MASTERARBEIT / MASTER’S THESIS

Titel der Masterarbeit / Title of the Master‘s Thesis

„Is the discrete semantic system (DSS) part of the more general conceptual or linguistic network?“

verfasst von / submitted by

Kimberly Brosche BSc BA MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 013
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Joint Degree Programme MEi:CogSci Cognitive Science
Betreut von / Supervisor:	Assoz. Prof. Dr. Attila Krajcsi

Abstract

The ubiquity of numbers in our culture confers especial importance on the cognitive mechanisms enabling numerical processing and comparison. When asked to identify the smaller/larger one of two Arabic numbers in comparison tasks, participants' performance improves with increasing numerical distance between the stimuli. This so-called symbolic distance effect can be attributed to the Discrete Semantic System (DSS), a linguistic and/or conceptual network of associations. Similar distance effects also occur when elements of non-numeric ordered lists such as letters and months are compared. This poses the question whether these distance effects might rely on a more general conceptual or linguistic network of associations or, in other words, whether the DSS is number-specific or domain-general. If the numeric and non-numeric distance effects rely on the same mechanism, their slopes would be expected to correlate. Hence, to address our research question, we subjected adult participants (n = 65) to three different comparison tasks, using Arabic numbers, letters of the alphabet and months of the year as stimuli. Subsequently, we calculated the reliability-adjusted correlations of the distance effect slopes for participants' error rate and reaction time across the three domains. In line with previous studies, significant distance effects were observed in all three domains for both error rate and reaction time. The distance effect slopes correlated moderately to strongly between numbers and months (error rate and reaction time) and between letters and months (error rate and reaction time). Further, we found weak correlations between the error rate distance effect slopes for numbers and letters as well as for the reaction time distance effect slopes for numbers and months. These results partly support the idea that the DSS drives symbolic comparisons beyond the number domain, at least in ordered lists as closely associated with numbers as months. However, participants likely applied different combinations of strategies to solve the comparison task, such as list retrieval and clustering, across the three domains, which could explain the imperfect correlations. That is, number and letter comparisons might rely on two distinct mechanisms, with month comparison partly drawing on both of them. Future studies should therefore aim to systematically manipulate and disentangle the properties of different ordered lists to elucidate how the human mind processes numeric and non-numeric symbolic comparison.

Keywords

Discrete semantic system, numerical cognition, comparison task, distance effect, semantic network

Zusammenfassung

Die Allgegenwart von Zahlen in unserer Kultur verleiht den kognitiven Mechanismen, die das Verarbeiten und Vergleichen von Zahlen ermöglichen, besondere Wichtigkeit. Wenn Versuchspersonen dazu aufgefordert werden, die kleinere/größere zweier arabischer Ziffern zu identifizieren, nimmt ihre Leistung mit zunehmender numerischer Distanz zwischen den beiden Stimuli zu. Dieser sogenannte symbolische Distanzeffekt kann dem „Discrete Semantic System" (DSS, zu Deutsch: Diskreten Semantischen System) zugeschrieben werden, einem linguistischen und/oder konzeptuellen Netzwerk von Assoziationen. Tatsächlich tritt ein ähnlicher Effekt auch auf, wenn Elemente nicht-numerischer geordneter Listen, wie Buchstaben oder Monate, verglichen werden. Dies wirft die Frage auf, ob diese Distanzeffekte auf einem generelleren konzeptuellen oder linguistischen Netzwerk von Assoziationen basieren könnten; in anderen Worten, ob das DSS ziffernspezifisch oder domänenübergreifend ist. Falls numerische und nicht-numerische Distanzeffekte auf demselben Mechanismus basieren, sollten ihre Steigungen (slopes) korrelieren. Um unsere Forschungsfrage zu beantworten, führten wir daher mit erwachsenen Versuchspersonen (n = 65) drei verschiedene Vergleichstests durch, in denen wir arabische Ziffern, Buchstaben des Alphabets und Monate des Jahres als Stimuli einsetzten. Anschließend berechneten wir die reliabilitäts-korrigierten Korrelationen der Steigungen der Distanzeffekte für die Fehlerrate und die Reaktionszeit der Versuchspersonen über die drei Domänen hinweg. Wie in vorherigen Studien konnte ein signifikanter Distanzeffekt in allen drei Domänen beobachtet werden, sowohl für die Fehlerrate als auch für die Reaktionszeit. Die Distanzeffektsteigungen korrelierten moderat bis stark zwischen Ziffern und Monaten (Fehlerrate und Reaktionszeit) und zwischen Buchstaben und Monaten (Fehlerrate und Reaktionszeit). Weiters fanden wir schwache Korrelationen zwischen den Fehlerraten-Distanzeffektsteigungen der Ziffern und Buchstaben als auch für die Reaktionszeit-Distanzeffektsteigungen der Ziffern und Monate. Diese Ergebnisse unterstützen teilweise die Idee, dass das DSS symbolischen Vergleichen jenseits der Zifferndomäne zugrunde liegt. Jedoch ist es wahrscheinlich, dass die Versuchspersonen je nach Domäne unterschiedliche Kombinationen an Strategien anwandten, wie etwa das Zurückgreifen auf Listen oder das Clustern, was die imperfekten Korrelationen erklären könnte. Das heißt, Ziffern- und Buchstabenvergleiche könnten auf zwei distinkten Mechanismen basieren, während Monatsvergleiche von beiden teilweise Gebrauch machen. Zukünftige Studien sollten daher anstreben, die Eigenschaften verschiedener geordneter Listen systematisch zu manipulieren und aufzuschlüsseln, um zu klären, wie das menschliche Gehirn numerische und nicht-numerische symbolische Vergleiche verarbeitet.

Acknowledgements

First, I wish to express my gratitude to my supervisor Attila Krajcsi for making my journey into a realm relatively unfamiliar to me, namely human psychology, such a pleasant, instructive and fun experience. I could always count on his motivated and motivating comments in which he tirelessly answered my innumerous and numerous questions that arose along the way, thereby guiding me through the twists and turns of this project. Further, I want to thank Tamás Szücs for patiently helping me to understand the details of his Javascript, which I was kindly allowed to base my script on. I am also deeply thankful for Miriam Simma's input that was an invaluable help in making the ideas in this thesis more comprehensible for an interdisciplinary audience. Importantly, this study would have been impossible without all the students who volunteered to compare symbols on a screen for an hour, which I highly appreciate. I especially want to highlight the role of my first pilot participant, Caroline-Chiara Zmeskal, for acting as a Guinea pig, for philosophizing about the last issues in the script that had to be ironed out and for ever-entertaining insights into the Hungarian language. Special praise also goes out to Elisabeth Zimmermann and Martyna Meyer, the Mei:CogSci coordinators in Vienna. They did not only lead me through the bureaucratic jungle surrounding the master's thesis procedure but also through the entire two-year adventure of this program. Similarly, my fellow CogSci students deserve a thank you for their thought-provoking and horizon-broadening input, and for creating an intellectual environment in which we could all thrive. Last but not least, I am deeply thankful to my human and non-human family and friends for their support, patience and encouragement throughout my studies.

Table of Contents

1. Introduction

Numbers pervade almost every aspect of our every-day lives. Both in non-symbolic and symbolic formats, numbers take on countless roles and functions. The nowadays indispensable symbols (e.g., “3”) and number words (e.g., “three”) humans have invented to facilitate trading, storekeeping and, of course, science, are seemingly unique in that they denote a numerosity rather than another semantic concept or referent. This poses the intriguing question whether numbers are also exceptional from a cognitive and evolutionary perspective. That is, whether the merits of numbers are unsurpassed by other ordered lists and to what extent the way we mentally represent numbers is unparalleled.

These questions are an integral part of the interdisciplinary endeavor of numerical cognition research. Its subject of investigation, numerical cognition, encompasses the numerical skills humans and non-human animals exhibit as well as more sophisticated mathematical abilities arising from these basic capacities (Kadosh & Dowker, 2015).

The centrality of numbers in human culture in conjunction with the question revolving around how exactly humans cognize and process numbers, which has been approached from a variety of perspectives, provide fertile ground for the interdisciplinary investigation presented in the current master’s thesis.

1.1 The relevance of numbers and numerical cognition

As for many cognitive phenomena, the first discipline to be captivated by numerical cognition was philosophy. By providing thought-provoking impulses, numerical cognition and the research thereof can influence major philosophical theories in cognitive science. For instance, a central theme in philosophy of cognition, embodiment, is colored by numerical cognition. The embodiment debate revolves around the questions to what extent and how our body shapes our cognitive experiences and processes (Shapiro, 2007). Despite numbers being regarded as the (or one of the) most abstract concept(s), some have argued that even numerical cognition is embodied (Fischer, 2012). For example, the SNARC (space-numerical association of response codes (Dehaene, 1992)) effect strongly suggests an inherent link between space and numbers. Both humans and non-human animals have been observed to recognize and respond to small numbers or numerosities faster and/or more accurately if these are presented on the left and if one needs to respond on the left, whereas the reverse is true for larger numbers or numerosities (Rugani & Hevia, 2016; Rugani & Rosa-Salva, 2021). These and other insights into entanglements between cognition and body (or environment) make numerical cognition highly relevant and insightful to philosophy (of cognition).

Apart from deeply theoretical and abstract matters, numerical cognition research also carries profound practical implications for education and therapy. Many seemingly abstract skills measured in numerical cognition tasks have been shown to correlate with arithmetic performance in children and adults. For example, individuals' accuracy in the perception of numerousness correlates with their math achievement (Halberda et al., 2008) and performance in both symbolic (Toll et al., 2015) and non-symbolic (Desoete et al., 2012) number comparison tasks was found to be an important predictor of kindergarteners' mathematical skills and disabilities. Additionally, characteristics such as the magnitude of the SNARC effect could even help to understand and gage cognitive aging (Uittenhove & Lemaire, 2015). In practice, these and similar findings can be utilized to better characterize and take into account individual differences in math performance and even pathologies such as dyscalculia.

Further, numbers have even elicited the interest of linguists. Researchers have inquired, for instance, whether and how strictly the cognitive processing of numerosities depends on language. Even though performance in numerical cognition tasks seems to be influenced by large disparities in "number-sensitivity" of the human participants' (native) language, some core capacities appear to be unaffected by such discrepancies. For example, several studies have probed the numerical competence of indigenous tribes whose languages only differentiate "one", "two" and "many" but do not contain a word for exact numerosities beyond two. Despite these participants being restricted in their manipulation of exact (larger) numerosities, non-linguistic abilities seem to remain largely unaffected by language (Gordon, 2004; Pica et al., 2004). In addition, subtle differences in numerical task performance between speakers of languages that either name the tens or the ones first have been shown to mainly be an artefact of input and output processes (Brysbaert et al., 1998). Taken together, numerical cognition does not seem to depend as strongly on language as previously thought.

The dispensability of language in at least some numerical cognition abilities is further backed up by the ample evidence suggesting quite advanced numerical cognition in non-verbal (or pre-verbal) beings. Both young human infants and non-human animals have been tested for their "sense of number", providing valuable insights into the phylogeny (evolution) and ontogeny (individual development) of numerical cognition. Findings yielded by these lines of research are progressively casting doubt on the notion that a sense for numbers is restricted to adult humans endowed with a sophisticated linguistic numbering/enumeration system.

1.2 Numerical cognition in humans and non-human animals

Being sensitive to quantities is crucial to animals' survival in the wild. For example, it is advantageous for animals to be capable of discriminating between two quantities of food and to compare different food patches (Nieder, 2020). Others have argued that it is even more important to keep track of individual animals, such as predators and social companions, which often includes grasping their (approximate) numerosity (Wynn, 1998; Wynn & Bloom, 1992). Thereby, basic numerical cognition confers an advantage on animals not only when trying to escape predators but also during decision-making in gregarious species (see also Piantadosi & Cantlon, 2017). Consequently, numerical cognition constitutes a fitness benefit and should therefore be favored by natural selection, as more apt individuals will be more likely to escape from predators or find sufficient food.

Hence, it comes with little surprise that at least basic forms of numerical cognition are widespread in the animal kingdom. This fact has not escaped the attention of researchers who have, thus, intensively explored numerical cognition in non-human animals.

Many of these research endeavors have primarily focused on non-human primates and particularly great apes. For example, monkeys can correctly order arrays of ten to 30 elements (e.g., dots) in ascending order after having been trained to do so with one to nine elements. When performing this task, they were equally accurate as adult human participants and, in addition, the monkeys were faster than their human counterparts (Cantlon & Brannon, 2006). In a follow-up study, the same monkeys and a human control group also proved capable of performing mental addition of these arrays. Participants of both species were sequentially presented with two arrays (e.g., three elements and then five) on a screen. Subsequently, they could choose between one array representing the correct numerical sum of the two previous arrays (e.g., eight in this example) and a distractor, i.e., an incorrect sum. Despite having to transfer the learned rule from the test set to the novel training set (partly containing combinations subjects had never been trained with) and even though cumulative surface area was controlled for, the monkeys performed similarly to the human participants (Cantlon & Brannon, 2007). An investigation of primates' performance in a simultaneous (i.e., the stimuli are presented simultaneously rather than sequentially) summation task was performed by Rumbaugh et al. (1987). Chimpanzees were trained and tested with two trays containing two wells each. These wells could be filled with chocolate treats, a highly palatable food reward for chimpanzees. When chocolates were present in both wells of a tray (i.e., four wells in total), chimpanzees had to "add up" the number of chocolates in the two wells of each tray to be able to choose the overall larger number of chocolate treats – which they did at high success rates. However, the authors acknowledge that it remains unclear which cognitive mechanism might have permitted chimpanzees to succeed in the summation task. For example, the chimpanzees could have employed between-set counting strategies. Alternatively, they could have even made use of the overall volume, rather than the numerosity, of the chocolate treats, given that it is not only methodologically difficult but also ecologically invalid to control for overall volume or surface area when presenting edible rewards as stimuli.

To avoid these and other pitfalls, a handful of studies have taken a radically different approach by teaching animal subjects number symbols. In Boysen and Berntson (1989), one female chimpanzee was trained to associate arrays of edible objects with placards depicting Arabic numbers from 1 to 3. She successfully transferred what she had learned to a wider range of numbers (0–15) and objects (e.g., inedible kitchenware). In a next step, she was required to select the correct sum of oranges hidden at two different locations in the enclosure. The same was afterwards carried out with the number placards, rather than oranges, being hidden. In both cases, the subject performed well above chance level in choosing the placard depicting the correct sum. Boysen and Berntson conclude that the chimpanzee had not only learned the correspondence between the novel symbols and the numerosity of object arrays but must also be capable of at least counting between sets, if not mental addition. Like many similar studies, however, Boysen and Berntson's only included a very limited sample size (in this case only one subject) and tested very unique individuals under exceptional conditions. Apes cross-fostered by humans and trained to communicate with humans are highly influenced by human culture. Despite the merits of these studies in situating the hard limits of a species' cognitive capacities under optimal (training) conditions, it is also worthwhile to disentangle the potential impacts of culture acting on both humans and human-reared apes.

A highly suitable model species to probe "core" elements of numerical cognition in non-human animals are young domestic chicks. First of all, being birds, chickens are only distantly related to humans and non-human primates. This, in turn, makes the discovery of human-like cognitive abilities in chickens particularly interesting as it suggests that these abilities are either as evolutionary ancient as the last common ancestor of primates (mammals) and birds, or that they have evolved (at least) twice independently in reaction to similar environmental challenges. Second, domestic chicks are precocial. This means that, soon after hatching, they are already mature enough to move around, forage etc. without requiring assistance or intensive care from their parents. Consequently, they can be subjected to behavioral experiments soon after hatching, which allows researchers to investigate (largely) experience-independent "core capacities" (Chiandetti & Vallortigara, 2018). In addition, chicks' propensity to spontaneously rejoin (the larger group of) artificial imprinting objects (e.g., ping-pong balls) they have been reared with can conveniently be exploited in (numerical) cognition tests even without extensive training (Chiandetti & Vallortigara, 2018; Marino, 2017; Regolin et al., 2005; Regolin & Vallortigara, 1995; Vallortigara et al., 1998). For these reasons, many endeavors to elucidate numerical cognition in non-human animals have focused on domestic chicks.

One such study has revealed that chicks are capable of discriminating between two small sets of objects based on numerosity (Rugani et al., 2008). Chicks were trained to peck an array of dots of specific number (e.g., two) in exchange for a food reward. When presented with two different arrays, chicks successfully chose the learned numerosity if the number of elements did not exceed three in either array. This finding was robust even when overall surface area and perimeter were controlled for (but for a critical discussion of such controls in general see: DeWind et al., 2015). However, one might wonder whether chicks can also perform such numerosity-based discrimination without undergoing prior training.

The question whether chicks spontaneously "imprint on a numerosity" and later identify the imprinting stimulus based on numerosity alone was addressed in Rugani et al. (2010a). In this study, chicks were reared with one to four artificial imprinting objects varying in shape and color. This phase was succeeded by a test in which chicks were given the choice between two sets of objects differing in their numerosity and/or overall volume. Even though chicks chose the larger number or volume of familiar objects whenever they were reared with sets of *identical* imprinting objects, they also seemed capable of recognizing the most familiar set of objects based on numerosity alone. That is, in one of the experiments, chicks were reared with a miscellaneous collection of objects *differing* in shape and color. In the test, the only feature one of the two sets presented to them shared with the imprinting set was the numerosity (but not the color or shape). In this condition, chicks spontaneously opted for the numerosity they had been reared with, irrespective of overall volume of the sets.

