role is much more limited than it was assumed previously. We also argue that while the DSS model may have some role in mathematical thinking, this role must also be limited since most identified DSS-based processing relies on statistical features of the stimuli instead of the semantic properties of the numbers. Still, higher-level mathematical problem-solving is possible in most literate humans, and appropriate representations should support this capability. If the ANS and the DSS have a limited role in this higher-level mathematical problem-solving, what other representations may be relevant?

To provide a hypothetical framework of number understanding and number representation, again, inspiration is taken from linguistic and conceptual models. We propose that number meaning and conceptual understanding of it has three main sources: perceptual anchoring, connections within the conceptual system, and dynamic semantic understanding (Figure 48).

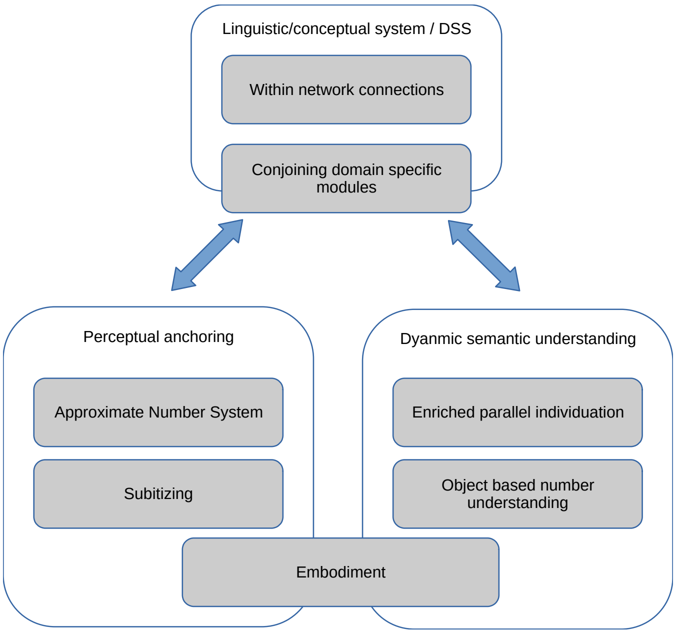


Figure 48. Hypothetical outline for possible sources of numerical meaning.

Regarding perceptual anchoring, one possible source of it may be the ANS, as explained above. Another source of perceptual anchoring may be subitizing, a relatively fast and accurate enumeration of perceptual objects (Kaufman et al., 1949). It is assumed that subitizing is supported by a pattern recognition mechanism (Krajcsi et al., 2013; Mandler & Shebo, 1982). Third, various examples of the embodiment models may be considered as perceptual anchoring, for example, finger use in various calculations (Crollen & Noël, 2015; Moeller et al., 2012; Wasner et al., 2016).

A second possible main source of number understanding can be the relations of the linguistic or conceptual systems. One such example is the relation within the DSS, such as the relation of small-large or even-odd nodes with the number nodes. This possible connection is similar to conceptual or linguistic models where meaning is partly represented in the connection of related concepts, such as the connection of “dog”, “barking”, and “bone”. Another possible connecting role of the language is that it may connect encapsulated domain-specific modules, making it possible to connect the information of the different modules that cannot be connected otherwise; for example, instead of handling “blue” and “left” independently, verbal children may understand what “left of the blue wall” could mean (Spelke & Tsivkin, 2001).

A third potential main source of number understanding is related to more dynamic semantic understanding where sources rely on analogy or metaphors of other domains. Dynamic is meant to be the opposite of static semantic knowledge. For example, in number addition, retrieving the sum of 3 and 2 can be static semantic knowledge that can be stored as a simple fact. Contrarily, the rule that the order of the addends does not influence the sum of the values and the justification of the rule are more dynamic knowledge that relies on the general properties of addition. One such example of dynamic understanding is the enriched parallel individuation model, which originally relied on the visual object tracking model, but later it relied on the object tracking model only metaphorically, assuming that key features of number handling and simple calculations are understood by the tracking of objects of a set (Carey, 2009; Carey & Barner, 2019). A similar model is built upon the phenomenon that young children can handle zero in simple numerical tasks, but they are unsure whether zero is a number (Krajcsi et al., 2021). It was argued that initially, integers are understood as properties of objects, and while the lack of objects (zero) can be handled, the value as a property cannot be considered when the objects the property refers to are missing (Krajcsi et al., 2021). This object-based conceptual understanding can be another source of number understanding. In another study, in contrast to the widely held belief, it was found that sign-value numbers (such as the Roman numbers) are easier to handle than place-value numbers (such as Arabic numbers) in simple multi-power numerical tasks (Krajcsi & Szabó, 2012). It was argued that multiple powers (such as ones, tens, and hundreds in a decimal system) are initially represented as a set of items, and sets of item groups, where a sign-value notation can be closer to this hypothetical representation than a place-value notation (Krajcsi & Szabó, 2012). Again, this item and group-based representation may serve as a dynamic conceptual base of semantic understanding. Finally, several components of embodiment models can be considered not only as perceptual anchoring but also as an analogy or metaphor of a mathematical concept, for example, walking and path as a metaphor for integers and related calculations (Moeller et al., 2012).

