



OPEN

# Professional mathematicians do not differ from others in the symbolic numerical distance and size effects

Mateusz Hohol<sup>1</sup>✉, Klaus Willmes<sup>2</sup>, Edward Nęcka<sup>3</sup>, Bartosz Brożek<sup>1</sup>, Hans-Christoph Nuerk<sup>4,5</sup> & Krzysztof Cipora<sup>6</sup>✉

The numerical distance effect (it is easier to compare numbers that are further apart) and size effect (for a constant distance, it is easier to compare smaller numbers) characterize symbolic number processing. However, evidence for a relationship between these two basic phenomena and more complex mathematical skills is mixed. Previously this relationship has only been studied in participants with normal or poor mathematical skills, not in mathematicians. Furthermore, the prevalence of these effects at the individual level is not known. Here we compared professional mathematicians, engineers, social scientists, and a reference group using the symbolic magnitude classification task with single-digit Arabic numbers. The groups did not differ with respect to symbolic numerical distance and size effects in either frequentist or Bayesian analyses. Moreover, we looked at their prevalence at the individual level using the bootstrapping method: while a reliable numerical distance effect was present in almost all participants, the prevalence of a reliable numerical size effect was much lower. Again, prevalence did not differ between groups. In summary, the phenomena were neither more pronounced nor more prevalent in mathematicians, suggesting that extremely high mathematical skills neither rely on nor have special consequences for analogue processing of symbolic numerical magnitudes.

Numerical knowledge is encoded in multiple formats serving specific functions<sup>1–3</sup>. The first kind of code contains the analogue representation of number magnitude; the second one encompasses the visual form of numbers; and the third one stores linguistic representations of numbers. Regarding the first code, namely, analogue magnitude, there is a large body of evidence for shared behavioural characteristics of comparative judgements on symbolic numbers, e.g., Arabic<sup>4–6</sup>, non-symbolic numerals, e.g., sets of dots<sup>7,8</sup>, and other continua including line length<sup>9</sup>, angle<sup>10</sup>, physical object size<sup>11,12</sup>, luminance<sup>13,14</sup>, and non-directly perceivable properties like intelligence<sup>15,16</sup>. Walsh<sup>17</sup> proposed the “theory of magnitude” (ATOM) for the processing of these and other continua, which can be thought of in terms of classification of “more or less than.”

**Characteristics of analogue numerical magnitude processing.** Analogue magnitude comparisons have been studied in different human cultures, languages and notations<sup>7,18,19</sup>, as well as age groups<sup>20,21</sup>. Moreover, there is an extensive knowledge base regarding magnitude comparisons in various non-human animal species<sup>22–24</sup>, from insects<sup>25</sup>, through fish<sup>26</sup>, amphibians<sup>27</sup>, and birds<sup>28</sup>, up to monkeys<sup>29</sup> and apes<sup>30</sup>. Taken together, these studies suggest presence of the analogue numerical magnitude representation among human beings and its deep evolutionary origins<sup>31</sup>. On the other hand, studying numerical magnitude comparisons in animals is basically limited to non-symbolic material. Analogue magnitude comparisons, both involving symbolic and non-symbolic numerical instances, are often assumed to be driven by Weber’s law and performed by a single

<sup>1</sup>Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Krakow, Poland. <sup>2</sup>Clinic of Neurology, University Hospital, RWTH Aachen University, Aachen, Germany. <sup>3</sup>Institute of Psychology, Jagiellonian University, Krakow, Poland. <sup>4</sup>Department of Psychology, University of Tübingen, Tübingen, Germany. <sup>5</sup>LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany. <sup>6</sup>Centre for Mathematical Cognition, Loughborough University, Loughborough, UK. ✉email: mateusz.hohol@uj.edu.pl; k.cipora@lboro.ac.uk

cognitive system called the Approximate Number System (ANS)<sup>19,32</sup>. However, other approaches have also been proposed<sup>33,34</sup>. One of them states that non-symbolic comparisons are handled by the ANS, while symbolic ones are served by the Discrete Semantic System (DSS) resembling the mental lexicon<sup>34</sup>.

Despite the analogue magnitude representation having been intensely studied in non-symbolic format, it also has a prominent role in symbolic number processing. Several models of symbolic number processing consider it as an important component<sup>1,4,35</sup>. The analogue magnitude representation underlying processing of *symbolic* numbers is the main focus of this study. The numerical distance effect (NDE) is one of the fundamental characteristics of analogue magnitude processing<sup>4,36</sup>. In case of the comparison of two numbers, performance is poorer (i.e., reaction times, henceforth RTs, are longer, and accuracy is lower) for numbers that are closer together (e.g., 4 and 5) than for ones that are farther apart (e.g., 1 and 5). The numerical size effect (NSE) is another manifestation of analogue magnitude processing<sup>37</sup>. For an identical numerical distance, performance is better when numbers are small (e.g., 1 and 5) than when they are large (e.g., 5 and 9). These effects were initially found for symbolic number processing<sup>4</sup> and later for non-symbolic numerosities<sup>7</sup>. Both effects can be well observed in tasks comprising judgments of single-digit numbers. Importantly, such tasks are less loaded with strategy use and domain-general processes, as compared to tasks using multi-digit numbers<sup>38–41</sup>. Here we are interested in the analogue magnitude processing of symbolic numbers. Thus, unless otherwise stated, we use the abbreviation NDE referring to the symbolic NDE. The same applies to the NSE.

The characteristics of basic analogue processing of symbolic numbers have been traditionally elucidated by Restle's<sup>5</sup> metaphor of the mental number line, where representations of numbers are organized as points in a spatial structure with larger distances between small numbers and smaller distances between large numbers (logarithmic compression)<sup>19,32</sup> or a more diffuse representation of large numbers. On the other hand, the NDE and the NSE can be accounted for without any recourse to spatial mapping<sup>42</sup>. Magnitude can be activated independently from its spatial association<sup>43</sup>.