Another property of numbers that chicks appear to be able to utilize is ordinality. Roughly, ordinality refers to the stable ordering of numbers, e.g., that 4 comes after 3 (Gallistel, 1993). To assess this, chicks were trained to find a hidden food reward in the n^{th} (e.g., 4^{th}) cup in a sagitally presented sequence of cups (Rugani et al., 2010b). In the test, the only cue chicks could rely on to find the food was the ordinal position of the cup as a) cups were now presented in a fronto-parallel line (i.e., the original arrangement was rotated by 90° so that the cups were now aligned from "left to right" and orthogonally to the direction into which chicks were released) and b) the inter-element space was varied in some conditions, making it impossible to find the food relying on the previous “absolute” position of the baited cup. Astonishingly, chicks succeeded in locating the baited cup even under these circumstances in the test, suggesting an ability to utilize ordinal information. Such a conclusion would be in line with the results of a study investigating transitive inference in adult domestic hens (Hogue et al., 1996). In Hogue et al., hens witnessed a familiar dominant individual defeat or being defeated by an unfamiliar hen during hierarchy formation. Remarkably, hens were less likely to afterwards attack the strangers that had defeated the familiar dominant hen. Although some alternative mechanisms could not convincingly be ruled out, the hens’ adjustment of behavior hints at their use of transitive inference during hierarchy formation.

Finally, chicks have also been shown to perform simple proto-arithmetic operations on small numerosities (Rugani et al., 2009). In Rugani et al.’s experiment, the chick was facing two screens that the chick knew to contain some of the artificial imprinting objects it had been reared with. The experimenter then visibly moved some of the objects (hanging on a thread) from one screen to the other screen. Based on their innate tendency, chicks were expected to rejoin the group that was larger after the transitions if they are able to perform these simple, approximate additions and subtractions. And indeed, chicks proved capable of handling small numerosities up to three even when other cues, such as the direction of the last transfer, were controlled for.

In summary, rather than being unique to humans, many basic cognitive abilities allowing the representation, processing and manipulation of numerosity are widespread and, presumably, phylogenetically ancient. The next chapter is dedicated to a task that assesses one of the most fundamental and evolutionarily adaptive (Nieder, 2020) components of numerical cognition: quantity comparison.

1.3 Symbolic and non-symbolic comparison tasks

As the name suggests, comparison tasks require (human and non-human) subjects to compare two simultaneously or sequentially presented stimuli. Typically, these are presented on a screen and subjects are asked to select the smaller or larger one of the stimuli/sets via the screen or via response buttons. From these responses, subjects’ error rate and reaction time can be calculated for each stimulus pair.

Apart from many of the animal studies mentioned above (e.g., Rugani et al., 2008), numerous studies with human participants have employed comparison tasks. In these tasks, the type of stimuli to be compared varied widely. Not only did the stimuli include discrete entities like dots on a screen (Kolkman et al., 2013; Lourenco et al., 2012; Toll et al., 2015) or sequentially played tones (Barth et al., 2005) but also continuous, physical features like the length of lines (Tudusciuc & Nieder, 2010).

Crucially, as opposed to (most) non-human animals (but see Terrace et al., 2003) and human infants, adult humans can not only be subjected to such *non-symbolic* comparison tasks (in which approximate numerousnesses of objects are to be compared (Dos Santos, 2022)) but can also be asked to select one of two *number symbols* representing exact numerosities (for a definition of symbolic and non-symbolic comparisons see: De Smedt et al., 2013).

Symbolic comparison tasks work analogously to non-symbolic comparison tasks, except for the way the quantities are presented. Instead of arrays of objects, participants are asked to select the smaller or larger one of two numbers (i.e., number symbols such as “3”). While this precisely taps into numerosity comparison, excluding any sort of continuous variables or other confounds present in non-symbolic comparison, such symbolic comparison tasks require the mastery of number symbols and are thus limited to (educated) humans.

Interestingly, the relationship between subjects’ performance in symbolic and non-symbolic comparison tasks still remains poorly understood. That is, while some have found an intra-individual correlation between the performances in these two tasks (Gilmore, 2015; Toll et al., 2015), others have not (Holloway & Ansari, 2015; Maloney et al., 2010). Despite this, two effects consistently occur when participants’ error rate and reaction time are analyzed in both symbolic and non-symbolic comparison tasks: the distance effect and the size effect.

1.4 The distance (and size) effect

The term distance effect refers to the observation that participants’ error rate and reaction time decrease as the numerical distance between the two stimuli to be compared grows (Dehaene et al., 1998; Moyer & Landauer, 1967). For example, determining that nine (the number symbol “9” or an array consisting of nine dots) is larger than two (the number symbol “2” or an array consisting of two dots) is easier than making the same judgment about eight and nine. These results are commonly interpreted as comparing “far” stimulus pairs being easier than comparing “close” stimulus pairs.

Analogously, the size effect predicts that the task’s difficulty increases with the magnitude of the numbers to be compared (Dehaene et al., 1998). That is, participants are more accurate and faster in trials involving smaller numbers (e.g., two and four) than when two larger numbers (e.g., seven and nine) are presented. However, the present thesis focuses primarily on the distance effect.

Distance effects have been reported for both symbolic and non-symbolic comparison tasks. Early evidence comes from Moyer and Landauer (1967) who found a distance effect for adult humans' reaction time and error rate when number symbols were compared. Similar findings have been obtained in the absence of symbolic or linguistic stimuli. For example, the abovementioned study in which chimpanzees learned to add up chocolate treats in the wells of two trays (Rumbaugh et al., 1987) as well as Cantlon and Brannon's (2006) comparative study on monkeys and humans both revealed a distance effect. In conjunction with results from human cultures largely lacking number words (Gordon, 2004; Pica et al., 2004), these and similar findings have prompted many to attribute the distance effect to a mechanism common to symbolic and non-symbolic magnitude processing in humans and animals (Gallistel & Gelman, 2000; Walsh, 2003).

1.5 The Analogue Number System

The most widely advocated among the proposed mechanisms is the so-called analogue number system (also referred to as approximate number system; ANS). The analogue number system is believed to support magnitude processing in animals and humans and to ultimately give rise to humans' sophisticated mathematical (e.g., arithmetical) abilities (Cantlon et al., 2009; Feigenson et al., 2004; Gallistel & Gelman, 2000). The representations underlying the processing of magnitudes via the analogue number system are inexact and, therefore, noisy.

This noisiness inherent to analogue magnitude processing is believed to be the root of the distance and size effects. According to the analogue number system, the noisiness of magnitude representations and, consequently, the difficulty of comparisons tasks, adhere to the Weber-Fechner law (Fechner, 1860). The Weber-Fechner law posits that the ratio between the two stimuli (magnitudes or psychophysical properties such as the weight of objects) determines the discriminability of the two stimuli. Analogously, the noisiness increases with growing magnitudes of the discriminations. Hence, both the distance and the size effect can be attributed to the ratio-based functioning of the ANS.

At the same time, the presence of these effects has also been taken as evidence for the crucial role of the ANS in both symbolic and non-symbolic number processing (e.g., Cantlon et al., 2009). That is, given that the distance effect occurs in both non-symbolic (in humans and animals) and symbolic number comparison tasks, it might seem plausible to assume that symbolic comparisons too are handled by the ANS, in virtue of the magnitudes they represent. However, this conclusion has been challenged by more recent evidence supporting an alternative explanation for the symbolic distance effect: the discrete semantic system.

1.6 The Discrete Semantic System

In contrast to the ANS, the discrete semantic system (DSS) attributes the symbolic distance and size effects to a network of semantic associations between nodes representing the digits (Krajcsi et al., 2016).

Such a semantic network has already been proposed for other types of mental representations. For example, the mental lexicon (Taft, 2015) is believed to be organized as a mental collection of words and their building blocks (Manova & Knell, 2021). Similarly, Ashcraft (1987) proposed a semantic network for arithmetic fact knowledge to account for differences in the difficulty of arithmetic problems. But how could a semantic network bring about the distance and size effects?

The DSS posits that the distance effect is the result of the associations between number nodes and concepts such as “small” or “large” (Krajcsi et al., 2016). Number symbols representing smaller magnitudes (e.g., “2”) should be more closely associated with the concept of “small” than those representing larger magnitudes (e.g., “9”). Therefore, the associations we have with “2” and “9” overlap less than those of “2” and “3”, leading to a higher degree of semantic similarity in the latter case. In consequence, successfully discriminating more distant number pairs in a symbolic comparison task is easier (i.e., participants’ error rate and reaction time will be lower) than discriminating close pairs.

While the distance effect can be accounted for by strength of association, the DSS ascribes the size effect to statistical frequencies of occurrence (Krajcsi et al., 2016). Given that we encounter smaller numbers more frequently, they might be easier to activate and process (cf. Ashcraft, 1987). As a result, comparisons involving smaller numbers (e.g., “2” and “4”) are easier and faster to solve than those involving larger numbers (e.g., “7” and “9”).

A corollary of the predictions made by the DSS is that the processing of symbolic magnitudes fundamentally differs from that of non-symbolic ones. On the one hand, the DSS does not deny the role of the ANS in non-symbolic magnitude processing. For such purposes, humans might still recruit the same mechanisms as non-human animals. On the other hand, the DSS does assert that humans’ processing of symbolic numbers – and therefore also what mechanisms are at work in symbolic comparison tasks – is crucially different from the system(s) dealing with non-symbolic magnitudes. Even though such a claim might seem counter-intuitive at first, in light of striking similarities such as the presence of the distance and size effects in both symbolic and non-symbolic comparison tasks, the following pieces of empirical evidence firmly undergird the predictions made by the DSS and invalidate those of the ANS.

The ANS postulates the same cause to underlie the distance and size effects, namely the inexactness of the magnitude representations which is dependent on both the ratio between the stimuli and their absolute size (Moyer & Landauer, 1967). If this is the case and the two effects are in fact two sides of the same coin, their slopes (of participants' performance plotted against the distance or size) should be highly correlated. However, counter to this prediction, such a correlation was only found in a non-symbolic but not in a symbolic number comparison task (Krajcsi, 2017).

This is in accordance with the findings obtained by experimentally manipulating the distance and size effects. Krajcsi and colleagues (2016) were able to show that the two effects can be dissociated when the frequency of the symbols (previously unknown numbers acquired by the participants in the course of the experiment) is manipulated. According to the ANS, both effects should solely depend on the quantity represented by a number symbol, regardless of the symbol's frequency. In contrast, the distance effect was retained while the size effect vanished when the frequency distribution of the artificial numbers was kept uniform (unlike for our every-day use of numbers, in which smaller numbers occur more frequently than larger ones).

In a similar vein, the ANS prescribes that the distance effect should solely reflect the distance between the quantities represented by the number symbols and not the relative strength of associations. To test this, Krajcsi and Kojouharova (2017) isolated the distance and size effect in an experimental set-up. Participants learned new number symbols corresponding to the Arabic numbers 1-3 and 7-9, leaving a gap between 3 and 7. As predicted by the frequency-based distance effect described in the DSS, the distance between 3 and 7 was treated as one unit. Therefore, the dissociation of the size and the distance effect in experiments utilizing artificial number symbols casts additional doubt on the role of the ANS in symbolic number processing.

In light of several faulty predictions of the ANS, we will now turn to the DSS and its implications. An intriguing question that arises is whether the semantic network of the DSS, which presumably generates the distance and size effects, is limited to the number domain or whether it could be more universal, encompassing also other ordered lists of symbols. In particular, one could expect the distance effect to occur in all lists whose elements could differ in their strength of association with concepts such as “small”, ”early”, “large” or ”late”.

1.7 The distance effect beyond the number domain

Confirmatory evidence for the claim that the distance effect arises from semantic and/or linguistic associations comes from other ordered lists that are almost as omnipresent in our lives as the numbers of the counting list, namely letters and months.

In an early series of experiments, Fairbank (1969) attempted to replicate and extend the findings of Moyer and Landauer (1967). Participants had to compare two numbers, two letters (A to I) or two months of the year (January to December). Doing so, Fairbank did not only find a distance effect for the numbers but noteworthily also for the letters and the months. Even though the reaction times were generally shorter for the numbers, these experiments were, to our knowledge, the first to hint at a similarity in the way numbers and other ordered lists are processed.

This avenue of research was soon advanced and Fairbank's methodology was refined. Applying a similar comparison task, Parkman (1971) attempted to rule out that the similarity between numbers and letters can simply be attributed to participants' use of the ordinal positions of the letters A to I (which could be viewed as equivalent to the numbers 1 to 9). Therefore, he presented the letters A through J to one group of participants and the letters K through T to another group of participants. Lending support to the result of Fairbank (1969), Parkman found a distance effect for both lists of letters. Subsequent investigations have since then applied a variety of experimental paradigms to deepen our understanding of distance effects in numeric and non-numeric symbolic domains.

For example, Moyer (1973) presented participants with animal names (e.g., "ant" vs. "frog") and asked them to identify the one larger in size. The animals were ranked according to their "ordinal position" and participants agreed to this ranking prior to the experiment. As for the number comparison, the greater the distance (difference between ordinal positions) between two animals the shorter was the participants' reaction time. Similarly, Foltz et al. (1984) measured the reaction time of participants identifying the larger one of two numbers (e.g., "5"), number words (e.g., "five") and object names that could be classified according to size (e.g., "apple" vs. "whale"). Reaction times were generally longest for object names and shortest for numbers presented as digits. Despite the difference in intercepts, a distance effect was clearly detected for all three domains.

Another study could show that the distance effect seems to extend to newly learned relations between object names (Potts, 1972). Participants were given a paragraph to read before the experiment that described the relationship between four animals. In one case, the animals were described as differing linearly in intelligence, in the other case they were ranked according to friendliness. This was aimed at allowing participants to form a linear sequence in the form "A > B > C > D" in memory before the task commenced. In the task, they had to judge whether statements of the form "A is more intelligent than B." were true or false. Indeed, participants' error rate was lower and their reaction time was shorter when the sentence confronted a remote pair (e.g., "A" and "D") than when a close pair was presented (e.g., "A" and "B").

One of the more recent studies exploring the distance effect in non-numeric domains employed a slightly different methodology: numbers and months were to be compared to a given standard (Morsanyi et al., 2016). In the task, participants had to indicate whether the presented number (1 to 9) or month (January to September) was smaller or larger than the standard, i.e., "5" or "May", respectively. Participants showed a higher performance (a combined measure of reaction time and accuracy) with numbers than with months and individual comparison skills even correlated with arithmetic abilities. Importantly, a distance effect was clearly present in both domains. However, a noteworthy caveat of using months as a control for numbers is that months are frequently substituted by numbers in every-day life (e.g., "May" is often written as "5" or "05"). This property, in turn, could have enabled participants to solve the number comparison task by transcoding (i.e., mentally converting) the months to numbers and comparing the *numbers* corresponding to the months rather than the months themselves. The authors counter this criticism arguing that, contrary to their findings, the reaction times for months would be lengthened by such an additional processing step. Nevertheless, other ordered lists less closely associated with numbers might be a better choice to elucidate the functioning of the symbolic distance effect.

To explore potential similarities between the distance effects in numbers and letters further, Van Opstal et al. (2008) availed themselves of a methodology resembling that of Morsanyi et al. More precisely, they asked participants to judge whether the presented number (ranging from 1 to 9) or letter (from J to R) was smaller or larger than a standard. This standard was "5" in the case of the numbers and "N" in the case of the letters. Before the target stimulus (the one to be compared to the standard) appeared on the screen, a priming stimulus was presented, which participants were instructed to ignore. While the priming distance effect (i.e., the smaller the distance between the target and the prime the *shorter* the reaction time (Koechlin et al., 1999)) only reached statistical significance for numbers, the comparison distance effect was present for both letters and numbers. These findings were confirmed in a follow-up experiment utilizing a different range of letters (I to Q).

Whether the similarities between the distance effects in the number and letter domains would also extend to the neural level was investigated in Goffin et al. (2020). Again, as in Van Opstal et al. (2008), a target stimulus was to be compared to a standard. Goffin et al. used the numbers from 2-8 and the corresponding letters (B to H) with the standards being 5 and E, respectively. Participants' response was assessed both from a neural (fMRI scans) as well as from a behavioral (reaction time and error rate) perspective. Interestingly, the expected neural effect (rebound effect) was evident in the fMRI data only for numbers but not for letters. Nevertheless, a behavioral distance effect emerged for both domains. One important merit of Goffin et al.'s study is that they actively investigated whether participants transcoded letters (like months, potentially) to numbers, since such a strategy could cause a misleading similitude between number and letter distance effects. That is, the authors asked one group of participants to solve arithmetic tasks using letters as operands (e.g., A + B = C because A corresponds to 1, B corresponds to 2 and C corresponds to 3). If transcoding letters to numbers is indeed what allows participants to solve the letter comparison task, the group that had received training in treating letters as numbers should be more successful in the comparison task than a control group. This was not the case, as the two experimental groups did not differ in their performance in the letter comparison task. Therefore, the authors conclude that the letter distance effect is the result of judgments about the ordinal relationships of letters, independent of the numbers they correspond to.

Alongside comparison tasks, another important line of evidence for distance effects in non-numeric ordered lists comes from ordering tasks. In these tasks, a group of stimuli (e.g., two or three) are presented and participants are required to evaluate whether they are presented in ascending or descending (or no) order. In such tasks, a reverse distance effect usually occurs for numbers. Contrary to the standard (canonical) distance effect observable in comparison tasks, the reverse distance effect is characterized by sequences with shorter distances (e.g., "1, 2, 3" as opposed to "1, 3, 5") being easier recognizable (resulting in shorter reaction times and error rates). While there is some controversy surrounding the origins of the reverse distance effect, e.g., serial scanning of the sequence as opposed to association-based mechanisms (Vos et al., 2017), it is generally believed to largely rely on familiarity (Turconi et al., 2006), i.e., resemblance to the counting list. Accordingly, short-distance, ascending sequences elicit a higher degree of familiarity than long-distance descending or unordered ones (Devlin et al., 2022). Given the similarity of the counting list with the alphabet or the sequence of months within a year, it comes with little surprise that ordering experiments have also revealed such a reverse distance effect in other, non-numeric domains.

For instance, the abovementioned study by Morsanyi et al. (2016) included an ordering task using numbers and months as stimuli, in addition to the comparison task. Participants were presented with a triad of numbers (1 to 9) or months (January to September) and asked to indicate whether they were in ascending order or order in which they appear within one calendar year. The reverse distance effect was only present in the subset of correctly ordered trials and, astonishingly, it reached statistical significance only for months but not for numbers.