This framework is admittedly speculative and hypothetical. Its main purpose is to provide a systematic framework relying on previous research in numerical cognition and models of language and concepts. This framework can also be used to motivate and inspire future studies, and it is the question of future works whether this general outline is appropriate or whether more fundamental modifications are needed.

Acknowledgments

I thank Kimberly Brosche, and Gábor Lengyel for their comments on an earlier version of the present chapter.

References

Algom, D. (2021). The Weber–Fechner law: A misnomer that persists but that should go away. *Psychological Review, 128*(4), 757–765. https://doi.org/10.1037/rev0000278

Brannon, E. M., & Merritt, D. J. (2011). Chapter 14—Evolutionary Foundations of the Approximate Number System. In S. Dehaene & E. M. Brannon (Eds.), *Space, Time and Number in the Brain* (pp. 207–224). Academic Press. https://doi.org/10.1016/B978-0-12-385948-8.00014-1

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, *18*(6), 425–428. https://doi.org/10.1016/j.cub.2008.02.052

Carey, S. (2009). *The origin of concepts* (1st ed.). Oxford University Press, USA.

Carey, S., & Barner, D. (2019). Ontogenetic origins of human integer representations. *Trends in Cognitive Sciences*, *23*(10), 823–835. https://doi.org/10.1016/j.tics.2019.07.004

Crollen, V., & Noël, M.-P. (2015). The role of fingers in the development of counting and arithmetic skills. *Acta Psychologica*, *156*, 37–44. https://doi.org/10.1016/j.actpsy.2015.01.007

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. Oxford University Press.

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P. Haggard, Y. Rossetti, & M. Kawato (Eds.), *Sensorimotor foundations of higher cognition: Vol. XXII* (pp. 527–574). Harvard University Press.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neural model. *Journal of Cognitive Neuroscience*, *5*(4), 390–407.

Dietrich, J. F., Huber, S., & Nuerk, H.-C. (2015). Methodological aspects to be considered when measuring the approximate number system (ANS) – a research review. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00295

Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*(4), 389–395. https://doi.org/10.3758/BF03202431

Hines, T. M. (1990). An odd effect: Lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*(1), 40–46. https://doi.org/10.3758/BF03202644

Hohol, M., Willmes, K., Nęcka, E., Brożek, B., Nuerk, H.-C., & Cipora, K. (2020). Professional mathematicians do not differ from others in the symbolic numerical distance and size effects. *Scientific Reports*, *10*(1), Article 1. https://doi.org/10.1038/s41598-020-68202-z

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*(4), 498–525.

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(8), 927–934. https://doi.org/10.3758/s13423-016-1175-6

Krajcsi, A., Chesney, D., Cipora, K., Coolen, I., Gilmore, C., Inglis, M., Libertus, M., Nuerk, H.-C., Simms, V., & Reynvoet, B. (2024). Measuring the acuity of the approximate number system in young children. *Developmental Review, 72*, 101131. https://doi.org/10.1016/j.dr.2024.101131

Krajcsi, A., & Kojouharova, P. (2017). Symbolic Numerical Distance Effect Does Not Reflect the Difference between Numbers. *Frontiers in Psychology, 8*. https://doi.org/10.3389/fpsyg.2017.02013

Krajcsi, A., Kojouharova, P., & Lengyel, G. (2021). Development of Preschoolers' Understanding of Zero. *Frontiers in Psychology*, *12*, 3169. https://doi.org/10.3389/fpsyg.2021.583734

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analogue number system: Different symbolic and nonsymbolic numerical distance and size effects. *Frontiers in Psychology, 9*. https://doi.org/10.3389/fpsyg.2018.00124