Some authors argue that phenomena quantified by the NDE and the NSE can also be accounted for by the numerical ratio effect. As the ratio effect depends on both the numerical distance between numbers to be compared and their absolute magnitudes, it considers both the NDE and the NSE. Performance decreases (longer RTs, higher error rates) with increasing numerical ratio. However, considering the ratio does not allow to investigate the NDE and the NSE separately. Research on the numerical ratio effect gained more popularity in studies investigating non-symbolic number processing<sup>44</sup>. Nevertheless, some studies used it to account for symbolic number processing effects<sup>45</sup> as well. Therefore, in addition to calculating the NDE and the NSE, we also studied ratio effects (see Supplementary Material 1 for more detail).

**Basic numerical processing and mathematical skills.** It is often argued theoretically that basic numerical processing serves as a scaffold for the acquisition of arithmetic concepts and full-fledged mathematical competences in general<sup>19,32,46,47</sup>. Investigating how characteristics of basic numerical processing actually relate to more advanced mathematical skills is one of the crucial aims of the differential psychology of mathematical cognition<sup>48,49</sup>. The idea is that more complex mathematical cognition cannot develop properly if the basic representations on which mathematics is based, are not properly built. However, evidence for such relationships between basic numerical representations and complex mathematical skills is mixed and differs largely depending on the signature of basic number processing under scrutiny. The performance on several tasks tackling some representations correlates positively with mathematical achievement, but other tasks reveal no correlation or somewhat mixed evidence<sup>50–52</sup>. Moreover, even different variants of the tasks considered to engage the same cognitive system could correlate with mathematical skills in different ways, depending on the stimuli. In particular, the association between numerical magnitude processing and mathematical achievement is stronger in magnitude tasks involving symbolic numbers in comparison to non-symbolic magnitude tasks<sup>53,54</sup>. Therefore, we decided to focus on symbolic number processing. However, this does not preclude incremental variance to be explained by non-symbolic processing (for reviews on relations between non-symbolic number processing and mathematical skills see<sup>48,52,55</sup>).

**Symbolic numerical distance effect and mathematical skills.** The analogue numerical magnitude representation is the most ubiquitous fundamental representation of numbers in the cognitive system. Its hallmark manifestation is the NDE. It is assumed that when the analogue numerical magnitude representation is precise and refined, the NDE should be smaller<sup>52,56</sup>. Conversely, a larger NDE seems to be associated with a more imprecise analogue magnitude representation.

Although the analogue numerical magnitude representation and the NDE constitute a fundamental representation and its manifestation respectively, studies show a somewhat inconsistent picture of relationships between the NDE and mathematical skills. Early studies demonstrated that the NDE decreases with age during childhood and stabilizes around the fourth grade<sup>57</sup>. This suggests that the NDE reflects a numerical representation which changes and gets refined during normal development. It is worth noting that this observation has been questioned because the NDE can be driven by changes in general RT pattern, and it is known that RTs become faster and less variable with age. Additionally, opposite effects can be found if effect size is considered rather than raw RT<sup>45</sup>. Nevertheless, despite its size changes with age, the NDE remains robust in adulthood. The results of several studies on groups with typical mathematical skills levels did not provide clearcut results: usually, the size of the NDE does not explain a substantial amount of variance in mathematical skills<sup>53,58</sup>. On the other hand, mathematical skills correlate at a moderate level with overall RT (i.e., not the NDE) in the magnitude classification task<sup>59</sup>, which may be indicative of easiness of access to numerical magnitude in general.

Studies on participant groups with mathematical difficulties have also provided inconclusive results ranging from a larger NDE<sup>60–62</sup> through no differences in the NDE<sup>63</sup> to a smaller NDE<sup>64</sup> when compared to groups without

mathematical difficulties. There are also case reports of a reverse NDE in dyscalculic individuals<sup>64,65</sup>. The NDE was also observed in a calculation prodigy<sup>66</sup>.

It seems that there is no genuinely strong and consistent relationship between the NDE and mathematical skills level when groups with typical mathematical abilities and mathematical difficulties are taken into account. However, all of these studies compared only participants with poor mathematical skills to a control group. Except for the one report of a calculation prodigy, groups displaying extremely high mathematical skills levels are largely understudied and to the best of our knowledge, the latter groups have not been systematically tested as regards the NDE. Professional mathematicians may differ from other groups in terms of their NDE. For instance, they differ from controls in their SNARC effect<sup>67</sup> (in the parity judgement task with symbolic numbers) and positive number mapping (in the number line task)<sup>68</sup>; thus, differences in basic numerical effects are possible. Such differences could be due to intense training and exposure to numbers leading to a more precise analogue magnitude representation, or because possessing a specific type of magnitude representation fosters the mastery of professional mathematical skills and thus helps one to become a professional mathematician.

**Symbolic numerical size effect and mathematical skills.** Compared to the NDE, our knowledge on the relationship between the size of the NSE and mathematical skills is considerably weaker. As Rousselle and Noël<sup>64</sup> reported, the NSE is reduced, similarly to the NDE, in children with mathematical learning difficulties as compared to children without difficulties. On the other hand, Núñez-Peña and Suárez-Pellicioni<sup>69</sup> found that highly math-anxious individuals revealed a larger NSE (the same is true regarding the NDE) than less math-anxious participants, suggesting that the first group is characterized by less precise access to numerical magnitude. In summary, although some studies report differences in the NSE and the NDE related to mathematical skills, no consistent picture has emerged yet.

**Individual prevalence of symbolic numerical distance and size effects.** As we have already mentioned, the NDE and the NSE are considered to be highly widespread among human cultures and age groups, yet little is known about their frequency of occurrence at the individual level. Recently, a distinction of group level psychological phenomena into dominant and indominant ones has been proposed<sup>70</sup>. Dominant phenomena are present in virtually all individuals and no one reveals a divergent effect. Indominant phenomena are present only in some individuals, and some individuals might also have reversed effects. Assuming the most popular view that NDE and NSE are hallmarks of the same and universal system of magnitude representation (sometimes even put together to constitute the numerical ratio effect) they should both be dominant, i.e., present in virtually all individuals. To the best of our knowledge, their individual prevalence has never been studied.