A more extensive comparative study was conducted by Vos et al. (2017). Triads of numbers (1 to 9), letters (A to I) and months (January to September) were presented to participants which then had to indicate whether the sequence was in order (ascending or descending) or not (unordered). They found the reaction times for months to be considerably longer than for the other domains and the error rate to be lowest for ascending sequences. Regarding the distance effect patterns, numbers, letters and months exhibited high degrees of similarity. That is, reverse distance effects were found for ascending and descending sequences for at least one of the response measures, reaction time or error rate, in all domains, while a standard distance effect emerged in unordered number and letter trials. Additionally, the finding that the distance effects were generally more pronounced and participants' performance was higher in ascending trials, led the authors to conclude that the (reverse) distance effect in all domains was driven by the varying strength of non-numeric associations. The authors theorized these associations to arise from differences in frequency of co-occurrence between the elements of the sequence, which are stronger for ascending sequences.

In contrast to Vos et al.'s findings, a previous study by Franklin et al. (2009) yielded slightly different results. Franklin and colleagues subjected participants to an ordering task in which ascending, descending or unordered sequences of three two-digit numbers (11 to 99) or months (January to December) were presented, some of which crossed a decade (e.g., 15 – 23 – 25) or year (e.g., December – January – February) boundary. Interestingly, for both numbers and months, a reverse distance effect was only detected when the sequence crossed a decade/year boundary, but the effect was absent when such a boundary was not crossed. That is, even though the distance effect patterns appear to be similar between numbers and months, the phenomenon still remains insufficiently understood.

In a similar vein, another body of evidence challenges the comparability of number and letter distance effects, at least in ordering tasks (Fulbright et al., 2003). When participants were asked to judge whether numbers, letters and shapes were in ascending order (according to the counting list, the alphabet or size, respectively), a *standard* distance effect in participants' reaction time only occurred for numbers and shapes but not for letters. In addition, the brain areas activated by the three tasks differed, as revealed by functional MR scans conducted during the task, with the letter task recruiting the largest variety of brain areas.

A very recent study (Anonymous, unpublished) combined an ordering and a comparison task to more comprehensively research the distance effect in numbers, letters and months. In the ordering tasks, participants had to judge whether a sequence of two-digit numbers, letters (B to L) or months (February to December) were in ascending order or not. In the comparison task, two stimuli out of the same stimulus pool were simultaneously presented and participants were instructed to select the larger one. A distance effect (reverse for the ascending ordering trials) was present for all domains in both tasks, even though the magnitude of the effect was highest for letters and smallest for numbers. In addition, the authors also found a correlation of the reaction times across the domains for both tasks. However, similar to a study by Attout et al. (2014) that correlated the magnitude of the number and letter distance effects, the correlation between the *slopes* of the distance effects was not assessed. This shortcoming is noteworthy, given that only the investigation of the correlations between the distance effect slopes can reveal more about the potentially shared mechanisms underpinning the comparison skills in different domains.

1.8 Research question

The multitude of studies outlined above paint a rather consistent picture: The distance effect does not seem to be unique to (symbolic and non-symbolic) representations of quantities. On the contrary, we have seen that the effect extends to letters of the alphabet, months of the year and even names of (ordered) objects such as animals. Most of the studies reviewed here revealed highly similar distance effect patterns for numeric and non-numeric domains. However, the correlations between the distance effect slopes were, to the best of our knowledge, never calculated. Consequently, these previous studies leave us in the dark about whether or not the distance effects can indeed be attributed to the same cognitive mechanism in all cases.

Given this knowledge gap, the present study aims to investigate whether the comparison distance effect relies on the same cognitive mechanism for number symbols and two non-numeric symbolic ordered lists, namely letters of the alphabet and months of the year. If they do rely on the same mechanism, namely a network of semantic and/or linguistic associations, this would imply that the discrete semantic system driving the symbolic numerical distance effect is domain-general. In analogy to the account of the symbolic numerical distance effect (Krajcsi et al., 2016), this could mean that letters at the beginning of the alphabet and months at the beginning of the year are more closely associated with concepts such as “small” or “early”, and letters and months at the end of their respective lists to be more closely associated with concepts such as “large” or “late”. The prediction deducible from a common mechanism underlying the distance effects is that the slopes of the distance effects, i.e., how steeply participants’ reaction time and error rate decrease with increasing distance between the stimuli, should be highly correlated across numbers, letters and months. Importantly, to avoid these correlations from becoming attenuated (i.e., to avoid a false negative result) due to imperfect intra-individual reliabilities of the distance effect slopes, we adjusted the observed correlations for the reliability of the distance effect slopes to yield the disattenuated (“true”) correlation (see Krajcsi, 2017). These disattenuated correlations can be expected to be near one and to differ significantly from zero if the DSS is domain-general and constitutes the basis of the symbolic comparison distance effects in both numeric and non-numeric domains.

2. Methods

To answer our research question, we conducted an online experiment in which participants were subjected to three different conditions of a comparison task: they were instructed to indicate the larger (or later) one of two numbers, letters and months of the year. Using participants’ reaction time and error rate, we subsequently calculated the individual distance effect slopes for all three domains to be able to assess the correlations.

2.1 Ethical approval

This study was approved by the ethics committee at Eötvös Loránd University, Hungary and is in accordance with Hungarian law. Participation in the study was voluntary. Participants gave their explicit consent to the anonymized processing of their data and they were free to revoke their consent and terminate the study at any time.

2.2 Participants

Participants were 65 undergraduate students (44 females, 20 males, one non-binary person; mean age = 21.69 years ; range = 18 to 33) enrolled in various majors at Eötvös Loránd University. They received partial course credits in exchange for their participation.

2.3 General procedure

Participants were tested in three conditions of a comparison task, one with numbers, one with letters and one with months of the year, via the online platform cognition.run. We conducted 720 trials per condition, amounting to an overall number of 2160 trials per participant and an overall duration of approximately 60 min per participant. All possible, non-identical combinations of stimuli within each domain were used and each pair was presented 20 times – 10 times with the larger/later stimulus being on the right and 10 with the larger/later stimulus being on the left. In each condition, the task was to identify which of the two presented stimuli comes later (in the number sequence, the alphabet or the year). The order of conditions was counter-balanced across participants, amounting to six different sequences to which participants were randomly assigned.

2.4 Stimuli

Simultaneously presented pairs of Arabic numbers, letters of the English alphabet and months of the year acted as stimuli.

For the number comparison, we included only single-digit numbers, i.e., the Arabic numbers from 1 to 9. To match the number range, we only presented nine letters of the English alphabet and nine months of the year (abbreviated to the first three upper-case letters to reduce reading time; as in Anonymous, unpublished). Because months are frequently depicted as numbers and, therefore, presumably strongly associated with them (as also noted by Anonymous, unpublished), it was important to also contrast numbers with a second type of ordered symbolic non-numeric list, namely letters of the alphabet. To additionally discourage participants from transcoding the letters or months to numbers (e.g., A corresponds to 1), we did not use the first nine elements (cf. criticism put forward by Parkman, 1971). Especially letters in the middle of the alphabet are less likely to be transcoded to numbers by participants than months or the first letters of the alphabet (cf. Goffin et al., 2020). Also, even though the study was conducted in Hungarian, we omitted Hungarian letters that are not part of the English alphabet (e.g., “á”, “ü” or “sz”) as well as months whose Hungarian abbreviations are longer than three letters (e.g., “SZEP” for “Szeptember”). The final list of stimuli comprised the months "JAN", "ÁPR", "MÁJ", "JÚN", "JÚL", "AUG", "OKT", "NOV" and "DEC", and the uppercase letters "F", "G", "H", "I", "J", "K", "L", "M" and "N". The resulting gap and inequity of distances between the stimuli caused by the omission of some letters and months should not affect our results, as Krajcsi and Kojouharova (2017) found that the distance effect in a comparison task utilizing new number symbols did not reflect the numerical distance between numbers (e.g., a distance of 2 between 3 and 5) but rather the distance relative to the subset of the stimuli presented.

2.5 Testing procedure

Each stimulus pair to be compared was simultaneously presented in white, centrally on a gray screen (see Figure 1). At the beginning of each block of trials (domain), instructions were given on the screen. All instructions encouraged participants to respond as fast and as accurately as possible. For the number comparison, participants were told to select the larger one of the two numbers presented on the screen. For the letters and months, they were instructed to select the letter or month that comes later in the alphabet or the year, respectively. All instructions were given in Hungarian. The stimuli were presented until the participant made a choice via the keyboard. Acoustic feedback was given to indicate to the participant whether the response was correct or incorrect. In between trials a blank, gray screen was shown for 700 ms (as in Krajcsi & Szűcs, 2022). After each block of trials (i.e., each condition), the participants were encouraged to take a break.

We automatically recorded the reaction time and whether or not the participant's response was correct for each trial.

Figure 1: Example of the stimulus presentation for numbers (top left), letters (top right) and months (bottom).

2.6 Follow-up questions

After the completion of all three blocks of trials, participants were directed to a follow-up questionnaire. They were asked for demographic information and we also tried to elucidate the strategies they used in the comparison tasks. For this latter purpose, we first asked participants to intuitively indicate whether they know the corresponding number to six letters (three of them had been used in the task) and the corresponding number to six months (three of these months had been used in the task). Afterwards, participants had to answer questions about their strategies to find out a) whether they transcoded letters or months to numbers, b) whether they used this strategy in the task and c) whether they represent the months of the year in a cyclic fashion (so that December comes before January).

The following questions were asked:

- What number corresponds to the letter "F"/”V”/”O”/”H”/”M”/”P”?
- What number corresponds to the month October/March/July/January/December/November?
- What (cognitive/mental) strategy did you use to solve the number comparison task?
- What (cognitive/mental) strategy did you use to solve the letter comparison task?
- What (cognitive/mental) strategies did you use to solve the month comparison task?
- I can easily convert letters of the alphabet to numbers (e.g., A = 1 etc.). [Answers: Yes; No; Only some; I don’t know]
- I can easily convert the months of the year to numbers (e.g., January = 1). [Answers: Yes; No; Only some; I don’t know]
- Did you convert the letters of the alphabet to numbers (e.g., A = 1) to solve the comparison task (e.g., A=1 and B=2, 2 > 1, so B comes after A)? [Answers: Yes; No; Other]
- Did you convert the months of the year to numbers (e.g., January = 1) to solve the comparison task (e.g., January=1 and February=2, 2 > 1, so February comes after January)? [Answers: Yes; No; Other]
- How do you represent months? [Answers: As a linear sequence that starts with January and ends with December; As a cycle so that December comes before January; Other]

The responses to the questions are descriptively reported in the results section below. Response times were not measured. Note that the indicated position of letters in the alphabet was deemed correct when it corresponded either to the position in the English alphabet or to that in the Hungarian alphabet.

2.7 Data preparation and analysis

All analyses were conducted in R version 4.2.2 (R Core Team, 2022). The packages ggplot2 (Wickham, 2016), dplyr (Wickham et al., 2023), splithalfr (Pronk et al., 2022), ie2misc (Embry et al., 2022), psychometric (Fletcher, 2022) and DescTools (Signorell, 2023) were used to plot the data, arrange and group the data, calculate Spearman-brown reliabilities, calculate the Median Absolute Deviation or to calculate confidence intervals of correlations and means, respectively.

Trials including M, N, January and December were excluded from all analyses, as visual inspection of the data revealed an end effect for the months January and December as well as for the last two letters, M and N. The end effect describes the phenomenon that, when a stimulus at the very beginning or the very end of the range employed in the task appears, the probability of this stimulus being the correct choice can take extreme values (e.g., 1 in the case of December and N, and 0 in the case of January), resulting in participants' error rate in these trials being disproportionally low.

To prevent the results from being distorted by random guessers, participants whose error rate lay more than two Median Absolute Deviations above the mean error rate in at least one domain were excluded from the analysis. Following this procedure, two participants were excluded due to too high error rate in the number trials, five were excluded for making too many errors in the letter trials, two participants committed too many errors in both the number and the month domain, another two participants made too many mistakes in the letter and month domain and one participants' error rate exceeded the threshold in all three domains. This resulted in a remaining sample size of 53 participants (35 females, 18 males; mean age = 21.65 years).

Based on the stimuli presented in each trial, the distance of each pair was calculated for the subsequent steps of the analysis. For example, the stimulus pairs "5 2" and "G J" were assigned a distance of 3. In the case of the months, omitted months (see section 2.4 Stimuli) were disregarded. That is, the pair "JAN ÁPR" was assigned a distance of 1 for the purpose of the analysis, irrespective of the missing months February and March. We considered only the distances within the task stimulus set, rather than those present in the complete list of months of the year. This decision was based on Krajcsi and Kojouharova's (2017) finding that, in a comparison task, newly learned number symbols' distances were perceived by participants based on the subset presented during the test rather than the numbers (values) they corresponded to.

In the next step, the mean error rates and the median reaction times were calculated per participant, domain (numbers, letters, months) and distance separately. Note that in all analyses involving reaction time, only correct trials were taken into account. On the basis of this, the individual distance effect slopes (i.e., the slopes of the relationship between distance and reaction time or error rate) were calculated for each participant. These slopes were tested for significant deviation from zero using one-sample t-tests – or one-sample Mann-Whitney-U-tests, in the event that a Shapiro-Wilk test indicated that the criterion of normality was not met.

To avoid an attenuation of the correlations between the slopes across domains due to low intra-individual reliability of the reaction time and error rate slopes, we calculated the reliability of the slopes to be able to later adjust the correlations to these (presumably imperfect) reliabilities (Spearman, 1904, 1910; cf. Krajcsi, 2017). For this purpose, each participant's trials were split into even and odd trails for each domain and the Spearman-Brown (Spearman, 1910) reliability between these two groups of trials was calculated. These reliabilities were subsequently used to disattenuate the correlations (see below).

In a last step, the confidence intervals of the disattenuated (i.e., reliability-adjusted) correlations were calculated for each domain combination (i.e., numbers – letters, numbers – months, letters – months) and response variable (reaction time and error rate). For each of these correlations, we calculated the 95% symmetrical confidence interval, based on the disattenuated correlation coefficient and the sample size, to check whether the reliability-adjusted intervals overlapped with zero. Whenever they did not overlap with zero, we concluded that the correlation was significant.

3. Results

3.1 Error rate and reaction time

Across all participants, the mean error rate was lowest in the number trials (0.026) and highest in the letter trials (0.085) with the month domain lying in between (0.0485). The same pattern emerged for the sample means of the individual medians of the reaction times. Participants were fastest to respond to number stimuli (536.829 ms), followed by the month trials (1106.366 ms) and the letter trials (1734.648 ms). A detailed overview of the mean error rates and reaction times per stimulus pair can be found in Tables 1 to 6.

Table 1: Means of the individual mean error rates (in %) per stimulus pair for the number trials. A green background shade indicates lower mean error rates, while a red background shade indicates higher mean error rates.

		Right								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Left	**1**		1.1%	0.8%	0.8%	0.4%	0.4%	0.8%	0.9%	0.6%
	2	1.3%		3.4%	2.1%	1.1%	1.1%	1.7%	1.1%	0.9%
	3	1.1%	1.9%		3.4%	3.4%	1.5%	1.7%	0.8%	0.4%
	4	0.8%	1.9%	1.5%		3.6%	2.3%	2.6%	0.6%	1.3%
	5	1.1%	1.7%	2.8%	4.7%		3.0%	4.2%	1.3%	2.5%
	6	0.4%	1.9%	1.1%	1.1%	3.6%		6.6%	4.3%	3.0%
	7	0.8%	2.5%	3.0%	3.2%	6.8%	10.8%		5.1%	4.8%
	8	1.3%	0.9%	1.5%	1.7%	2.6%	6.0%	8.9%		9.6%
	9	0.8%	0.9%	1.7%	1.7%	2.1%	3.2%	5.3%	9.1%	

Table 2: Means of the individual mean error rates (in %) per stimulus pair for the letter trials. The stimuli M and N were excluded and their values are therefore not depicted. A green background shade indicates lower mean error rates, while a red background shade indicates higher mean error rates.

		Right								
		F	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
Left	**F**		6.4%	6.4%	10.0%	7.4%	2.3%	2.8%		
	G	10.0%		13.6%	13.2%	11.9%	6.3%	4.0%		
	H	7.9%	15.7%		14.3%	10.2%	5.1%	4.3%		
	I	8.9%	16.8%	11.5%		6.2%	5.3%	6.6%		
	J	6.0%	14.2%	13.0%	10.2%		6.8%	8.3%		
	K	4.0%	7.0%	6.6%	9.8%	8.7%		10.6%		
	L	3.0%	6.4%	7.4%	6.8%	9.1%	13.6%			
	M									
	N									

Table 3: Means of the individual mean error rates (in %) per stimulus pair for the month trials. The stimuli Jan and Dec were excluded and their values are therefore not depicted. A green background shade indicates lower mean error rates, while a red background shade indicates higher mean error rates.

		Right								
		JAN	**APR**	**MAJ**	**JUN**	**JUL**	**AUG**	**OKT**	**NOV**	**DEC**
Left	**JAN**									
	APR			6.9%	7.9%	3.5%	2.7%	1.3%	0.8%	
	MAJ		6.5%		9.0%	4.4%	5.0%	2.9%	2.5%	
	JUN		9.2%	13.2%		1.5%	3.3%	1.5%	2.5%	
	JUL		5.2%	3.6%	1.0%		4.0%	4.2%	3.1%	
	AUG		0.8%	2.5%	2.1%	8.3%		11.7%	10.2%	
	OKT		1.5%	2.9%	1.5%	3.3%	14.2%		9.4%	
	NOV		1.7%	2.7%	1.7%	2.9%	11.9%	9.0%		
	DEC									

Table 4: Means of the individual medians of the reaction time (in ms) per stimulus pair for the number trials. A green background shade indicates shorter mean reaction times, while a red background shade indicates longer mean reaction times.

		Right								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Left	**1**		539	530	505	493	492	511	493	483
	2	559		610	544	549	527	540	502	500
	3	510	594		610	559	524	547	523	511
	4	497	549	582		597	581	554	551	526
	5	512	539	558	604		591	587	547	543
	6	511	507	531	570	598		657	589	583
	7	501	527	565	563	590	673		634	592
	8	484	500	513	535	536	607	643		647
	9	495	509	509	542	530	583	588	655	

Table 5: Means of the individual medians of the reaction time (in ms) per stimulus pair for the letter trials. The stimuli M and N were excluded and their values are therefore not depicted. A green background shade indicates shorter mean reaction times, while a red background shade indicates longer mean reaction times.