Krajcsi, A., & Szabó, E. (2012). The role of number notation: Sign-value notation number processing is easier than place-value. *Frontiers in Psychology*, *3*(463). https://doi.org/10.3389/fpsyg.2012.00463

Krajcsi, A., Szabó, E., & Mórocz, I. Á. (2013). Subitizing Is Sensitive to the Arrangement of Objects. *Experimental Psychology*, *60*(4), 227–234. https://doi.org/10.1027/1618-3169/a000191

Leth-Steensen, C., Lucas, J., & Petrusic, W. M. (2011). Modelling SNARC by using polarity codes to adjust drift rates. *Proceedings of the 27th Annual Meeting of the International Society for Psychophysics, Herzliya, Israel*, *27*, 357–362.

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292–1300. https://doi.org/10.1111/j.1467-7687.2011.01080.x

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22. https://doi.org/10.1037/0096-3445.111.1.1

Moeller, K., Fischer, U., Link, T., Wasner, M., Huber, S., Cress, U., & Nuerk, H.-C. (2012). Learning and development of embodied numerosity. *Cognitive Processing*, *13*(1), 271–274. https://doi.org/10.1007/s10339-012-0457-9

Moyer, R. S., & Landauer, T. K. (1967). Time required for Judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Odic, D., & Starr, A. (2018). An Introduction to the Approximate Number System. *Child Development Perspectives*, *12*(4), 223–229. https://doi.org/10.1111/cdep.12288

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019. https://doi.org/10.1177/0956797613482944

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: An investigation of underlying mechanism. *Cognition*, *133*(1), 188–200. https://doi.org/10.1016/j.cognition.2014.06.011

Parrish, A. E., & Beran, M. J. (2022). Approximate Number System (ANS). In J. Vonk & T. K. Shackelford (Eds.), *Encyclopedia of Animal Cognition and Behavior* (pp. 381–386). Springer International Publishing. https://doi.org/10.1007/978-3-319-55065-7_1068

Pinhas, M., & Tzelgov, J. (2012). Expanding on the Mental Number Line: Zero Is Perceived as the “Smallest.” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1187–1205. https://doi.org/10.1037/a0027390

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*(3), 416–442. https://doi.org/10.1037/0033-2909.132.3.416

Rinaldi, L., & Marelli, M. (2020). The use of number words in natural language obeys Weber's law. *Journal of Experimental Psychology: General*, *149*(7), 1215–1230. https://doi.org/10.1037/xge0000715

Sella, F., Slusser, E., Odic, D., & Krajcsi, A. (2021). The emergence of children's natural number concepts: Current theoretical challenges. *Child Development Perspectives*, *15*(4), 265–273. https://doi.org/10.1111/cdep.12428

Spelke, E. S., & Tsivkin, S. (2001). Initial knowledge and conceptual change: Space and number. In *Language acquisition and conceptual development* (pp. 70–100). Cambridge University Press. https://doi.org/10.1017/CBO9780511620669.005

Sternberg, R. J., & Ben-Zeev, T. (1996). *The Nature of Mathematical Thinking*. Psychology Press.

Stoianov, I., & Zorzi, M. (2012). Emergence of a "visual number sense" in hierarchical generative models. *Nature Neuroscience*, *15*(2), 194–196. https://doi.org/10.1038/nn.2996

Szkudlarek, E., & Brannon, E. M. (2017). Does the Approximate Number System Serve as a Foundation for Symbolic Mathematics? *Language Learning and Development*, *13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Testolin, A., Zou, W. Y., & McClelland, J. L. (2020). Numerosity discrimination in deep neural networks: Initial competence, developmental refinement and experience statistics. *Developmental Science*, *23*(5), e12940. https://doi.org/10.1111/desc.12940

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*(1), 66–80. https://doi.org/10.3758/BF03196349

Verguts, T., & Van Opstal, F. (2014). A Delta-Rule Model of Numerical and Non-Numerical Order Processing. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1092–1012. https://doi.org/10.1037/a0035114

Wasner, M., Nuerk, H.-C., Martignon, L., Roesch, S., & Moeller, K. (2016). Finger gnosis predicts a unique but small part of variance in initial arithmetic performance. *Journal of Experimental Child Psychology*, *146*, 1–16. https://doi.org/10.1016/j.jecp.2016.01.006