**Objectives of the present study.** First of all, in the present study, we aim to investigate the relationship between two phenomena characterizing the analogue representation of number magnitude—namely, the NDE and the NSE—and the mathematical skills level operationalized in terms of formal education. To this end, we tested four groups of participants: professional mathematicians, engineers, social scientists, all at the level of advanced doctoral studies in their respective domain, and a reference group using the symbolic magnitude classification task with single-digit Arabic numbers. Secondly, our goal was to investigate the individual prevalence of the NDE and the NSE.

At the group level, we expected to replicate the NDE and the NSE. Taking into account that previous studies provided a somewhat inconsistent picture of relationships between the analogue numerical magnitude representations and mathematical skills, as discussed earlier, it is hard to state directional hypotheses regarding the NDE and the NSE of professional mathematicians. However, because both these phenomena are hallmarks of the analogue representation of numerical magnitude, we expected that, if they relate to mathematical skills, the direction should be the same for both NDE and NSE. In particular, the following scenarios seem possible:

- Professional mathematicians do not differ from other groups in their symbolic NDE and NSE effects. This scenario is supported by the observation that the size of the NDE does not typically account for a considerable amount of variance in mathematical skills.
- Professional mathematicians have weaker symbolic NDE and NSE than other groups, because smaller effects are typically considered to be indicators of a more precise analogue numerical magnitude representation.
- Professional mathematicians have stronger symbolic NDE and NSE compared to other groups. We do not see a strong theoretical justification for this scenario. However, mathematicians generally constitute an understudied group. Their analogue magnitude representation may be more flexible in comparison to others, leading to stronger NDE and NSE.

Regarding individual prevalence, we expect to find that the NDE and the NSE go hand-in-hand as dominant phenomena<sup>70</sup>, since they are both assumed to characterize the same aspect of basic symbolic numerical processing, namely an analogue magnitude representation.

## Method

**Participants.** The magnitude classification task was performed by four groups of participants. There were 100 participants (47 female) in total. Their mean age was 25.2 years (SD = 3.7, range 18–35 years). All participants had normal or corrected-to-normal vision and were native Polish speakers. All participants provided informed consent, and the methods and procedures conformed to recognized ethical guidelines for testing human partici-

pants. The study was approved by the Ethics Committee for Experimental Research at the Institute of Psychology, Jagiellonian University.

Participants constituted the following groups: (1) mathematicians (henceforth M,  $n = 14$ ; 2 females; mean age 28.2)—PhD studies in mathematics; (2) engineers (henceforth E;  $n = 15$ ; 2 females mean age 28.1)—PhD studies in fields other than mathematics but requiring advanced math in everyday professional work (e.g., telecommunication, chemistry); (3) social scientists (henceforth S;  $n = 15$ ; 2 females; mean age 27.5)—PhD studies in social sciences (i.e., psychology, sociology, philosophy, law); (4) a reference group (henceforth R;  $n = 56$ ; 39 females; mean age 23.1)—individuals recruited from the general population. The inclusion criterion for the first three groups was to be at least advanced in doctoral studies (the exact dissertation topic approved by the departmental council). Although the educational background in the R group varied, nobody met the inclusion criteria for M, E, and S, nor was a PhD student nor psychology undergraduate student. The members of the groups M, E and S are the same participants as described in Cipora et al.<sup>67</sup>.

The large male overrepresentation in the M and the E groups reflects the general gender proportion in these cohorts. We reached out to all of the eligible PhD students and young PhDs of the Cracovian universities and tested all of those, who agreed to participate. Subsequently, we matched the gender proportion in the social sciences group (S). As the reference group (R) was initially tested in the context of another study (but using exactly the same task), and was recruited from general population, we tested all participants who signed up for the study. Given the gender proportion imbalance between the R group and M, E and S groups, in the Supplementary Material 2 we present all analyses considering male participants only and between gender comparisons regarding measures of interest to examine whether the observed differences were driven by a male advantage in the M and E groups.

The participants in the first three groups self-reported right-handedness. In the fourth group 52 participants were right-handed and 4 were left-handed. In the M, E, and S groups, the inclusion criterion was based on the writing hand. The handedness of participants in the R group is reported here based on this information as well. However, these participants also answered Oldfield's handedness questionnaire<sup>71</sup>, which allows determination of handedness in a more fine-grained manner. The Oldfield questionnaire score for each participant in the R group is included in the shared data file.

**Materials.** We used a computerized magnitude classification task with symbolic Arabic numbers. The participant task was to decide whether a visually presented single-digit number was smaller or larger than 5 using the Q and P keys on a standard QWERTY keyboard. Both speed and accuracy were stressed in the instruction. All stimuli were presented in black font (size 30) against a light grey background (210 210 210 in RGB notation) to avoid sharp contrasts. The task comprised two blocks with reverse response key assignment. In each block, each number (1, 2, 3, 4, 6, 7, 8, 9) was presented 30 times that gives 240 experimental trials per block. In total, 480 experimental trials of the magnitude classification task were administered to each participant. Trial order was randomized with the restriction that each number could not appear more than two times in a row. Short training sessions preceded blocks. Each training session comprised 16 trials (each number presented twice). Accuracy feedback was presented following incorrect responses, and information about response mapping was present in the bottom line of the screen. The order of blocks was counterbalanced among participants. In experimental blocks, no feedback and information about response key assignment was present. Each trial started with an eye fixation cross presented for 300 ms. Subsequently, the target number was presented until the participant responded or for a maximum duration of 2 s. The next trial started after 500 ms of blank screen presentation. A standard, MS Windows compatible computer running DMDX<sup>72</sup> was used to present stimuli and collect responses. Documented experimental procedures are also shared at the Open Science Framework (<https://doi.org/10.17605/OSF.IO/MSDNR>).