		Right								
		F	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**
Left	**F**		1737	1905	1634	1596	1470	1371		
	G	1909		2060	1984	1948	1578	1529		
	H	1935	2360		2093	2098	1850	1621		
	I	1694	2002	2123		1890	1710	1652		
	J	1737	2057	2318	2048		1736	1627		
	K	1603	1771	2044	1833	2014		1902		
	L	1454	1658	1899	1817	2014	2134			
	M									
	N									

Table 6: Means of the individual medians of the reaction time (in ms) per stimulus pair for the month trials. The stimuli Jan and Dec were excluded and their values are therefore not depicted. A green background shade indicates shorter mean reaction times, while a red background shade indicates longer mean reaction times.

		Right								
		JAN	APR	MAJ	JUN	JUL	AUG	OKT	NOV	DEC
Left	JAN									
	APR			1347	1353	1233	1125	1023	960	
	MAJ		1445		1298	1256	1141	1053	999	
	JUN		1413	1375		1131	1183	1026	1011	
	JUL		1311	1274	1212		1124	1050	1004	
	AUG		1153	1182	1167	1218		1177	1104	
	OKT		1072	1082	1044	1073	1285		1170	
	NOV		950	968	959	936	1124	1208		
	DEC									

3.2 Distance effects

The individual slopes of the distance effect for the error rate in the number domain differed significantly from zero (Mann-Whitney U-test: $p < 0.001$; see Figure 2). The mean of all slopes, i.e., average distance effect across all participants, was -0.005 with the 95%-confidence interval ranging from -0.007 to -0.004.

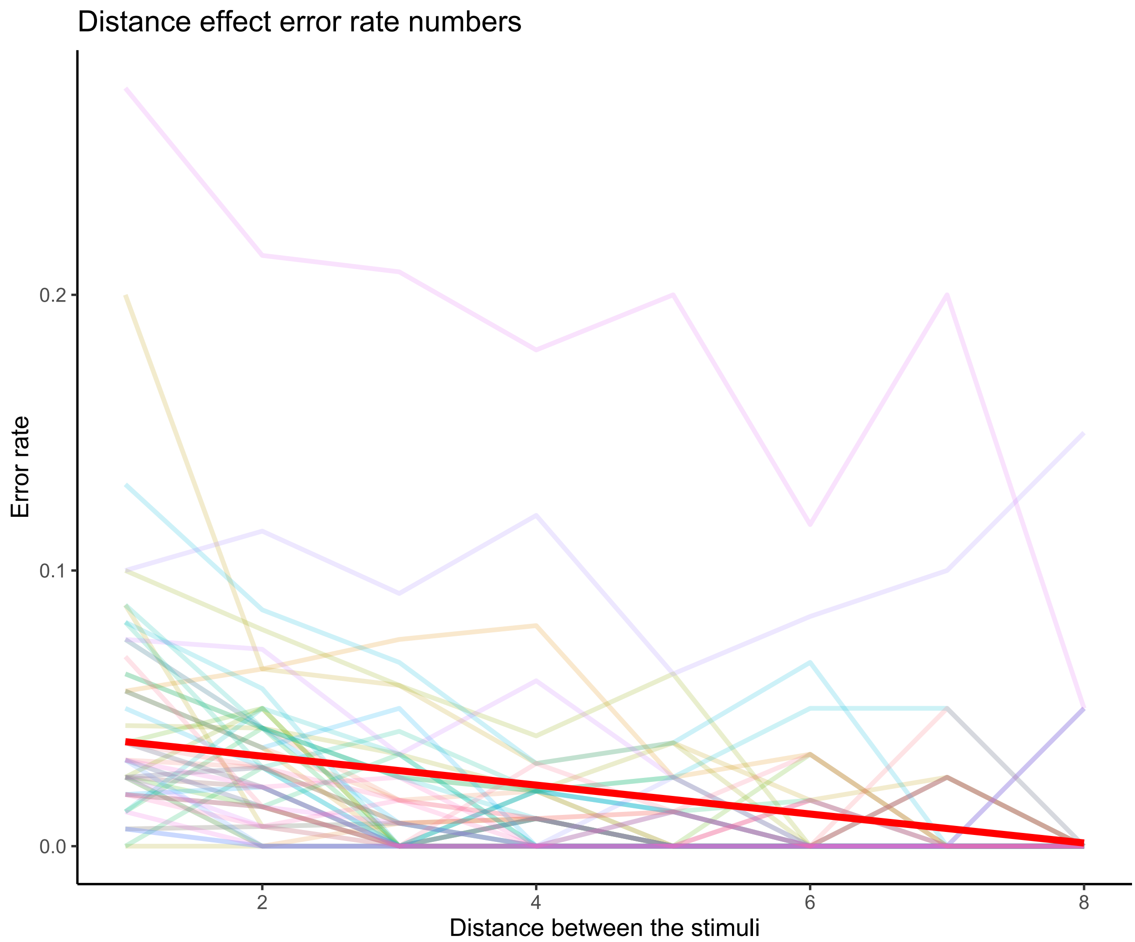

Figure 2: Distance effect for participants' mean error rate in the number domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

Similarly, a significant (Mann-Whitney U-test: $p < 0.001$) distance effect emerged for the error rate in the letter trials (see Figure 3). For the letters, the mean of all individual distance effect slopes amounted to -0.017, with the 95% -confidence interval ranging from -0.022 to -0.012.

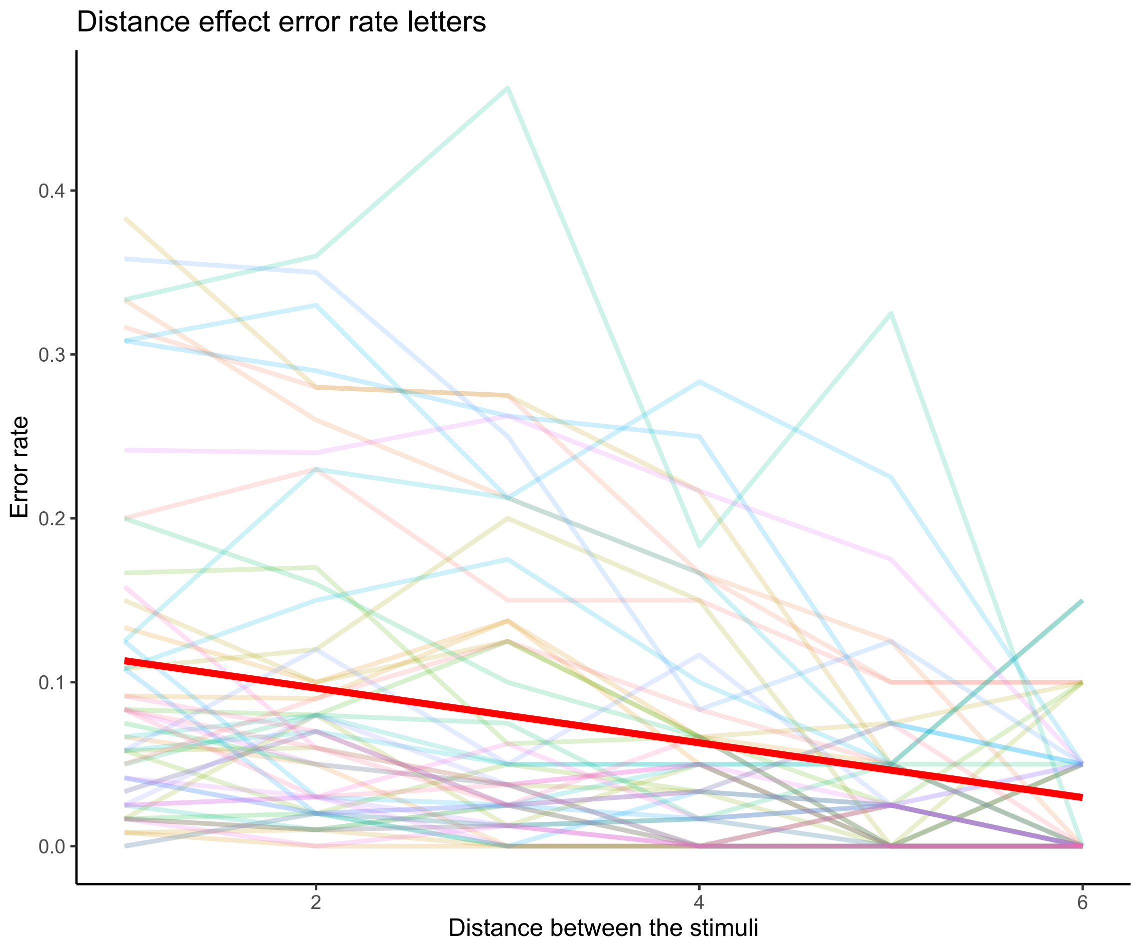

Figure 3: Distance effect for participants' mean error rate in the letter domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

Also for the month trials, error rates significantly declined with increasing distance between the stimuli (t-test: $p < 0.001$; see Figure 4). Across all participants, the slope of the distance effect was -0.013, with the 95%-confidence interval ranging from -0,016 to -0.011.

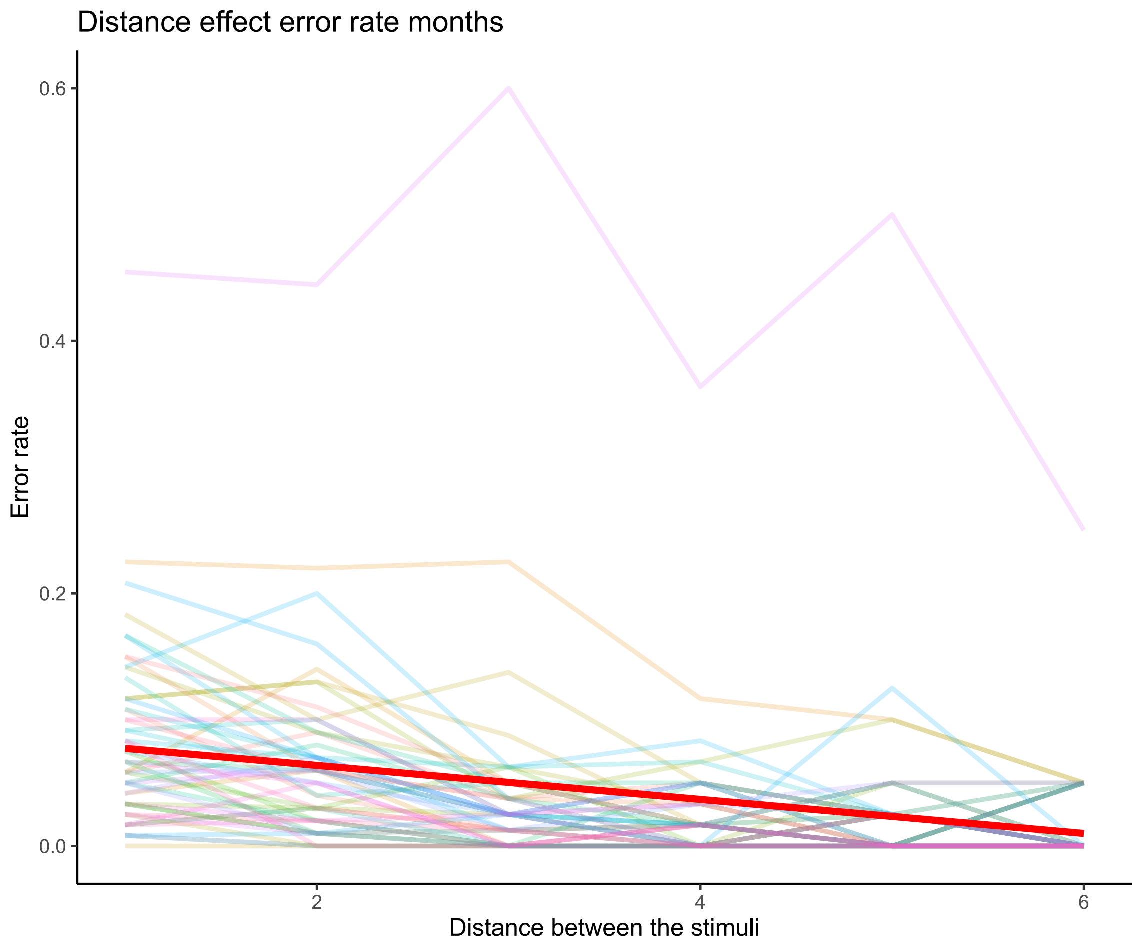

Figure 4: Distance effect for participants' mean error rate in the month domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

Turning to the reaction time, we detected a significant (Mann-Whitney U-test: $p < 0.001$) distance effect for the reaction time in the number domain, see Figure 5. The mean slope of the individual distance effects was -15.300, with the confidence interval ranging from -17.609 to -12.991.

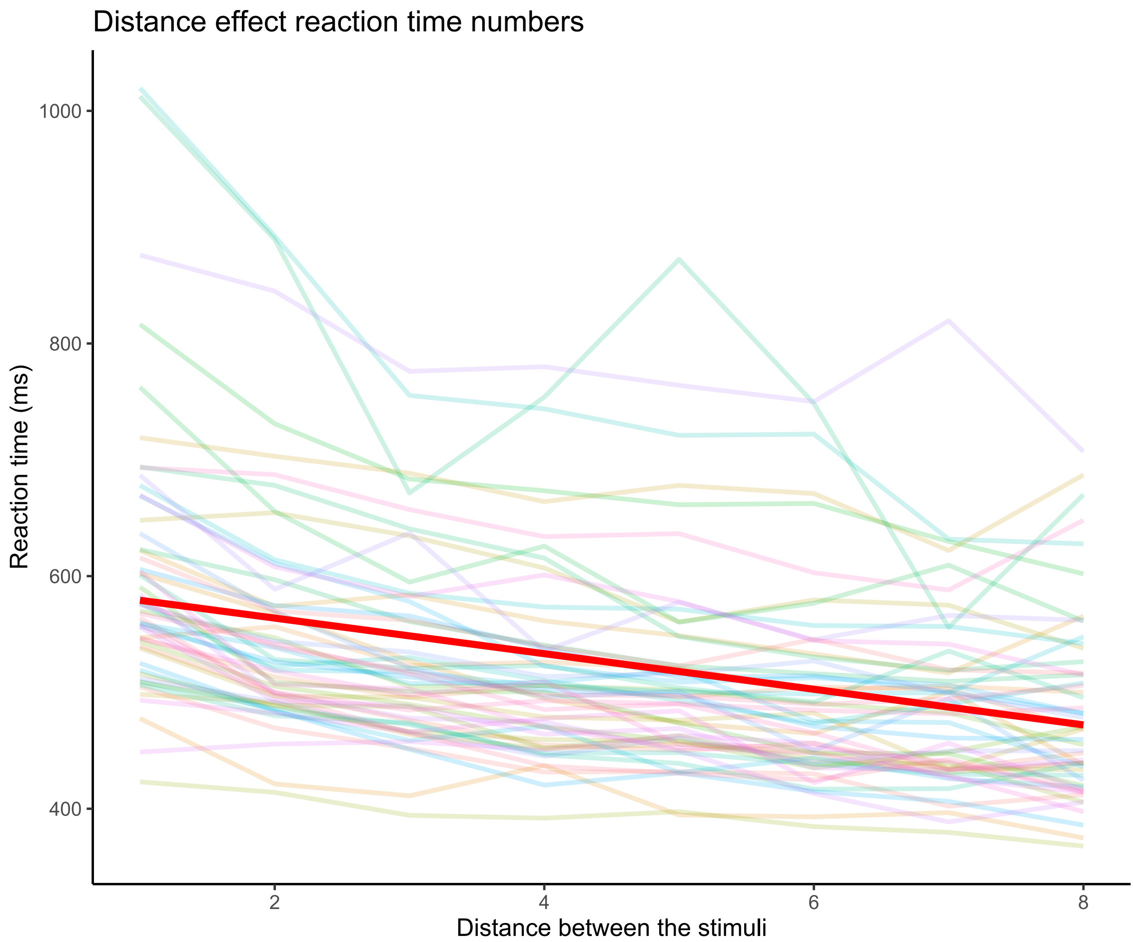

Figure 5: Distance effect for participants' median reaction time (in ms) in the number domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

As for the numbers, the distance effect for the reaction time was also highly significant (Mann-Whitney U-test: $p < 0.001$) for letter trials (see Figure 6). For these trials, the mean of all individual slopes was -107.425, with the 95%-confidence interval ranging from -135.820 to -79.030.