**Procedure.** The task was performed as part of a numerical cognition test battery. First of all, informed consent was obtained from all participants. Subsequently, participants sat in front of the computer and performed computerized tasks. The parity judgment task was administered first (results were reported in<sup>67,73</sup>; the raw data from the parity judgment task can be found at <https://osf.io/tw843/>). Afterwards, participants started the magnitude classification task. The magnitude classification task lasted approximately 12 min. After completion of the magnitude classification task, other tasks followed, differing between the M, E, S, and R groups. These tasks and their results are not reported here.

**Analysis.** Data processing and analysis were conducted in the R language<sup>74</sup>. Both the data and analysis script are available at the Open Science Framework (<https://doi.org/10.17605/OSF.IO/MSDNR>).

To control for the stability of our data, we estimated the reliability of all effects of interest. This was done using a split-half method (Spearman-Brown corrected for double test length). A detailed description of the algorithm can be found in the Supplementary Material to Cipora et al.'s work<sup>75</sup>.

In the analysis, both frequentist and Bayesian approaches were used, so that we can provide evidence supporting existing effects or null effects. The NDE and the NSE were quantified by means of multiple regression analyses on RTs aggregated for each number for each participant separately. RTs were regressed on the numerical distance from the criterion value of 5 and on the numerical magnitude of numbers. Magnitude and distance predictors are orthogonal, so there is no collinearity problem. Slopes corresponding to these predictors were measures of the NDE and the NSE, respectively. A complimentary analysis considering the ratio effect is available in Supplementary Material 1. The bimanual setup with reverse response-to-key assignment allows for the measurement of the SNARC effect as well (this analysis is presented in Supplementary Material 3). In the case of numerical distance, negative slopes correspond to the typical NDE, the more negative they are, the stronger the

NDE is. In the case of the NSE, positive slopes correspond to the regular size effect, and the larger they are, the stronger the effect is. To test whether the effect is present at the sample/group level, slopes were tested against 0 by means of the one-sample  $t$ -test (one-sided: for negative values for the NDE and positive values for the NSE). Both frequentist and Bayesian  $t$ -tests were used. Group comparisons were conducted by means of UNIANOVA and the BF<sub>s</sub> were computed with the `anovaBF` function of the R package `BayesFactor`<sup>76</sup>.

In the following step, we aimed to investigate the presence of the NDE and the NSE at the individual level. Specifically, the regression method does not allow for making inferences about the presence of effects of interest at the individual level. This is possible with a bootstrapping approach<sup>75</sup>. Here we adapted a H0 bootstrapping approach proposed by Cipora et al.<sup>75</sup>. Specifically, we aimed to check whether finding the NDE / NSE as empirically observed in each participant is likely when the null hypothesis holds, i.e. the RT pattern of a given participant does not depend on the numerical magnitude of numbers in a magnitude classification task. Therefore, separately for each participant we randomly sampled (with replacement) 8 sets of 60 trials. Subsequently, these sets were arbitrarily assigned numbers 1, 2, 3, 4, 6, 7, 8, 9 and the corresponding numerical distances from 5. These were used as predictors in a regression analysis similar to the one used to estimate the empirically observed distance and size effects. The bootstrapping procedure was repeated 5,000 times. The slopes from these bootstrap based regressions were considered as possible outcomes of the analysis if there is no NDE and no NSE. Subsequently, we checked whether empirically observed slopes were outside the mid 90% of the distribution of bootstrap slopes (i.e., the 90% H0 confidence intervals). In case of the NDE, if the empirical slope was  $< 0$  and it was outside the 90% CI, a participant was considered to show a reliable NDE. If the slope was positive and it was outside the 90% CI, the participant was considered to show a reliable reverse NDE. If the slope was within the 90% CI, the participant was considered as not showing a reliable NDE. For the NSE, the classification is similar except that positive slopes correspond to the typical effect and negative ones to a reverse NSE.

In the last step, we compared groups with respect to the proportion of participants displaying reliable typical, non-reliable, or reliable reverse effects. (In Supplementary Material 4 we present a complete correlation matrix of all the measures we used in this study).

To check for robustness of our results, we conducted the same analysis for standardized slopes<sup>75</sup>. The results remained unchanged, so we do not report them in detail, however, they are available for inspection along with other shared analyses (<https://doi.org/10.17605/OSF.IO/MSDNR>).

**Data preprocessing.** Data from two participants (one from the M group and one from the E group) were discarded from further analysis during preliminary data screening because of excessive error rates (49.5% and 50%). These participant errors can be attributed to confusion over experimental instructions (i.e., one block comprised mostly correct responses and the other mostly errors). All of the following results do not consider the data from these two individuals. Overall accuracy on the magnitude classification task was 96.9%. The ANOVA on the transformed accuracy data [ $2 \cdot \arcsin(\sqrt{\text{proportion correct}})$ ] revealed significant between-group differences  $F(3, 94) = 3.30$ ,  $p = 0.024$ ,  $\eta^2_p = 0.095$ ,  $\text{BF} = 2.21$ . Post hoc analysis (HSD corrected) revealed that the E group had a significantly higher performance than the R group ( $p = 0.047$ ). However, due to very high overall performance, errors were not further analysed. Subsequently, the RT data were filtered. First, correct responses with RTs  $< 200$  ms (0.05% of all trials) were treated as anticipations and not further analysed. Eventually, a sequential trimming method<sup>77</sup> was applied: for each participant RTs outside  $\pm 3\text{SD}$  from the individual mean were discarded. Ultimately, 91.1% of RT data was considered in the main analysis.

## Results

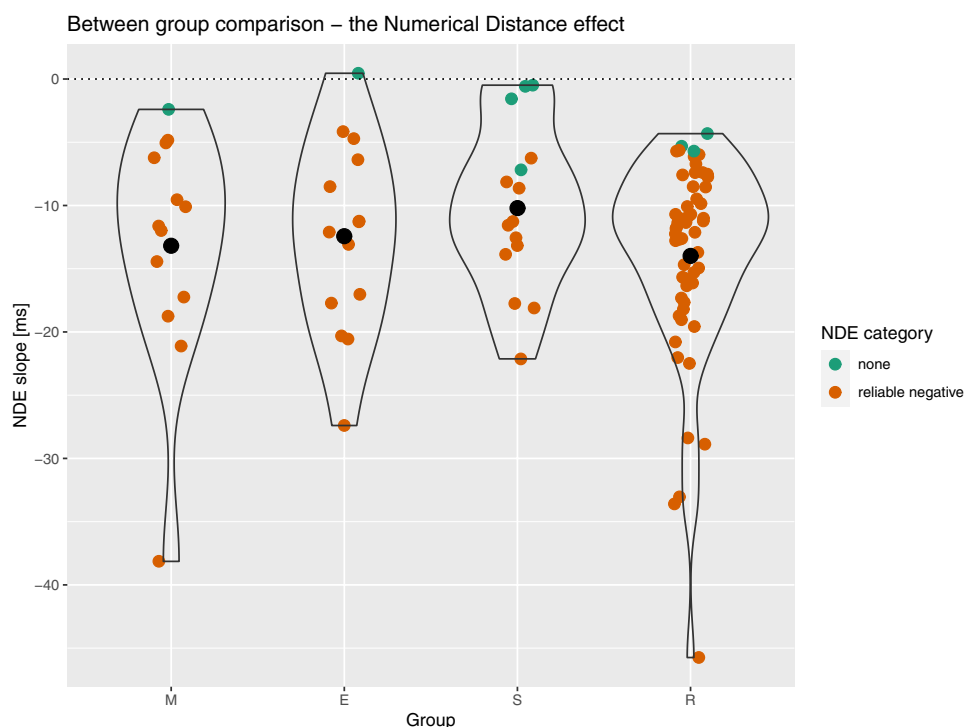
**Overview and reliability.** Overall mean RT was 504 ms ( $\text{SD} = 79$ ). There was no between group difference in mean RT  $F(3, 94) = 1.41$ ,  $p = 0.246$ ,  $\eta^2_p = 0.05$ ,  $\text{BF} = 0.31$ . Overall mean intraindividual variability in RT was 90 ms ( $\text{SD} = 36$ ). The intraindividual variability in RT did not differ between groups,  $F(3, 94) = 0.52$ ,  $p = 0.669$ ,  $\eta^2_p = 0.02$ ,  $\text{BF} = 0.13$ . Reliability (Split-half, Spearman-Brown corrected) of the NDE was 0.84 and reliability (Split-half, Spearman-Brown corrected) of the NSE was 0.70.

**The numerical distance effect.** A summary of the results is presented in the left part of Table 1. The NDE was robust at the whole sample level as well as in each group, separately. This is supported by both frequentist and Bayesian analysis. However, the NDE did not differ between groups  $F(3, 94) = 0.93$ ,  $p = 0.428$ ,  $\eta^2_p = 0.03$ ,  $\text{BF} = 0.20$ . The results of the bootstrapping analysis are summarized in Fig. 1. Distributions of the NDE slopes largely overlap between groups. Virtually all participants except one displayed negative slopes (indicating the typical NDE). A total of 91% of the participants had a reliable NDE. Importantly, none of the participants had a reliable reverse NDE. Proportions of participants revealing reliable slopes did not differ between groups (Fisher exact test,  $p = 0.085$ ).

**The numerical size effect.** A summary of the results is presented in the right part of Table 1. The NSE was robust at the whole sample level as well as in the R and the E group. It did not reach significance in any other group. However, in the M, E, and S groups Bayesian evidence was highly inconclusive. Again, groups did not differ regarding the NSE,  $F(3, 94) = 0.59$ ,  $p = 0.626$ ,  $\eta^2_p = 0.02$ ,  $\text{BF} = 0.13$ . The results of the bootstrapping analysis are summarized in Fig. 2. Interestingly, only 29% of the participants displayed a reliable NSE. On the other hand, 11% of the participants had a reliable reverse NSE. Most importantly, 60% of participants did not display a reliable NSE. The proportions of participants displaying reliable, reliable reverse, and no NSE did not differ between groups (Fisher exact test  $p = 0.629$ ).

Group	Numerical distance							Numerical size						
	Mean (SD)	<i>t</i> -test*	<i>d</i>	BF10	Proportion			Mean (SD)	<i>t</i> -test*	<i>d</i>	BF10	Proportion		
					Reliable (%)	Reliable reversed (%)	No reliable (%)					Reliable (%)	Reliable reversed (%)	Not reliable (%)
Overall	-13.08 (7.90)	<b><i>t</i>(97) = -16.39, <i>p</i> &lt; 0.001**</b>	1.66	> 10 <sup>26</sup>	91	0	9	1.20 (3.16)	<b><i>t</i>(97) = 3.77, <i>p</i> &lt; 0.001</b>	0.38	142.00	29	11	60
M	-13.19 (9.41)	<b><i>t</i>(12) = -5.06, <i>p</i> &lt; 0.001</b>	1.40	234.00	92	0	8	0.81 (3.20)	<i>t</i> (12) = 0.92, <i>p</i> = 0.189	0.25	0.63	38	15	46
E	-12.43 (7.60)	<b><i>t</i>(13) = -6.12, <i>p</i> &lt; 0.001</b>	1.64	1,391.00	93	0	7	1.68 (3.19)	<b><i>t</i>(13) = 1.97, <i>p</i> = 0.035</b>	0.53	2.31	29	14	57
S	-10.21 (6.47)	<b><i>t</i>(14) = -6.12, <i>p</i> &lt; 0.001</b>	1.58	1846.00	73	0	27	0.37 (2.06)	<i>t</i> (14) = 0.69, <i>p</i> = 0.250	0.18	0.48	27	20	53
R	-13.98 (7.96)	<b><i>t</i>(55) = -13.15, <i>p</i> &lt; 0.001</b>	1.76	> 10 <sup>15</sup>	95	0	5	1.40 (3.39)	<b><i>t</i>(55) = 3.08, <i>p</i> = 0.002</b>	0.41	19.10	27	7	66

**Table 1.** The numerical distance and the numerical size effects. \*One sample *t*-test against zero (one sided); \*\*significant results are marked with a bold font; *M* mathematicians, *E* engineers, *S* social scientists, *R* reference group.