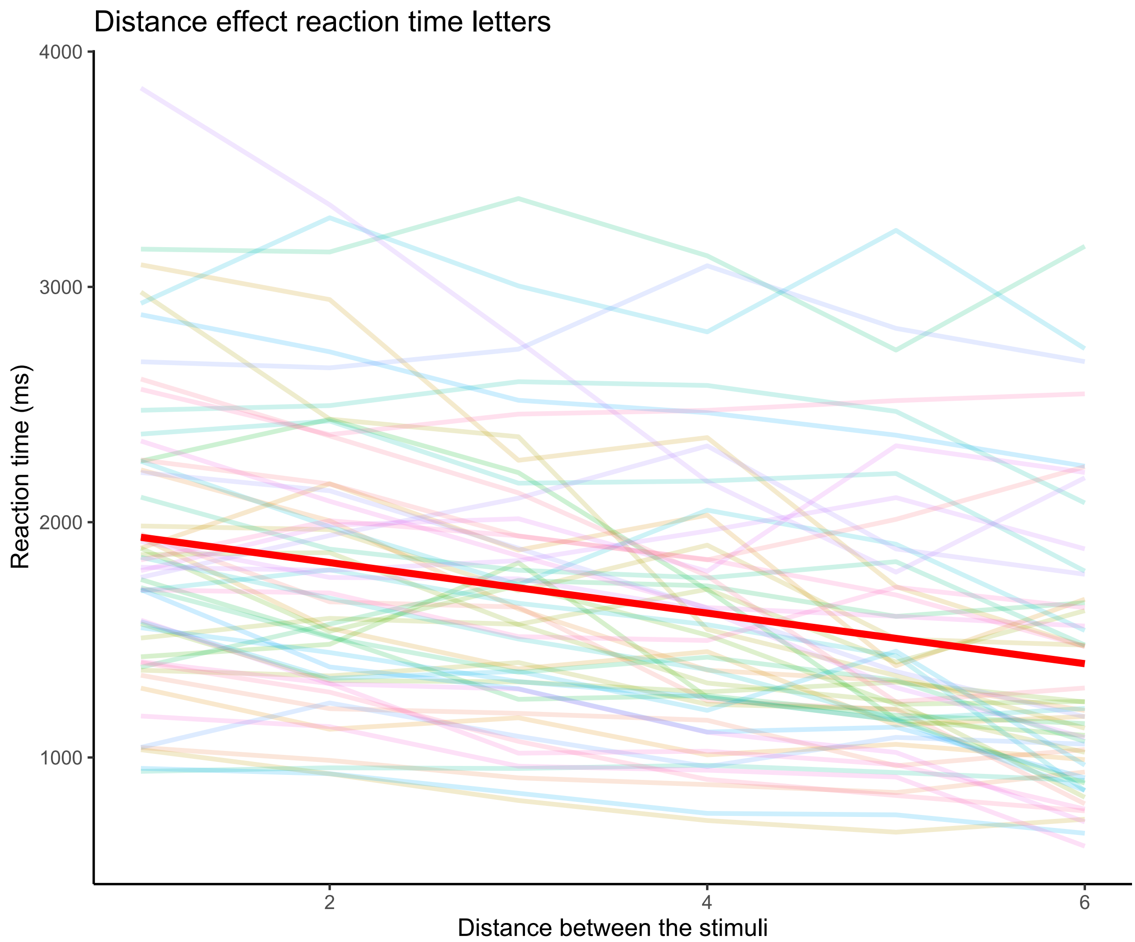

Figure 6: Distance effect for participants' median reaction time (in ms) in the letter domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

Finally, we also observed a significant distance effect for the reaction time when participants compared months of the year (t-test: $p < 0.001$; see Figure 7). For the reaction time in month trials, the mean of all individual distance effect slopes amounted to -50.464, with the 95%-confidence interval ranging from -59.863 to -41.064.

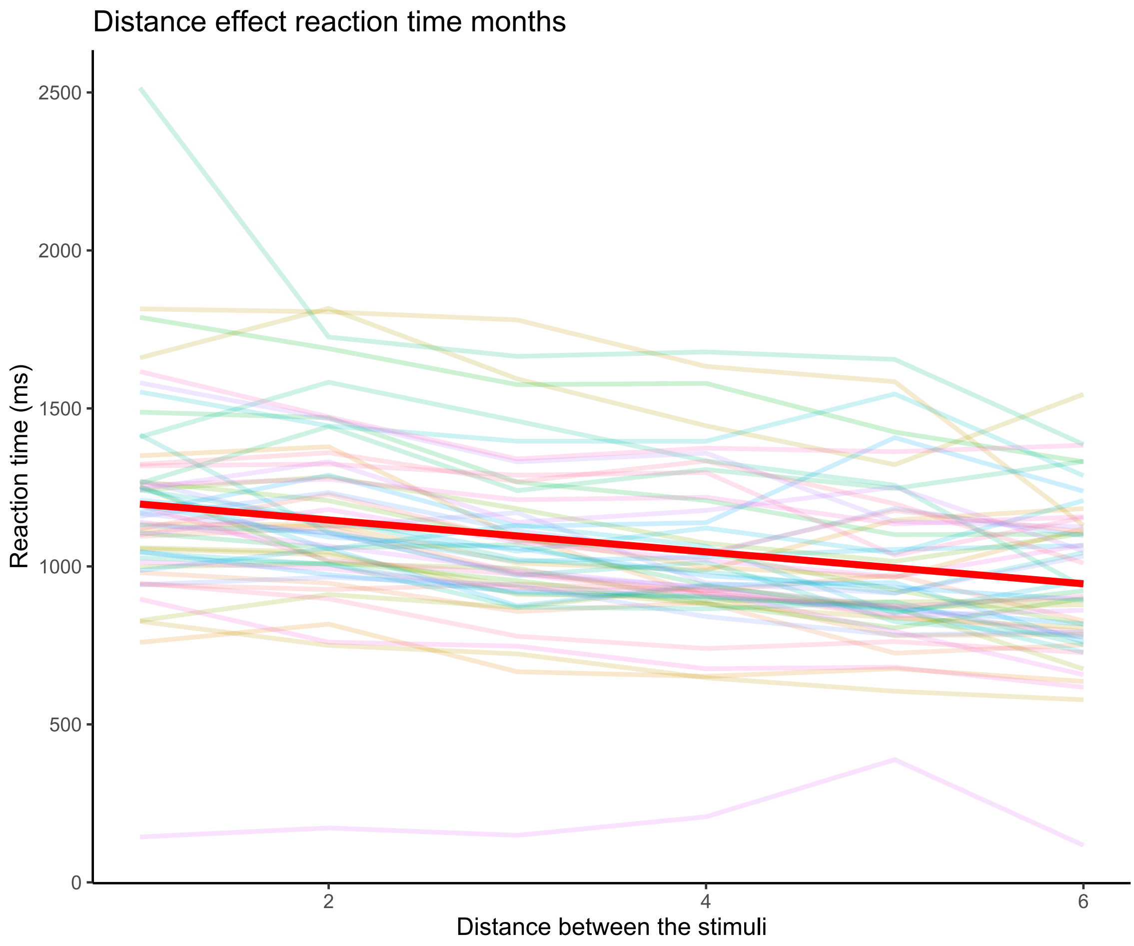

Figure 7: Distance effect for participants' median reaction time (in ms) in the month domain. Each line corresponds to one individual participant. The mean distance effect slope is depicted in red.

3.3 Reliabilities

The intra-individual Spearman-Brown reliabilities between the even and odd trials can be seen in Table 7. Reliabilities for the error rate and reaction time were high in the number and letter domain, while they were only sufficient for the month comparison task.

Table 7: Intra-individual Spearman-Brown reliabilities for all three domains and both response variables. A value of 1 indicates perfect reliability.

Domain	***Error rate***	***Reaction time***
Numbers	0.808	0.795
Letters	0.807	0.708
Months	0.629	0.487

3.4 Correlations

A correlation matrix summarizing the correlation coefficients between the distance effect slopes of the three domains for both response variables, error rate and reaction time, is provided in Table 8. More details can be found in the following sections.

Table 8: Correlation matrix containing the disattenuated correlations between the distance effects for numbers, letters and months. Correlations of the error rate distance effects are given in the cells with blue frames (above the gray diagonal), while the correlations between the reaction time distance effects can be found in the cells with orange frames. A darker shade of green indicates a stronger correlation.

	Numbers	Letters	Months
Numbers		0.351	0.834
Letters	0.044		0.711
Months	0.306	0.532	

3.4.1 Numbers and letters – error rate

The 95% confidence interval of the raw (not disattenuated/reliability-adjusted) correlation between the distance effect slopes of the error rate in the number domain and those in the letter domain ranged from 0.014 to 0.514, while the confidence interval of the disattenuated correlation ranged from 0.089 to 0.567. Therefore, the correlation's deviation from zero was significant, despite the correlation remaining weak. The relationship between the slopes of the two domains is visualized in Figure 8.

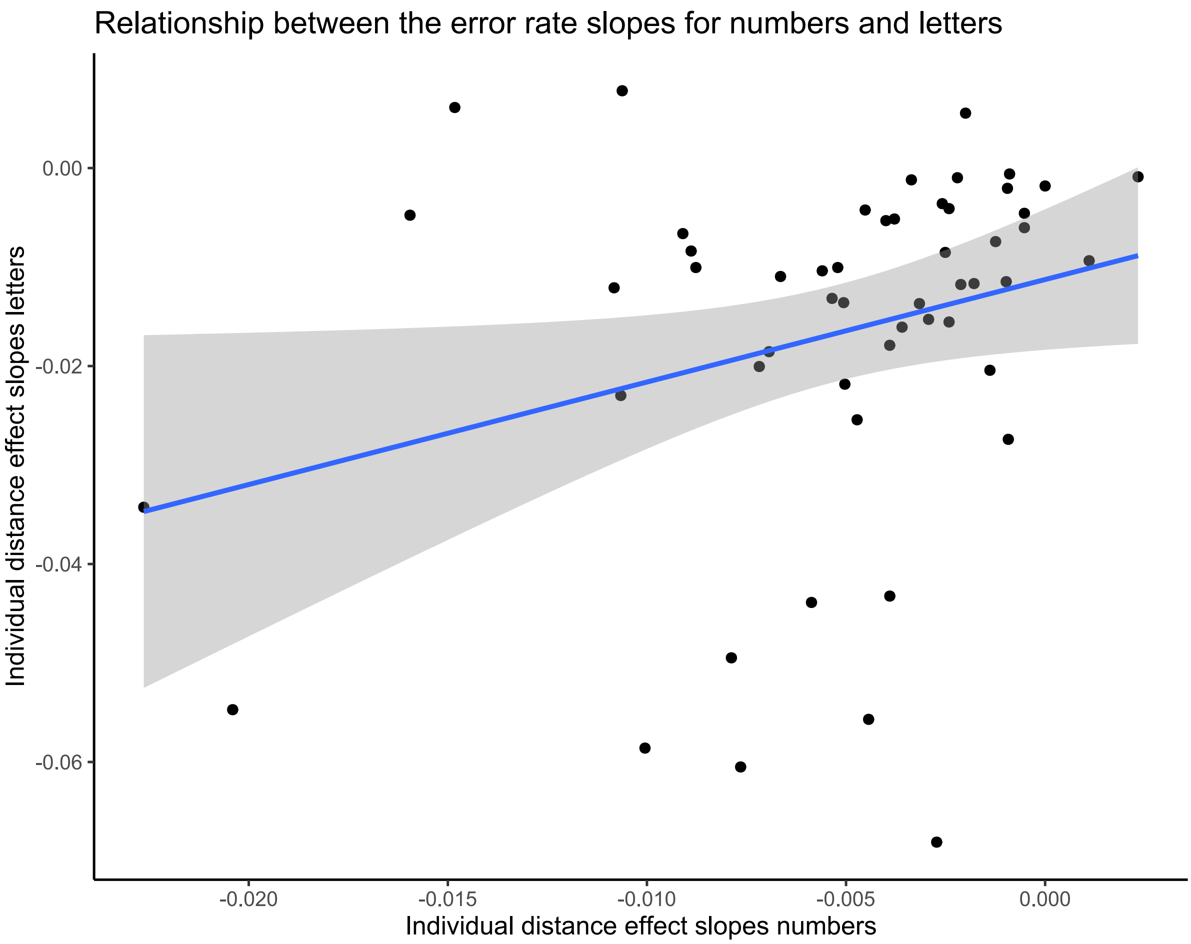

Figure 8: Relationship between the individual distance effect slopes for the error rate between the number domain and the letter domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.4.2 Numbers and months – error rate

A strong correlation emerged between the individual distance effect slopes for the error rate in the number domain and the slopes in the month domain (see Figure 9), which deviated significantly from zero, with the 95% confidence interval of the raw correlation ranging from 0.387 to 0.745 and the confidence interval of the disattenuated correlation ranging from 0.728 to 0.901.

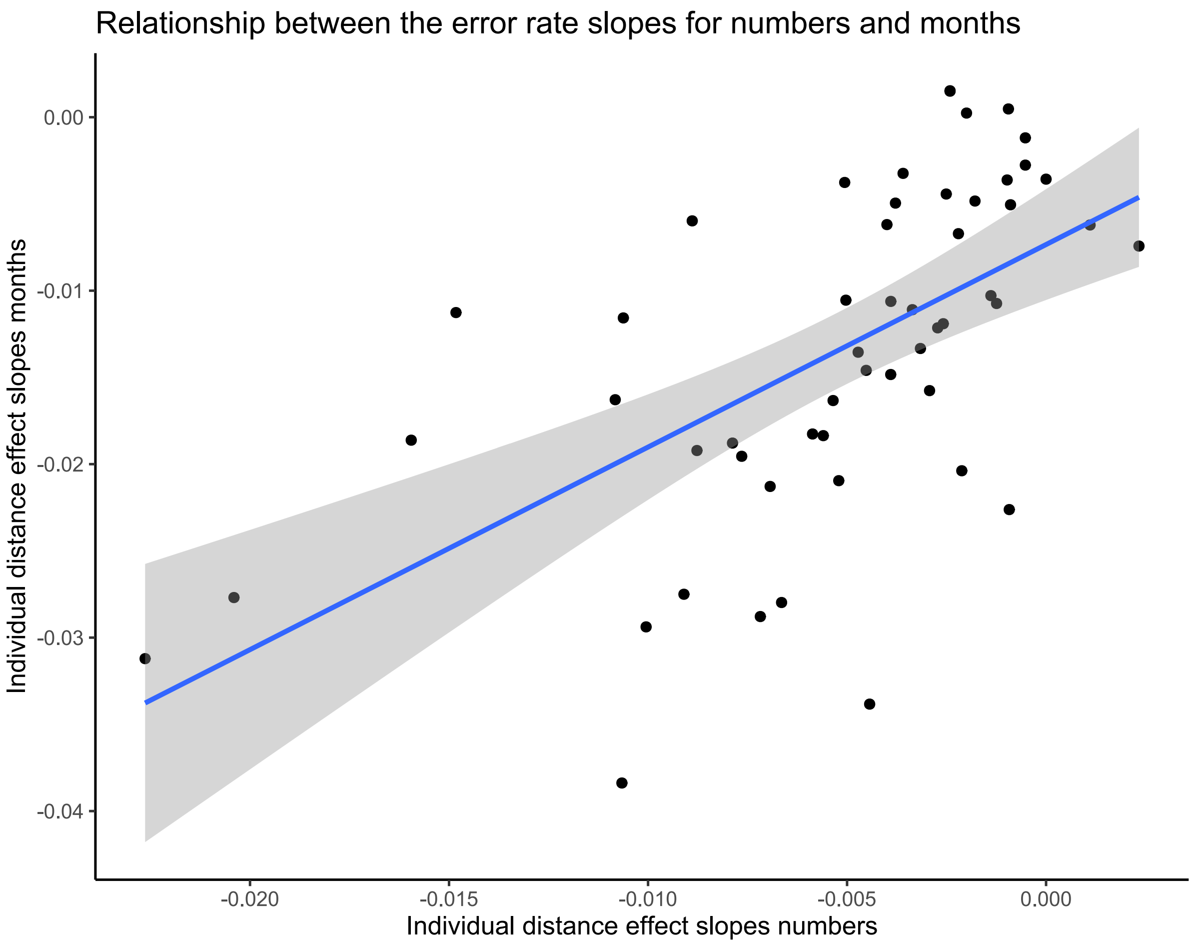

Figure 9: Relationship between the individual distance effect slopes for the error rate between the number domain and the month domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.4.3 Letters and months – error rate

Similarly, the correlation between the individual distance effect slopes for the error rate of the letter and month domains deviated significantly from zero, even though the correlation was not as strong as between numbers and months. The 95% confidence interval of the raw correlation ranged from 0.274 to 0.683 and that of the disattenuated correlation ranged from 0.546 to 0.823. The relationship is illustrated in Figure 10.

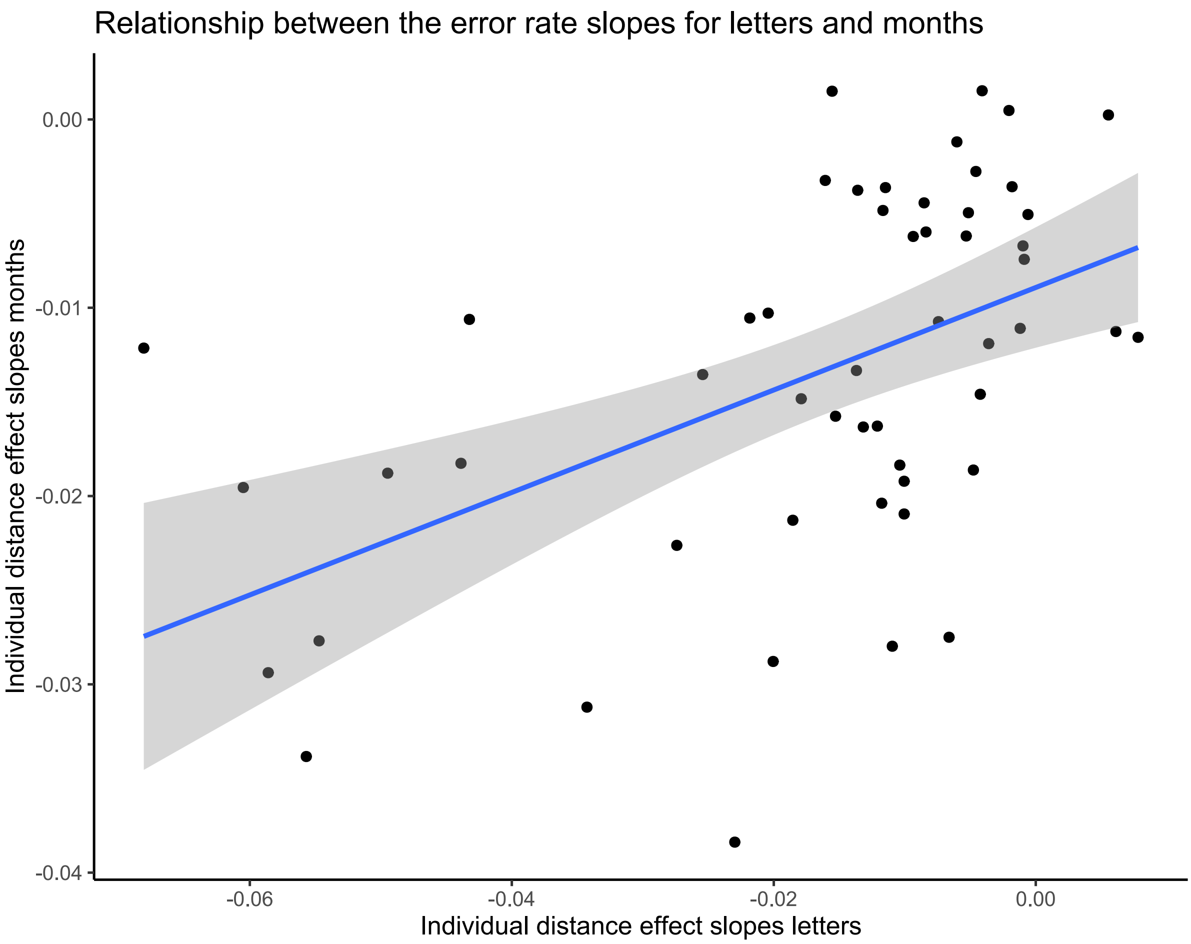

Figure 10: Relationship between the individual distance effect slopes for the error rate between the letter domain and the month domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.4.4 Numbers and letters – reaction time

The individual distance effect slopes for participants' median reaction times did not correlate significantly between numbers and letters. The 95% confidence interval of the raw correlation ranged from -0.238 to 0.302. Given that also the 95% confidence interval of the disattenuated correlation, which ranged from -0.229 to 0.310, overlapped with zero, we can conclude that the correlation does not significantly deviate from zero. This lack of significant correlation is also reflected in Figure 11.