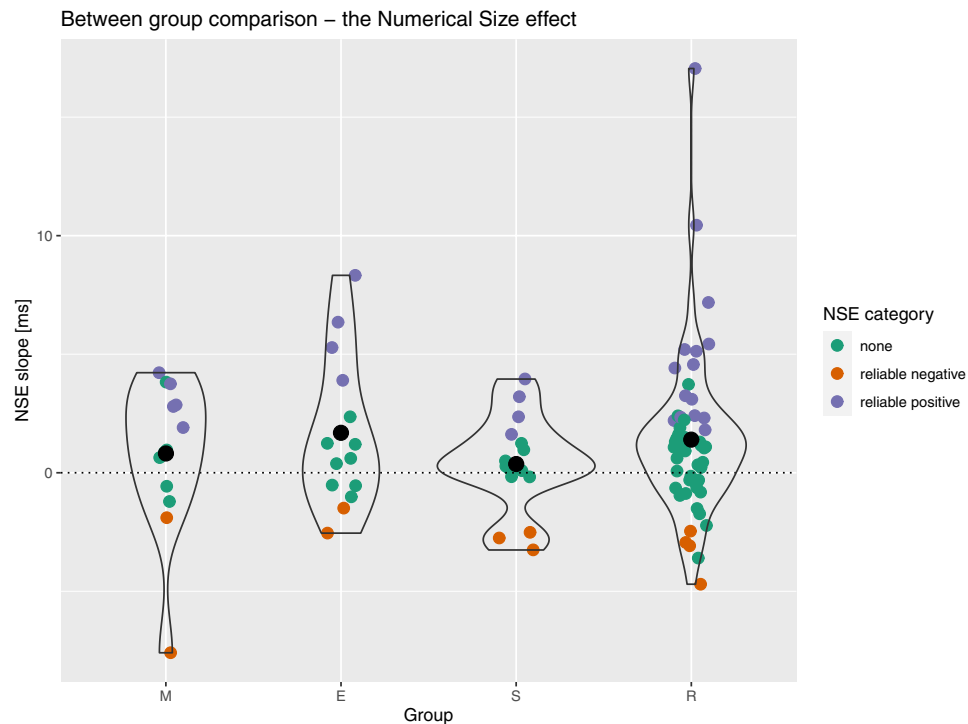


**Figure 1.** Summary of the NDE results. *M* = mathematicians, *E* = engineers, *S* = social scientists, *R* = reference group. Black dots represent group means. Coloured points (horizontally jittered) depict individual participants. Colour represents whether a given participant had a reliable NDE or not (as calculated using the bootstrapping method).

## Discussion

**Overview.** This study aimed to investigate how two hallmark effects of the analogue representation of symbolic numerical magnitude, namely the NDE and the NSE, relate to mathematical skills. In particular, we were interested in these phenomena at a very high level of mathematical skills operationalized in terms of formal education. For this purpose, we recruited four groups of participants: professional mathematicians, engineers, social scientists, and a reference group. We administered the magnitude classification task with single-digit Arabic numbers. Secondly, we checked the individual prevalence of the phenomena of interest, i.e., how many individuals reveal a reliable NDE and NSE.

We did not find between group differences, despite replicating a robust group level NDE both at the whole sample level and in each group separately. Bayesian analysis provided direct support for no between group differences in NDE. The NSE was also robust at the whole sample level, in the reference group and in the engineers



**Figure 2.** Summary of the NSE results. M = mathematicians, E = engineers, S = social scientists, R = reference group. Black dots represent group means. Coloured points (horizontally jittered) depict individual participants. Colour represents whether the given participant had a reliable size effect, a reliable reverse size effect, or no reliable size effect (as calculated using the bootstrapping method).

group, but it did not reach significance in any other group. On the other hand, again, groups did not differ with respect to the NSE, which was also supported by Bayesian evidence. Analysis of individual prevalence also did not reveal any between group differences as regards both phenomena under scrutiny. Notably, the lack of between group differences cannot be accounted for by unequal proportions of male and female participants: there was no overall gender difference as regards the NDE and the NSE, and the above patterns of results were similar when only male participants were analysed (see Supplementary Material 2).

Nevertheless, the analysis of individual prevalence provided insight into general properties of the NDE and NSE. The NDE seems to be reliably present in virtually all participants, and none of the participants revealed a reliable reverse NDE. The results pattern for the NSE was different. There were several individuals who had a reliable reverse NSE, but more than 60% of the participants did not display either a reverse or a typical NSE.

**Analogue magnitude processing of symbolic numbers in mathematicians: The numerical distance and size effects.** Taking into account that due to intense training and daily exposure to symbolic numbers, professional mathematicians might be expected to differ in their ability to access number magnitude information (or that the specific type of analogue representation has helped them to become professional mathematicians). On the other hand, previous studies suggest that the size of the NDE remains unchanged until adulthood<sup>20,21</sup>, and it presumably does not depend on the mathematical skills level<sup>58</sup>, at least when groups without learning problems are considered. Our study generally aligns with this pattern of results. Notably, even when groups with mathematical difficulties are taken into account, available evidence together suggests no differences in the measured NDE.

Regarding the NSE, the only intergroup differences reported up to date concern individuals with mathematical learning problems and math-anxiety. Reports on the relationship between the level of mathematical competence and the NSE are not known to us. Rousselle and Noël<sup>64</sup> found that mathematical learning problems go hand in hand with a reduced NSE, while Núñez-Peña and Suárez-Pellicioni<sup>69</sup> showed a correlation between high levels of math-anxiety and a larger NSE. Although the results of Rousselle and Noël are hard to elucidate theoretically, Núñez-Peña and Suárez-Pellicioni suggest that highly math-anxious individuals have a less precise access to numerical magnitude.

In our study we found no differences between mathematicians and other groups. Importantly, the same pattern of results remained unchanged irrespective of the approach we used. It held both for unstandardized and standardized NDE and NSE slopes, as well as when we took into account proportions of participants revealing reliable effects.

**What constitutes extremely high mathematical skills level?** Testing extreme groups within a given domain has provided instructive insights in several fields of psychology and cognitive science<sup>78</sup>. Extreme groups may reveal effects that are blurred in typical level groups (e.g., due to limited variance). Also, the field of numerical cognition has gained valuable insights by testing extreme groups in terms of mathematical skills. Investigations on mathematical cognition have mostly been focused on groups displaying mathematical difficulties<sup>79–81</sup>. Studies performed on groups displaying high mathematical skills levels are rare, and “high-level math ability” has been inconsistently defined; sometimes as prodigious calculators or sometimes as professional mathematicians. Prodigious calculators are individuals who perform complex arithmetic tasks very fast and efficiently<sup>82</sup>. Nevertheless, their skills are usually limited to a set of arithmetic problems and originate mostly from extended drill<sup>19</sup>. On the other hand, academic mathematical expertise is typically understood as being able to swiftly operate on mathematical theorems, concepts, conduct rigorous proofs, and discover new mathematical laws in a creative way<sup>83–85</sup>, which may doubly dissociate from calculation proficiency<sup>86</sup>.

Prodigious calculators described in the literature can hardly be considered to show extremely high levels of mathematical skills. Mathematical education does not solely aim at excelling in mental calculation procedures, it instead aims at an increased understanding of mathematical concepts and operating on them, aspects typically not mastered by prodigious calculators<sup>86</sup>. This means that previous research showing the NDE in a calculation prodigy<sup>66</sup> cannot be generalized to professional mathematicians. Although many fields of professional practice, like engineering, require familiarity with advanced mathematical tools, individuals involved in them share the above characteristics of professional mathematical activity very rarely. As in our previous study<sup>67</sup>, here, we considered engineers and mathematicians as separate groups.

Despite the fact that the existing knowledge-base is quite narrow, previous studies revealed that professional mathematicians differ from non-mathematicians in several cognitive aspects. Although there is no space here to review all of them, these differences result from a configuration of domain-general factors such as fluid intelligence<sup>67</sup>, arithmetic operations skills<sup>68,87–89</sup>, and spatial-numerical mappings<sup>67,68</sup>. Analogue processing of symbolic numerical magnitudes constitutes the understudied category of basic numerical skills of professional mathematicians. It is worth noting that this category should be considered as distinct from spatial-numerical associations<sup>90</sup>.

**The construct validity of the common magnitude system for symbolic number magnitude effects.** The correlation between the NDE and the NSE was low, even when adjusting for non-perfect task reliabilities (Supplementary Material 4). Importantly, the correlation vanished when we controlled for overall RT. This finding is somehow surprising since both phenomena are considered to be manifestations of an analogue magnitude representation<sup>19,32</sup>. Usually in diagnostics, for measures taken to indexing the same underlying construct, one would expect and require high intercorrelations of related measures. It is important to note that the reliability of our measures was still satisfactory, so that reliability issues cannot account for the low correlation and not at all, for the null correlation, when overall RT is partialled out. For these reasons, these results challenge the construct validity of a common magnitude system like the ANS for the case of symbolic number magnitude processing.

On the other hand, our findings can be accounted for by another framework for magnitude processing, namely Krajcsi’s DSS framework<sup>34,91,92</sup>. It postulates that the symbolic NDE and NSE emerge from two independent mechanisms. The DSS stores symbolic numbers in an amodal way as nodes within a conceptual network. Numerical properties are encoded as the connections between these nodes. The NDE reflects the semantic distance between the numbers, while the NSE appears as the consequence of the frequency of symbols, i.e., the fact that larger numbers are typically less frequent than smaller ones in daily life<sup>93</sup>. On the other hand, non-symbolic comparative judgements could still be processed by the ANS. In contrast to the ANS framework that assumes a strong correlation between the symbolic NDE and NSE, within the DSS framework, they could be partly dissociated. Indeed, Krajcsi<sup>91</sup> found no correlation between the NDE and the NSE when participants processed symbolic numbers. Our results corroborate this pattern and could be interpreted in a line with the DSS framework of human magnitude representation of symbolic numbers.

**The individual prevalence of symbolic numerical distance and size effects.** We found a reliable NDE for almost all participants. On the other hand, prevalence of the NSE was much lower. Only 29% of participants displayed a reliable NSE, a reverse reliable NSE occurred for about 11% of the individuals, while 60% did not have a reliable NSE.

Using the recent distinction of psychological phenomena into dominant and indominant ones<sup>70</sup>, we can conclude that although the NDE is a dominant phenomenon occurring in all individuals (like the Stroop effect), and the reverse effect is not observed at all, the NSE seems to be indominant. In the field of numerical cognition a pattern similar to the NSE was recently observed for the SNARC effect that appears to be reliably present only in about 45% of individuals<sup>75</sup>. Note, however, that these phenomena belong to different categories of basic numerical cognition: the NSE and the NDE do not have a spatial component<sup>42,43</sup>. The dissociation we found calls for future research, because it seems to be at odds with the most common interpretation of the NDE and the NSE.