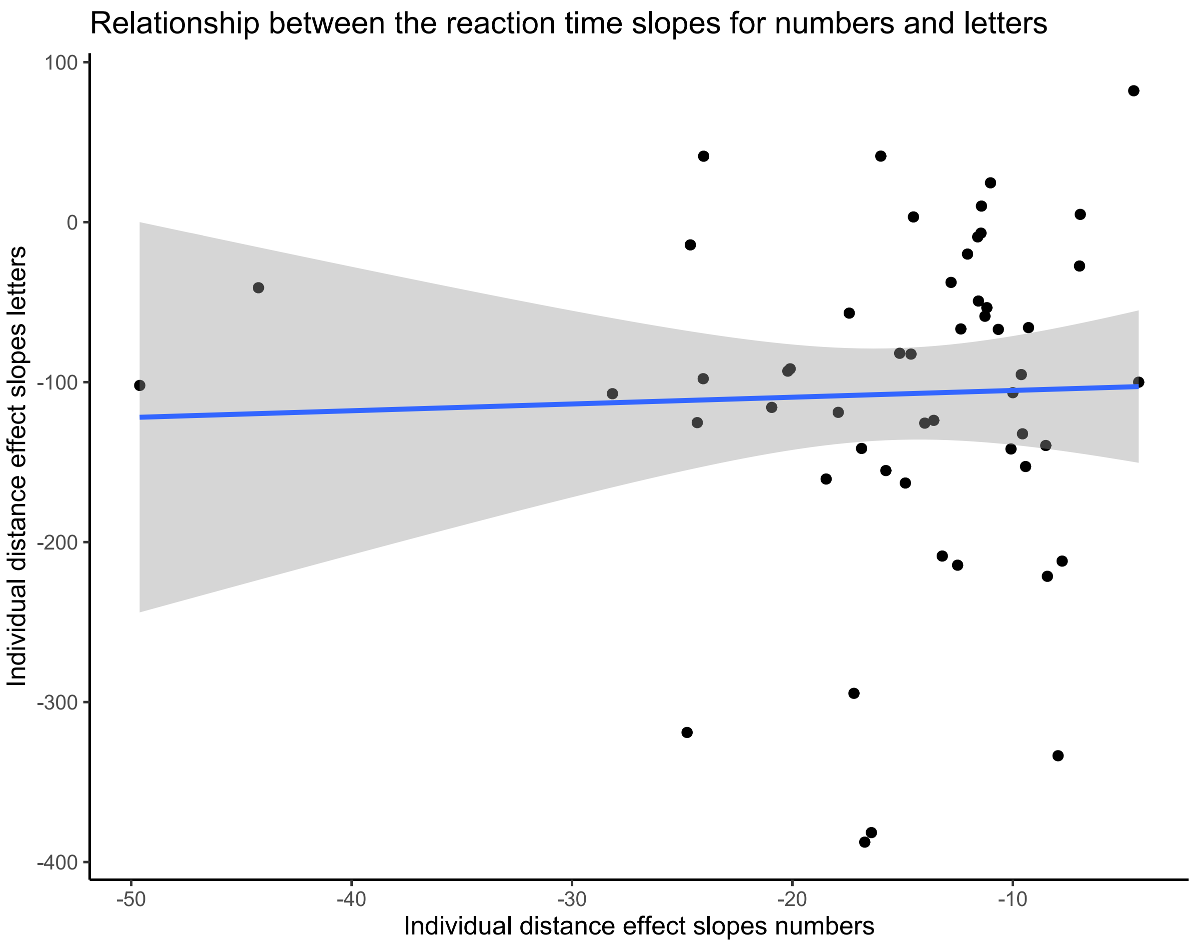

Figure 11: Relationship between the individual distance effect slopes for the median reaction time between the number domain and the letter domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.4.5 Numbers and months – reaction time

We found a weak correlation between numbers and months regarding the distance effect slopes for the reaction times (see Figure 12). The 95% confidence interval of the raw correlation ranged from -0.084 to 0.438. In contrast, the confidence interval of the disattenuated correlation ranged from 0.039 to 0.532, which means that the correlation's deviation from zero was weak but significant.

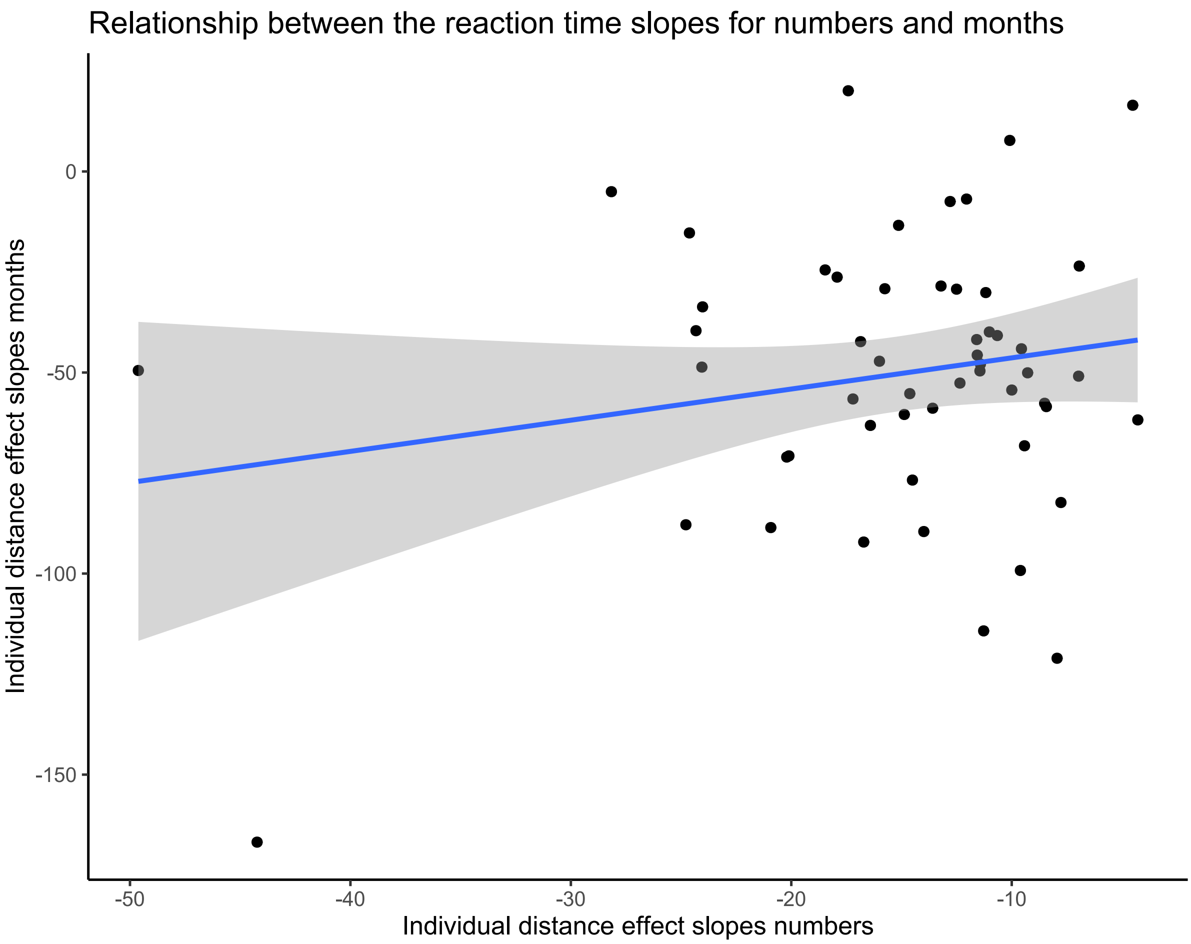

Figure 12: Relationship between the individual distance effect slopes for the median reaction time between the number domain and the month domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.4.6 Letters and months – reaction time

Finally, the distance effect slopes for the reaction time were moderately correlated between letters and months (see Figure 13). That is, the 95% confidence interval of the raw correlation ranged from 0.063 to 0.549 and that of the disattenuated correlation ranged from 0.305 to 0.701 and, therefore, the correlation differed significantly from zero.

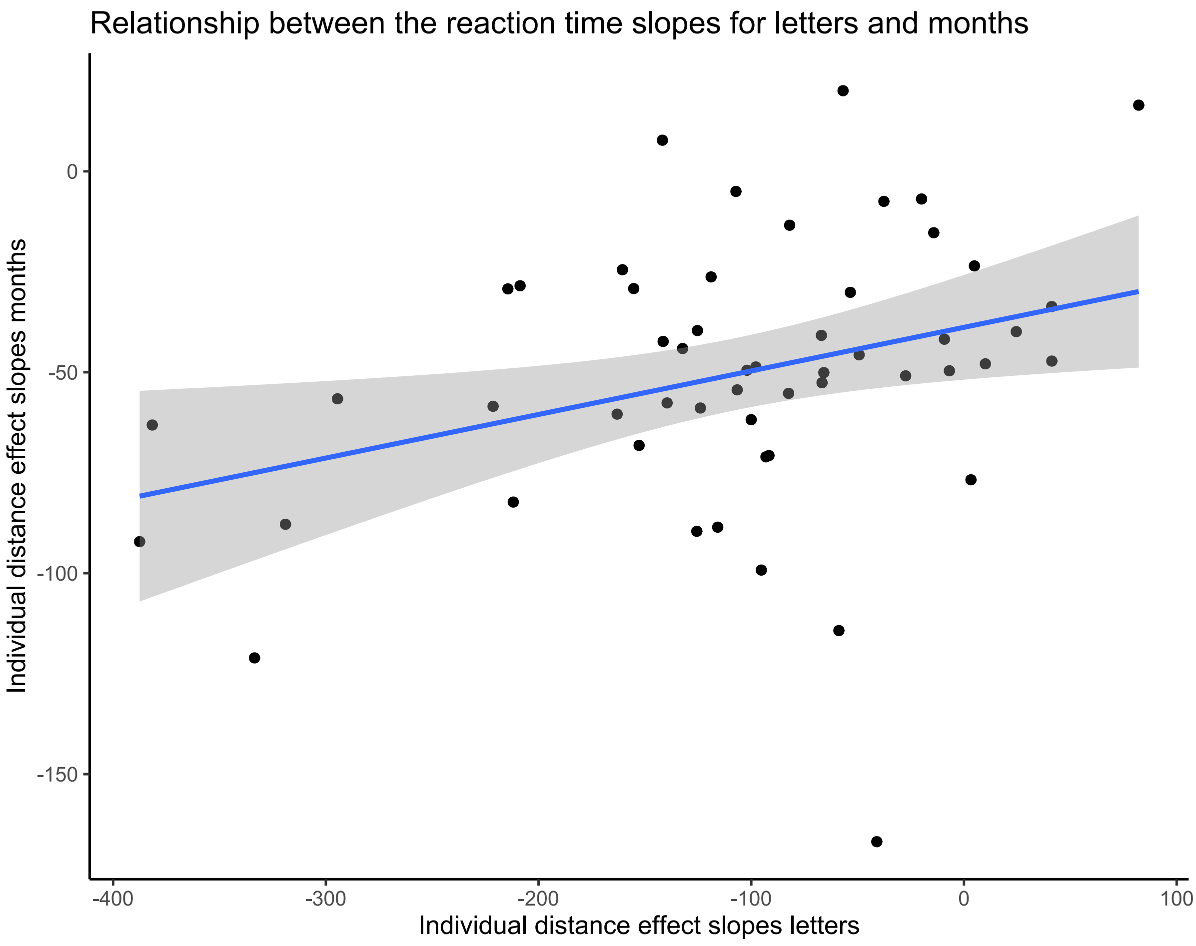

Figure 13: Relationship between the individual distance effect slopes for the median reaction time between the letter domain and the month domain. Each dot represents the individual distance effect slopes of one participant. Note that the raw, i.e., not yet reliability-adjusted, correlation is depicted here.

3.5 Responses to the questionnaire

Out of the 65 participants that completed the main task, 59 filled out the follow-up questionnaire. In the following, these participants' responses to the questions are reported descriptively.

3.5.1 Transcoding letters to numbers

Even though only three participants stated that they can easily and spontaneously convert (all) letters to numbers (e.g., A = 1) and only 16 participants felt confident about at least some letters, participants generally showed a high success rate when asked for the number corresponding to specific letters. 56% of participants gave the correct response (both the position in the English and in the Hungarian alphabet were counted correct) for the letter "F", 35% gave the correct number corresponding to the letter "V", 36% could convert "O" to a number, 52% correctly identified the number corresponding to the letter “H” and 38% gave the correct response for “M” and “P”, respectively. While no participant spontaneously reported (in response to an open-ended question) having used a number-based strategy, six participants disclosed having transcoded letters to numbers at least occasionally when explicitly asked about this strategy.

3.5.2 Transcoding months to numbers

Given that months are routinely associated with and replaced by numbers, it comes with little surprise that converting months to numbers seemingly posed less of a challenge to participants than transcoding letters to numbers. In total, 52 participants stated that they can easily convert at least some months to numbers. When asked about specific months, for the months October, March, July, January and November more than 98% of participants correctly entered the corresponding number (e.g., January = 1), while the success rate was even at 100% for the month December. Six participants spontaneously reported having transcoded months to numbers to solve the task and 29 participants agreed that they used this strategy in at least some of the trials.

3.5.3 Mental representation of the months and of the year

As participants’ mental image of the months and of their order within the year can potentially influence the results of the comparison task, we explicitly asked about mental representations. That is, some participants might represent the year linearly, i.e., a sequence starting with January and ending with December, while for others it might be intuitive that months are arranged in a circular fashion, with January coming after December (as discussed in Vos et al., 2017). In response to this question, 14 participants reported a circular representation, whereas the remaining 45 participants picture the months within the year more or less linearly. In response to the open-ended question about their strategy in the month comparison task, two participants spontaneously mentioned (linear) mental imagery and one participant referred to a mental circular representation of the year that enabled them to solve the comparison task.

3.5.4 Strategies in the number comparison task

When asked to describe the strategy they applied to solve the number comparison task, the vast majority of participants did not give concrete answers other than having tried to identify the larger one of the numbers (i.e., reiterating the instruction) or referred to intuition. Four participants mentioned an end effect, i.e., they reported having realized that 1 is always the incorrect and 9 always the correct answer if these stimuli were included in a pair. One participant mentioned a verbal strategy that consisted in repeating the depicted numbers aloud to be better able to comprehend them. Four participants expressed some confusion over the instructions of the task, e.g., because they erroneously assumed they should focus on font size rather than numerical size.

3.5.5 Strategies in the letter comparison task

As opposed to the number comparison, fewer participants, namely 15, described no particular strategy but rather attributed their success with the letter trials to intuition. Another 15 participants explicitly mentioned having recited the alphabet (silently or aloud) for at least some of the trials. An end effect, i.e., that the very first (F and G) and last (M and N) elements were easier to process, was reflected in 19 of the responses. Eight participants mentioned a grouping principle, by which they subdivided the stimulus range to allow for faster comparison. Other strategies, namely using proper names associated with these letters, focusing on the shape of the letters and writing down the relevant part of the alphabet on paper, were mentioned by one participant each. It is important to note that many participants reported more than one strategy and some explicitly mentioned having switched strategies either across trials (i.e., with time) or across stimuli (depending on the exact letters presented).

3.5.6 Strategies in the month comparison task

At least for some month trials, 18 participants seem to have responded intuitively and spontaneously without being aware of a particular strategy. Statements hinting at an end effect, i.e., that January and December were easier to compare than other months, occurred 19 times. As mentioned above, six participants reported having transcoded months to numbers and three people visualized the months in front of them, one of which in a circular fashion. As opposed to the letter trials, only two participants stated having recited the list of months and three read the presented stimuli out aloud. Three participants expressed confusion over and/or reported having applied specific strategies for the months whose names look very similar, i.e., January (JAN), June (JÚN) and July (JÚL). Most interestingly, a clustering/grouping strategy was more prevalent, or at least recalled more frequently, than in the other domains. A total of 21 participants spontaneously referred to seasons when describing their strategy. That is, they first grouped months into seasons (e.g., April – spring, October – autumn) to facilitate cross-season comparisons.

4. Discussion

The present project aimed at assessing whether the discrete semantic system is limited to symbolic number processing or whether it is domain-general. More precisely, we enquired whether the association-based mechanisms postulated by the DSS can also account for comparison distance effects in the domains of letters and months. In line with the findings of previous studies, we expected to see a significant distance effect in all domains. In addition, we predicted that, if number, letter and month comparisons indeed rely on the same system, the distance effects should not only be present, but the distance effect slopes of individual participants' error rate and reaction time should also correlate strongly across the three domains. We found a distance effect – for both error rate and reaction time – in all three domains. However, the pairwise correlations between the distance effects of the domains only partially support the view that the DSS is domain-general. A more detailed overview of the findings is given in the following.

4.1 Overview of the results and interpretation in light of previous work

We observed significant distance effects in all three domains for both response variables (error rate and reaction time), thus corroborating previous evidence of distance effects in letter and month comparison and/or ordering tasks (Anonymous, unpublished; Fairbank, 1969; Franklin et al., 2009; Fulbright et al., 2003; Goffin et al., 2020; Morsanyi et al., 2016; Parkman, 1971; Van Opstal et al., 2008; Vos et al., 2017). Also consistent with these previous works is that comparing numbers was seemingly easiest and comparing letters most difficult for participants. First, the mean error rate across all participants in month trials and especially in letter trials was considerably higher than in number trials. Second, even within the correct trials, on average, participants took almost twice as long to respond to month stimuli and more than three times as long to identify the later one of two letters than to successfully compare two numbers. Finally, this is also reflected in the distance effect, in that the mean distance effect slopes were steepest for letters and flattest for numbers. In summary, participants committed most mistakes, took longest to respond correctly and showed the starkest relative increase in performance with increasing distance in the letter domain, whereas they were most accurate and quickest with the numbers, the stimulus type they presumably compare most routinely in every-day life.

Simply detecting a distance effect in all three domains is, however, insufficient to gain insights into the processing of the stimuli and, thereby, the origin of the distance effects. In analogy to the distance effects in symbolic and non-symbolic number comparison tasks (Krajcsi et al., 2016, 2018; Krajcsi & Kojouharova, 2017), it is conceivable that, albeit present in all domains, the distance effects in number, letter and month comparison tasks are driven by distinct mechanisms (cf. Lyons et al., 2016). Hence, only the correlations between the individual distance effect slopes across the three domains can reveal whether the DSS is indeed the source of all three distance effects.

The results of the correlational analyses lend partial support to the potential domain-generality of the DSS. The strongest evidence for the domain-generality of the DSS would have consisted in a high correlation between number and letter distance effects, as letters are presumably less likely to be transcoded to numbers during the task than are months (Fias et al., 2007). Counter to our predictions and counter to neuro-imaging data attributing number and letter comparison to shared neural structures (Fias et al., 2007; but see Zorzi et al., 2011), the correlation between the number and letter distance effects was weak for the error rate and non-existent for the reaction time.

Importantly, the weakness or absence of these correlations cannot simply be attributed to a lack of reliability, i.e., excessive noise, in our data, since we applied a disattenuation procedure which corrects the correlations according to the reliabilities of the potential correlates. In light of the lack of strong correlations between number and letter distance effects, we can, therefore, not confidently conclude that the DSS also supports letter comparison.

Numbers and months, in contrast, appear to be processed highly similarly in comparison tasks. For this combination of domains, we found a strong correlation between the error rate distance effect slopes and a moderate correlation between the reaction time distance effect slopes. These findings are in line with the idea that the distance effects in number and month comparison tasks rely on a shared mechanism.

Nevertheless, one could argue that this correlation does not necessarily reflect any domain-generality of the DSS, but that months are processed by the DSS in virtue of them being transcoded to numbers. That is, as outlined in the introduction and methods section, months are often used interchangeably with the number that represents their position within the year. If participants transcoded months to numbers to solve the task, the observed correlation would not be particularly surprising: participants would simply process numbers in both cases but take a “detour” in the month comparison task. Indeed, about half of the participants indicated in the questionnaire that they at least occasionally transcoded months to numbers to solve the task.

Still, the conclusion that the transcoding strategy alone, and not a domain-general DSS, can account for the correlations of the number and month distance effect slopes is not fully supported by our data. One noteworthy consideration is that, if months are transcoded to numbers in the comparison task, one would expect that participants’ reaction time in month trials would be substantially lengthened by the additional processing step. Even though the mean reaction time in the month domain (1106.366 ms) exceeded that in the number domain (536.829 ms), it is questionable whether the time would have sufficed to transcode months to numbers before responding. In support of this doubt, Fias et al. (2007) found that participants took on average 2276 ms to solve a letter comparison task by transcoding *letters* to numbers. This is considerably longer than the reaction times we observed. However, months are presumably more easily (and faster) transcoded to numbers than letters. This is also reflected in our participants’ higher accuracy in transcoding months, rather than letters, to numbers in the follow-up questionnaire.

However, if participants transcoded months but not letters to numbers in our study, there is seemingly no reason why the distance effect slopes should correlate between letters and months. In contrast, this is exactly what we found. The moderate to strong correlations between the error rate and reaction time distance effects of letter and month trials suggest that months are processed very similarly to letters. What remains perplexing is how it is possible that months correlate strongly (at least in terms of the error rate) with both numbers and letters, but numbers and letters only correlate very weakly with one another. After all, if month and number comparisons rely on the same mechanism and letter and month comparisons rely on the same mechanism, it is difficult to imagine how number and letter comparisons could rely on *different* mechanisms. There are several tentative explanations that could be put forward to unravel our patterns of results, especially the unexpectedly weak correlations between number and letter distance effect slopes.

4.2 Strategies beyond the DSS and limitations

The weak correlation between number and letter distance effect slopes is not necessarily at odds with the domain-generality of the DSS. Instead, some of the assumption on which we built our predictions might be flawed.

An important caveat potentially affecting the correlation is that processing, and in particular comparing, letters might be a less automatic process than in the case of numbers (Fulbright et al., 2003; Lyons et al., 2016). Developing the argument even further, one could imagine that this lack of familiarity could affect participants' strategy in the comparison task. In analogy to young children that successfully learn how to recite the counting list before they grasp the cardinality principle (i.e., that the last number of a counting list denotes the numerosity of the counted set) (Wynn, 1990), participants in our study might have relied on lists more strongly when comparing letters and months than when comparing numbers. Two important limitations of this line of explanation should be noted though. For one thing, ordinality persists to be of crucial importance to number processing too, even after having mastered the cardinality principle (Sasanguie & Vos, 2018); and for another, as noted above, if the letter and number comparison were simply more difficult for participants (which is supported by our data), this difference should not *per se* impact the correlations as long as the relative difficulty of the three tasks is consistent across individual participants.

What could, however, affect the correlations between the domains is *intra*-individual variation of strategies. For example, individual participants could have utilized a list-retrieval strategy (reciting the alphabet or the months of the year) for some trials but not for others. Indeed, many participants stated having switched strategies across trials, either because familiarity with the task increased or because of the specific stimuli presented to them. In addition, strategy switching has also been reported for multiplication tasks. In this context, it has been hypothesized that participants resort to a backup strategy whenever retrieval of the multiplication result from a network of associations does not succeed or takes too long (Siegler, 1988; see also Ashcraft, 1987). Even more interestingly, Hamilton and Sanford (1978) asked participants about their strategy after each trial of a two-letter ordering task. They found that participants recited the alphabet more frequently to reach the solution in short-distance trials than in long-distance trials, which the authors therefore conclude to be the cause of the observed standard distance effect. A similar phenomenon, i.e., participants resorting to list retrieval in a subset of trials, could have occurred in our study, entailing a number of noteworthy consequences.

Inconsistent individual use of association-based comparison and list retrieval could explain the unexpectedly weak correlations between the number and letter distance effect slopes, as hitherto unconsidered mechanisms might be at work in "list retrieval" trials. To be more concrete, order judgments and, thereby, effects that usually occur in ordering tasks, could play into the comparisons. Ordering tasks are generally characterized by reverse distance effects (especially in trials in which the stimuli are correctly ordered), which is assumed to be the result of sequence familiarity, e.g., “3 – 4 – 5” being more familiar and more frequently encountered than “3 – 7 – 9” (Devlin et al., 2022; Turconi et al., 2006). Accordingly, one can conjecture that such reverse distance effects were present in some of our (letter) trials and might have interfered with the otherwise standard distance effect (cf. Franklin et al., 2009; Morsanyi et al., 2016). On the one hand, this explanation is not entirely consistent with the highly significant standard distance effects we observed in all three domains. On the other hand, especially when only two (rather than three or more) stimuli are presented simultaneously, the distance effect is not necessarily reversed even in ordering tasks (Hamilton & Sanford, 1978; Lyons et al., 2016). Hence, it is conceivable that participants’ reliance on ordinality in some trials affected the correlations without being reflected in the distance effects themselves.

Another potential ramification of participants’ stronger use of list retrieval in the letter (and month) trials concerns the gaps we introduced into the stimulus space. As outlined in the methods section, by exclusively using letters of the English alphabet and omitting Hungarian letters such as “Gy”, the stimulus space might have seemed incomplete (“gapped”) to Hungarian speakers (who, after having received instructions in Hungarian, might have expected to see these Hungarian letters too). This decision was based on the abovementioned study by Krajcsi and Kojouharova (2017), in which the distance effect in a comparison task with newly learned number symbols did not reflect the numerical distances between the (Arabic) numbers these new symbols corresponded to but the distances between the symbols presented in the task (disregarding any gaps resulting from omissions). Therefore, our decision was also contingent on the assumption that, in our study, participants would employ the same strategy as in Krajcsi and Kojouharova (2017) also for the letter comparison. If so, the gaps should have indeed not influenced participants’ performance. For instance, participants should have treated the comparison between “G” and “H” as having a distance of 1, despite the actual distance in the Hungarian alphabet being 2 (due to the presence of the letter “Gy” between “G” and “H”). If, on the other hand, participants (occasionally) deployed a retrieval-based strategy by reciting the Hungarian alphabet, these gaps could make a crucial difference and could distort the distances. Consequently, if the strategy was used inconsistently across trials, participants might have treated some trials we had devised as distance-two trials as distance-three trials and so forth. This, in turn, would make the group of letter trials of a certain distance inhomogeneous and, thereby, interfere with the correlation between the number and letter distance effects.

Finally, a lower degree of familiarity and a higher reliance on list retrieval in one domain could distort the temporal patterns of participants' performance across trials. For instance, it is plausible to assume that participants first needed to familiarize with the specific stimulus space (subset of letters form the middle of a gapped alphabet) they were confronted with before transitioning to more automatic responses. Indeed, some participants reported having switched strategies (e.g., from list retrieval to more automatic comparison processes) after the first few letter trials.

Such a temporal pattern and an initial lack of familiarity also suggest an interesting explanation for the disagreement between our results and those of Fias et al. (2007) who found highly similar neural activation patterns between a number and a letter comparison task. That is, Fias et al. trained their participants on both comparison tasks, number and letters, until they reached a learning criterion (80% correct) before testing commenced. Undoubtedly, this resulted in a higher degree of familiarity with the task in their participants than in ours. One could hence surmise that a controlled training procedure prior to testing would strengthen the correlation between number and letter distance effects.

The individual diversity of strategies applied could even feed back into the temporal patterns of participants' performance. The use of backup strategies, in our case reciting the alphabet, could affect the strength of the associations between the letters; and stronger associations could again favor the primary strategy (e.g., seemingly automatic, association-based comparison) with time (cf. Siegler, 1988). Such a domain-specific learning effect, in turn, could have further decreased the correlations between domains, in particular between numbers and letters. However, this concern can be dispelled in virtue of the disattenuation procedure we applied to the raw correlations. If learning effects should indeed have led to inconsistencies across trails within a domain, this should have been captured in the reliabilities of the variables and, thereby, been taken into account when adjusting the correlations.

Even if the intra-individual variation in strategy use across the letter trials is what made the correlation between the number and letter distance effect slopes almost undetectable, we still have not answered the question posed earlier: How is it possible that the month distance effect slopes correlate with both those of letters and those of numbers when the correlation between the number and the letter distance effect slopes is almost undetectable? After all, many of the arguments elaborated in the previous paragraphs similarly apply to months. For instance, like letters, pairs of months are not frequently compared in every-day life. In addition, some participants stated having employed a retrieval based strategy also for months (at least in some trials), potentially evoking a partly reverse distance effect. And, finally, we also omitted some months, resulting in gaps comparable to those in the letter stimulus space. In short, the use of different strategies across letter trials cannot *per se* explicate why the letter distance effect slope correlated with the month distance effect slope but not (or only weakly) with the number distance effect slope.

What could, however, account for these seemingly contradictory findings is the varying extent to which the domain-specific mix of mechanisms (strategies) *overlaps* between the domains. It is conceivable that the month comparison task shares more of these underlying mechanisms with both the letter and the number comparison (i.e., the months take an "intermediate" position between letters and numbers) than letters share with numbers. For instance, some month trials might be similar to number trials if participants indeed transcoded months to numbers and, on the other hand, other month trials might have resembled some letter trials in virtue of participants' list retrieval strategy.

Indeed, list retrieval and stronger reliance on ordinality could even be what *caused* the distance effect in the letter and, potentially, the month domain. Similar to what Hamilton and Sanford (1978) proposed, list retrieval could have occurred more frequently in short-distance trials, thereby lengthening the reaction time and, speculatively, increasing the error rate in these trials. In consequence, this could have brought about a distance effect via an alternative, unexpected mechanism.

Further, it is conceivable that even in the case of numbers*,* comparison is performed employing several distinct mechanisms. The two crucial properties of symbolic numbers are cardinality (quantity of a set) and ordinality (position within a set). Cardinality and ordinality rely on distinct cognitive mechanisms and follow differential developmental trajectories (Lyons et al., 2016), but could still jointly influence the distance effect. For example, Marshuetz and Smith (2006) argue that what they call magnitude codes could even play into the processing of order, in addition to inter-item associations. Even though inexact representations of quantities have been shown to not underlie the symbolic number comparison distance effect (Krajcsi et al., 2016, 2018), it is possible that the associations within the DSS comprise more than just ordinality. Potentially, a crucial difference in the way numbers, letters and months are processed and compared is already evident in our formulation of the research question and of the instructions in the task. That is, there could be a difference between selecting the smaller/larger element and selecting the earlier/later element that goes beyond linguistic convention. As a result of numbers' double nature, imperfect correlations between numeric and non-numeric symbolic distance effect slopes could arise due to only partially overlapping mechanisms and/or a varying predominance of cardinality or ordinality across trials.

In a similar vein, strategies common between letter and month comparisons might hinge on the fact that both are less abstract than numbers and, hence, carry a wealth of semantic associations. In other words, the associations posited by the DSS might be the most salient feature of numbers (in the task) but only one of several relevant features of letters and months (for a related discussion see Morsanyi et al., 2016). Assuming that the only information conveyed by a number symbol is the representation of a quantity (cardinality), on the one hand, and its ordinal position, on the other hand, there is not much else participants could utilize to facilitate the differentiation and comparison of two number symbols. The idea becomes more tangible when turning to letters and especially months, whose information content is considerably richer. For example, we associate months with festivities (e.g., "December is the month of Christmas" or "It's my grandmother's birthday in April"), deadlines (e.g., "September is when the semester ends"), other months starting with the same letter (e.g., "July is one of 'January – June – July' which I have to pay special attention to"; similar to what one participant described) or, last but not least, seasons. Indeed, almost half of the participants spontaneously reported the use of a season-facilitated strategy in at least some trials. This means that identifying the later one of two months of different seasons (e.g., October and July) was possible via a "fast track" and only if the two months fell within the same season (e.g., August and July) did participants need to resort to comparing them singly. Interestingly, some participants mentioned a similar strategy for the letter domain, meaning that they subdivided the stimulus range into smaller subgroups. This is in line with what Hamilton and Sanford (1978) described; namely, that their participants started to recite the alphabet at certain "anchoring points" and not necessarily from the start. Such a clustering strategy could hence constitute another feature shared between letters and months. Clustering could introduce an additional processing step into the letter and month comparison. Similar to multi-digit number comparison (Poltrock & Schwartz, 1984; Verguts & De Moor, 2005), comparing two letters or months might hence consist of two sequential processing steps, presumably affecting both the reaction time and the distance effect. In support of this, the reverse distance effects in ordering tasks have been shown to be sensitive to whether or not the presented stimuli cross a decade (in the case of two-digit numbers) or year (in the case of months) boundary (Anonymous, unpublished; Franklin et al., 2009). In conjunction with the transcoding of some months to numbers, the lower abstractness of letters and months as well as the clustering strategy could hence account for the correlations – and lack thereof – we found.

In summary, one core assumption that motivated the present study might not have been entirely justified, namely that the presence of distance effects in all three domains is indicative of a shared underlying mechanism. Consequently, it might not have been entirely justified to enquire whether this shared mechanism consists in a domain-general DSS. Even though hypothesizing that different ordered symbolic lists are processed similarly is more defendable than the long-held misconception that the symbolic and non-symbolic number distance effects share a common cause (cf. Krajcsi et al., 2018), and even though we still firmly believe that the DSS plays an essential role at least in number comparison, complementary strategies (and the intra-individual variation of these) ought not to be underestimated. That is, while the letter and month comparisons might still rely on differential associations with “small”/”early” and “large”/”late”, as number comparison does, additional, domain-specific factors and strategies might be at play in the letter and month domains. Especially for the letters, these factors could have caused the predictions we derived from the DSS to be inaccurate or at least imprecise. To henceforth avoid these shortcomings, we can formulate multiple recommendations for future studies based on the findings of the present study.

4.3 Future avenues

First, future studies should take into account potential learning effects. As expounded above, participants could not only gradually switch strategies or become more familiar with the stimulus list across trials, but the alternative strategies themselves (e.g., reciting the alphabet) could influence the processing of subsequent trials. Hence, it would be desirable to systematically control for these effects. This could be achieved, for example, by subjecting one group of participants to a training procedure until they reach a learning criterion (cf. Fias et al., 2007) and comparing their performance in the test to that of a control group that has not undergone task-specific training.

Second, future studies should measure the influence of other intra-individually consistent parameters that could cause correlations similar to those predicted by the DSS. The reason is that, even if perfect correlations between the distance effect slopes of all three domains were found, these need not automatically be the result of a common mechanism underlying the three comparison tasks. Instead, the correlations could be driven by parameters that affect participants’ performance in all three tasks. That is, individual attributes, such as a participant’s working memory (Lyons & Beilock, 2009), general intelligence or arithmetic skills (Morsanyi et al., 2016), could determine the slope of the distance effect in all cases. Therefore, future studies could aim to disentangle a shared underlying mechanism, like the DSS, from these other confounds.

Finally, comparison tasks should be performed on non-numeric ordered lists stripped of any semantic information (such as season). As such a list probably does not exist, an artificial list of newly invented number symbols could be taught to participants (similar to Krajcsi & Kojouharova, 2017). And, to assess whether familiarity with (comparing) the stimuli and reliance on list retrieval is indeed what caused some of our negative results, one of these lists should be taught by reciting the list (e.g., “?” – “@” – “$”), i.e., the way children usually learn the alphabet and, initially, the counting list (cf. Merkley et al., 2016; Zhao et al., 2012). At the same time, a second list of symbols would not be presented as a list, but by reinforcing the one-to-one correspondences between a new symbol and an Arabic number (e.g., “@” = 2), as it was done in Krajcsi and Kojouharova (2017). Subsequently, comparison tasks could be conducted with both lists to be able to correlate the distance effect slopes intra-individually.

4.4 Summary and conclusions

The present study was the first to not only investigate the *presence of,* but also the *correlations between* the slopes of, numeric and non-numeric symbolic distance effects. We enquired whether the Discrete Semantic System, which has previously been suggested to underpin symbolic number comparison and to cause the respective distance effect, is also responsible for the distance effects seen when letters of the alphabet and months of the year are compared. We found moderate to strong correlations between participants’ individual distance effect slopes for numbers and months, and for letters and months, respectively, but only a weak correlation in the case of numbers and letters. This suggests that our assumption, namely that similar distance effects in number, letter and month comparison tasks indicate a common mechanism, and our resulting question of whether this mechanism could be grounded in the DSS, could have been flawed. In contrast, at least two different mechanisms might be at work, the combination of which might overlap more for some pairs of domains (e.g., numbers and months as well as letters and months) than for others (e.g., numbers and letters). For example, reciting the alphabet more frequently in short-distance trials and grouping months into seasons to facilitate far-distance comparisons could cause the distance effect in these domains. The relative role of such other mechanisms that might be operating alongside, or even within, the DSS remains to be explored to be able to determine the domain-generality of the DSS and to assess the uniqueness of number processing.

References

Anonymous. (unpublished). *Is it All About Numbers? Domain-general and Domain-Specific Processing in Ordering and Comparison Tasks*.

Ashcraft, M. (1987). Children's Knowledge of Simple Arithmetic: A Developmental Model and Simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal Methods in Developmental Psychology* (pp. 302–338). Springer. https://doi.org/10.1007/978-1-4612-4694-7_9

Attout, L., Fias, W., Salmon, E., & Majerus, S. (2014). Common Neural Substrates for Ordinal Representation in Short-Term Memory, Numerical and Alphabetical Cognition. *PLOS ONE*, *9*(3), e92049. https://doi.org/10.1371/journal.pone.0092049

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*(39), 14116–14121. https://doi.org/10.1073/pnas.0505512102

Boysen, S., & Berntson, G. (1989). Numerical Competence in a Chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, *103*, 23–31. https://doi.org/10.1037/0735-7036.103.1.23

Brysbaert, M., Fias, W., & Noël, M.-P. (1998). The Whorfian hypothesis and numerical cognition: Is 'twenty-four' processed in the same way as 'four-and-twenty'? *Cognition*, *66*(1), 51–77. https://doi.org/10.1016/S0010-0277(98)00006-7

Cantlon, J. F., & Brannon, E. M. (2006). Shared System for Ordering Small and Large Numbers in Monkeys and Humans. *Psychological Science*, *17*(5), 401–406. https://doi.org/10.1111/j.1467-9280.2006.01719.x

Cantlon, J. F., & Brannon, E. M. (2007). Basic Math in Monkeys and College Students. *PLOS Biology*, *5*(12), e328. https://doi.org/10.1371/journal.pbio.0050328

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

Chiandetti, C., & Vallortigara, G. (2018). Chicken – Cognition in the Poultry Yard. In F. Amici & N. Bueno-Guerra (Eds.), *Field and Laboratory Methods in Animal Cognition: A Comparative Guide* (pp. 97–118). Cambridge University Press. https://doi.org/10.1017/9781108333191.007

De Smedt, B., Noel, 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), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neurosciences*, *21*(8), 355. https://doi-org.uaccess.univie.ac.at/10.1016/S0166-2236(98)01263-6

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*(1), 64–81. https://doi.org/10.1348/2044-8279.002002

Devlin, D., Moeller, K., Reynvoet, B., & Sella, F. (2022). A critical review of number order judgements and arithmetic: What do order verification tasks actually measure? *Cognitive Development*, *64*, 101262. https://doi.org/10.1016/j.cogdev.2022.101262

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

Dos Santos, C. F. (2022). Re-establishing the distinction between numerosity, numerousness, and number in numerical cognition. *Philosophical Psychology*, *35*(8), 1152–1180. https://doi.org/10.1080/09515089.2022.2029387

Embry, I., Hoos, A., & Diehl, T. H. (2022). *ie2misc: Irucka Embry's Miscellanous USGS Functions* (R package version 0.9.0). https://CRAN.R-project.org/package=ie2misc

Fairbank, B. A. (1969). *Experiments on the Temporal Aspects of Number Perception* [Dissertation, The University of Arizona]. https://www.proquest.com/docview/302381873/citation/DF0293D4A1834092PQ/1

Fechner, G. T. (1860). *Elemente der Psychophysik*. Breitkopf u. Härtel. http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10255088-3

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Fias, W., Lammertyn, J., Caessens, B., & Orban, G. A. (2007). Processing of Abstract Ordinal Knowledge in the Horizontal Segment of the Intraparietal Sulcus. *The Journal of Neuroscience*, *27*(33), 8952–8956. https://doi.org/10.1523/JNEUROSCI.2076-07.2007

Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cognitive Processing*, *13*(1), 161–164. https://doi.org/10.1007/s10339-012-0477-5

Fletcher, T. D. (2022). *psychometric: Applied Psychometric Theory* (R package Version 2.3). https://CRAN.R-project.org/package=psychometric

Foltz, G. S., Poltrock, S. E., & Potts, G. R. (1984). Mental comparison of size and magnitude: Size congruity effects. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *10*(3), 442–453. https://doi.org/10.1037//0278-7393.10.3.442

Franklin, M., Jonides, J., & Smith, E. (2009). Processing of order information for numbers and months. *Memory & Cognition*, *37*, 644–654. https://doi.org/10.3758/MC.37.5.644

Fulbright, R., Manson, S., Skudlarski, P., Lacadie, C., & Gore, J. (2003). Quantity Determination and the Distance Effect with Letters, Numbers, and Shapes: A Functional MR Imaging Study of Number Processing. *AJNR. American Journal of Neuroradiology*, *24*, 193–200.

Gallistel, C. R. (1993). A conceptual framework for the study of numerical estimation and arithmetic reasoning in animals. In *The development of numerical competence: Animal and human models* (pp. 211–223). Lawrence Erlbaum Associates, Inc.

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

Gilmore, C. (2015). Approximate Arithmetic Abilities in Childhood. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (pp. 310–330). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.013.006

Goffin, C., Vogel, S. E., Slipenkyj, M., & Ansari, D. (2020). A comes before B, like 1 comes before 2. Is the parietal cortex sensitive to ordinal relationships in both numbers and letters? An fMRI-adaptation study. *Human Brain Mapping*, *41*(6), 1591–1610. https://doi.org/10.1002/hbm.24897

Gordon, P. (2004). Numerical Cognition Without Words: Evidence from Amazonia. *Science*, *306*(5695), 496–499. https://doi.org/10.1126/science.1094492

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

Hamilton, J. M. E., & Sanford, A. J. (1978). The Symbolic Distance Effect for Alphabetic Order Judgements: A Subjective Report and Reaction Time Analysis. *Quarterly Journal of Experimental Psychology*, *30*(1), 33–41. https://doi.org/10.1080/14640747808400652

Hogue, M. E., Beaugrand, J. P., & Laguë, P. C. (1996). Coherent use of information by hens observing their former dominant defeating or being defeated by a stranger. *Behavioural Processes*, *38*(3), 241–252. https://doi.org/10.1016/s0376-6357(96)00035-6

Holloway, I. D., & Ansari, D. (2015). Numerical symbols: An overview of their cognitive and neural underpinnings. In *The Oxford handbook of numerical cognition* (pp. 531–551). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.001.0001

Kadosh, R. C., & Dowker, A. (Eds.). (2015). *The Oxford handbook of numerical cognition* (1. ed.). Univ. Press.

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*, 1882–1905. https://doi.org/10.1037/0096-1523.25.6.1882

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction*, *25*, 95–103. https://doi.org/10.1016/j.learninstruc.2012.12.001

Krajcsi, A. (2017). Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychonomic Bulletin & Review*, *24*(3), 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., Lengyel, G., & Kojouharova, P. (2016). The Source of the Symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *7*, 1795. https://doi.org/10.3389/fpsyg.2016.01795

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic Number Comparison Is Not Processed by the Analog Number System: Different Symbolic and Non-symbolic Numerical Distance and Size Effects. *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00124

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

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., & Beilock, S. L. (2009). Beyond quantity: Individual differences in working memory and the ordinal understanding of numerical symbols. *Cognition*, *113*(2), 189–204. https://doi.org/10.1016/j.cognition.2009.08.003

Lyons, I. M., Vogel, S. E., & Ansari, D. (2016). On the ordinality of numbers: A review of neural and behavioral studies. *Progress in Brain Research*, *227*, 187–221. https://doi.org/10.1016/bs.pbr.2016.04.010

Maloney, E. A., Risko, E. F., Preston, F., Ansari, D., & Fugelsang, J. (2010). Challenging the reliability and validity of cognitive measures: The case of the numerical distance effect. *Acta Psychologica*, *134*(2), 154–161. https://doi.org/10.1016/j.actpsy.2010.01.006

Manova, S., & Knell, G. (2021). *Chapter 17. Two-suffix combinations in native and non-native English*. John Benjamins Publishing Company. https://benjamins.com/catalog/cilt.353.17man

Marino, L. (2017). Thinking chickens: A review of cognition, emotion, and behavior in the domestic chicken. *Animal Cognition*, *20*(2), 127–147. https://doi.org/10.1007/s10071-016-1064-4

Marshuetz, C., & Smith, E. E. (2006). Working memory for order information: Multiple cognitive and neural mechanisms. *Neuroscience*, *139*(1), 195–200. https://doi.org/10.1016/j.neuroscience.2005.08.024

Merkley, R., Shimi, A., & Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations: The role of magnitude and ordinal information. *ZDM*, *48*(3), 279–289. https://doi.org/10.1007/s11858-015-0751-y

Morsanyi, K., O'Mahony, E., & Mccormack, T. (2016). Number comparison and number ordering as predictors of arithmetic performance in adults: Exploring the link between the two skills, and investigating the question of domain-specificity. *The*

Quarterly Journal of Experimental Psychology, *70*, 1–54. https://doi.org/10.1080/17470218.2016.1246577

Moyer, R. S. (1973). Comparing objects in memory: Evidence suggesting an internal psychophysics. *Perception & Psychophysics*, *13*(2), 180–184. https://doi.org/10.3758/BF03214124

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

Nieder, A. (2020). The Adaptive Value of Numerical Competence. *Trends in Ecology & Evolution*, *35*(7), 605–617. https://doi.org/10.1016/j.tree.2020.02.009

Parkman, J. M. (1971). Temporal aspects of digit and letter inequality judgments. *Journal of Experimental Psychology*, *91*(2), 191–205. https://doi.org/10.1037/h0031854

Piantadosi, S. T., & Cantlon, J. F. (2017). True Numerical Cognition in the Wild. *Psychological Science*, *28*(4), 462–469. https://doi.org/10.1177/0956797616686862

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and Approximate Arithmetic in an Amazonian Indigene Group. *Science*, *306*, 499–503. https://doi.org/10.1126/science.1102085

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

Pronk, T., Molenaar, D., Wiers, R. W., & Murre, J. (2022). Methods to split cognitive task data for estimating split-half reliability: A comprehensive review and systematic assessment. *Psychonomic Bulletin & Review*, *29*(1), 44–54. https://doi.org/10.3758/s13423-021-01948-3

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Regolin, L., Rugani, R., Pagni, P., & Vallortigara, G. (2005). Delayed search for social and nonsocial goals by young domestic chicks, *Gallus gallus domesticus*. *Animal Behaviour*, *70*(4), 855–864. https://doi.org/10.1016/j.anbehav.2005.01.014

Regolin, L., & Vallortigara, G. (1995). Perception of partly occluded objects by young chicks. *Perception & Psychophysics*, *57*(7), 971–976. https://doi.org/10.3758/BF03205456

Rugani, R., Fontanari, L., Simoni, E., Regolin, L., & Vallortigara, G. (2009). Arithmetic in newborn chicks. *Proceedings of the Royal Society B: Biological Sciences, 276*(1666), 2451–2460. https://doi.org/10.1098/rspb.2009.0044

Rugani, R., & Hevia, M.-D. de. (2016). Number-space associations without language: Evidence from preverbal human infants and non-human animal species. *Psychonomic Bulletin & Review*, *24*(2), 352. https://doi.org/10.3758/s13423-016-1126-2

Rugani, R., Kelly, D. M., Szelest, I., Regolin, L., & Vallortigara, G. (2010b). Is it only humans that count from left to right? *Biology Letters*, *6*(3), 290–292. https://doi.org/10.1098/rsbl.2009.0960

Rugani, R., Regolin, L., & Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *Journal of Experimental Psychology. Animal Behavior Processes*, *34*, 388–399. https://doi.org/10.1037/0097-7403.34.3.388

Rugani, R., Regolin, L., & Vallortigara, G. (2010a). Imprinted numbers: Newborn chicks' sensitivity to number vs. continuous extent of objects they have been reared with. *Developmental Science*, *13*, 790–797. https://doi.org/10.1111/j.1467-7687.2009.00936.x

Rugani, R., & Rosa-Salva, O. (2021). Spatial–Numerical Association in Nonhuman Animals. In A. B. Kaufman, J. C. Kaufman, & J. Call (Eds.), *The Cambridge Handbook of Animal Cognition* (pp. 602–620). Cambridge University Press. https://doi.org/10.1017/9781108564113.032

Rumbaugh, D. M., Savage-Rumbaugh, S., & Hegel, M. T. (1987). Summation in the chimpanzee (*Pan troglodytes*). *Journal of Experimental Psychology: Animal Behavior Processes*, *13*, 107–115. https://doi.org/10.1037/0097-7403.13.2.107

Sasanguie, D., & Vos, H. (2018). About why there is a shift from cardinal to ordinal processing in the association with arithmetic between first and second grade. *Developmental Science*, *21*(5), e12653. https://doi.org/10.1111/desc.12653

Shapiro, L. (2007). The Embodied Cognition Research Programme. *Philosophy Compass*, *2*(2), 338–346. https://doi.org/10.1111/j.1747-9991.2007.00064.x

Siegler, R. S. (1988). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology. General*, *117*(3), 258–275. https://doi.org/10.1037//0096-3445.117.3.258

Signorell, A. (2023). *DescTools: Tools for Descriptive Statistics* (R package version 0.99.48). https://CRAN.R-project.org/package=DescTools

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

Taft, M. (2015). The Nature of Lexical Representation in Visual Word Recognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook of Reading*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199324576.013.6

Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science*, *14*, 66–73. https://doi.org/10.1111/1467-9280.01420

Toll, S. W. M., Van Viersen, S., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). The development of (non-)symbolic comparison skills throughout kindergarten and their relations with basic mathematical skills. *Learning and Individual Differences*, *38*, 10–17. https://doi.org/10.1016/j.lindif.2014.12.006

Tudusciuc, O., & Nieder, A. (2010). Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans. *Experimental Brain Research*, *207*(3), 221–231. https://doi.org/10.1007/s00221-010-2452-7

Turconi, E., Campbell, J. I. D., & Seron, X. (2006). Numerical order and quantity processing in number comparison. *Cognition*, *98*(3), 273–285. https://doi.org/10.1016/j.cognition.2004.12.002

Uittenhove, K., & Lemaire, P. (2015). Numerical cognition during cognitive aging. In *The Oxford handbook of numerical cognition* (pp. 345–364). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.001.0001

Vallortigara, G., Regolin, L., Rigoni, M., & Zanforlin, M. (1998). Delayed search for a concealed imprinted object in the domestic chick. *Animal Cognition*, *1*(1), 17–24. https://doi.org/10.1007/s100710050003

Van Opstal, F., Gevers, W., De Moor, W., & 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

Verguts, T., & De Moor, W. (2005). Two-digit Comparison. *Experimental Psychology*, *52*(3), 195–200. https://doi.org/10.1027/1618-3169.52.3.195

Vos, H., Sasanguie, D., Gevers, W., & Reynvoet, B. (2017). The role of general and number-specific order processing in adults' arithmetic performance. *Journal of Cognitive Psychology*. https://doi.org/10.1080/20445911.2017.1282490

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

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed. 2016 Edition). Springer.

Wickham, H., Francois, R., Henry, H., Müller, K., & Vaugh, D. (2023). *dplyr: A Grammar of Data Manipulation* (R package version 1.1.0). https://CRAN.R-project.org/package=dplyr

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. (1998). An evolved capacity for number. In *The evolution of mind* (pp. 107–126). Oxford University Press.

Wynn, K., & Bloom, P. (1992). The origins of psychological axioms of arithmetic and geometry. *Mind & Language*, *7*, 409–416. https://doi.org/10.1111/j.1468-0017.1992.tb00313.x

Zhao, H., Chen, C., Zhang, H., Zhou, X., Mei, L., Chen, C., Chen, L., Cao, Z., & Dong, Q. (2012). Is Order the Defining Feature of Magnitude Representation? An ERP Study on Learning Numerical Magnitude and Spatial Order of Artificial Symbols. *PLoS ONE*, *7*(11), e49565. https://doi.org/10.1371/journal.pone.0049565

Zorzi, M., Di Bono, M. G., & Fias, W. (2011). Distinct representations of numerical and non-numerical order in the human intraparietal sulcus revealed by multivariate pattern recognition. *NeuroImage, 56*, 674–680. https://doi.org/10.1016/j.neuroimage.2010.06.035