**Limitations of the study and future research directions.** In this study, we used the symbolic magnitude classification task with single-digit Arabic numbers. There are at least two limits of our investigation. The first one concerns the relative ease of the task. The goal of this study was to examine whether professional mathematicians do already differ in basic numerical effects, which are supposed to index basic numerical representations like the symbolic magnitude representation. To best our knowledge, no one has tested before the hallmark effect of the symbolic magnitude processing, the symbolic numerical distance effect in professional mathemati-



cians. Since the symbolic numerical distance effect is observed in the overwhelming majority of studies with single-digit numbers, we decided to use single-digit stimuli. We believe that our approach is simple but not too simple since we found stable and robust effects of interest in all tested groups. Undoubtedly, further research should test whether professional mathematicians differ from other populations in the processing of more complex, i.e., multi-digit, numerical material. We hypothesize that using two-digit numbers could change the pattern of results, especially reveal differences in the NDE and NSE between mathematicians and other groups, because they are more complex. Especially, place-value processing as an additional process that is automatically initiated, whenever two-digit numbers are processed<sup>94</sup>. However, we claim that presumed differences would not only be due to the numerical difficulty level of the task (larger number size, less frequent numbers, additional place-value processing), but other factors producing considerable differences in single-digit and multi-digit number processing as the involvement of strategies and domain general-processes<sup>38–41</sup>. To sum up, our data should be interpreted with care. The fact that professional mathematicians do not differ from controls in the most basic tasks and effects of symbolic numerical magnitude processing does not imply that there will also be null differences in more complex stimuli or tasks.

The second limitation concerns a non-symbolic dimension of the analogue number magnitude representation. Although we are aware that a link between the non-symbolic magnitude processing and mathematical skills is well documented<sup>48,52,55,95</sup>, here we tested all the groups using only the symbolic magnitude classification task since the symbolic number processing accounts the more variance in mathematical skills<sup>53,54</sup>. A well-established tradition of mathematical cognition research, from Moyer and Landauer's study<sup>4</sup>, through Dehaene's triple code framework<sup>1</sup>, to more recent computational model<sup>35</sup>, accounts the processing of both symbolic and non-symbolic numerical magnitudes within the single analogue cognitive system usually called ANS<sup>19,32</sup>. Moreover, the shared behavioural patterns in the analogue magnitude processing, including all the numerical<sup>4–8</sup>, and non-numerical instances<sup>9–16</sup> could be accounted within a unified ATOM framework<sup>17</sup>. These frameworks focus on similarities, but this does not exclude the existence of differences within the magnitude system depending on the task. This is substantiated in the fact that the symbolic NDE is uncorrelated with the non-symbolic NDE<sup>96</sup>. Moreover, Krajcsi's framework discussed above assumes that symbolic and non-symbolic numerical magnitudes are handled by independent cognitive systems (namely non-symbolic by ANS and symbolic by DSS)<sup>34,91</sup>. Therefore, a generalization of the behavioural pattern we found in the symbolic processing domain to the non-symbolic one would be unsound. Regardless of the applied theoretical framework, differences between professional mathematicians and other groups in analogue magnitude tasks involving non-symbolic numbers (and non-numerical material) could exist and should be accounted for in subsequent studies.

## Conclusions

Analogue magnitude is one of multiple mental representations of symbolic numbers, and the NDE and the NSE constitute its primary instances. Even though these phenomena were previously revealed in various human cultures and age groups, as well as in non-human animals, there is no consensus about their relationship to mathematical skills. Furthermore, so far, studies on the NDE and the NSE have been carried out at the group level, with nothing known about their individual prevalence. Testing professional mathematicians, engineers, social scientists, and the reference group, we found no between group differences in the NSE and the NDE. This observation allows us to infer that the professional training and practice of mathematicians does not change their analogue magnitude representation of numbers or alternatively, that possessing a specific type of magnitude representation does not foster mastering professional mathematical skills. Looking at the prevalence of the phenomena for single-digit numbers at the individual level, we found a reliable NDE in almost all participants, whilst the prevalence of a reliable NSE was surprisingly much lower. This indicates that the former effect is dominant, whereas the latter is indominant. This last conclusion especially calls for further research on whether NDE and NSE truly reflect properties of the same system of representing magnitude, which is assumed to be universal in all humans. Importantly, this conclusion cannot be accounted for by the fact that NSE was just weaker than the NDE (e.g., because we used single-digit numbers only): when controlling for the mean RT, NDE and NSE did not correlate with each other (see Table S2 in Supplementary Material 4).

## Data availability

Experimental task used in the study as well as both the data and analysis script are available at the Open Science Framework (<https://doi.org/10.17605/OSF.IO/MSDNR>).

Received: 30 January 2020; Accepted: 11 June 2020

Published online: 13 July 2020

## References

- Dehaene, S. Varieties of numerical abilities. *Cognition* **44**, 1–42 (1992).
- Dehaene, S. & Cohen, L. Towards an anatomical and functional model of number processing. *Math. Cogn.* **1**, 83–120 (1995).
- Klein, E. *et al.* Considering structural connectivity in the triple code model of numerical cognition: Differential connectivity for magnitude processing and arithmetic facts. *Brain Struct. Funct.* **221**, 979–995 (2014).
- Moyer, R. S. & Landauer, T. K. Time required for judgements of numerical inequality. *Nature* **215**, 1519–1520 (1967).
- Restle, F. Speed of adding and comparing numbers. *J. Exp. Psychol.* **83**, 274–278 (1970).
- Dehaene, S., Dupoux, E. & Mehler, J. Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 626–641 (1990).
- Buckley, P. B. & Gillman, C. B. Comparisons of digits and dot patterns. *J. Exp. Psychol.* **103**, 1131–1136 (1974).
- van Oeffelen, M. P. & Vos, P. G. A probabilistic model for the discrimination of visual number. *Percept. Psychophys.* **32**, 163–170 (1982).
- Johnson, D. M. Confidence and speed in the two-category judgment. *Arch. Psychol.* **241**, 1–52 (1939).

10. Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P. & Orban, G. A. Parietal representation of symbolic and nonsymbolic magnitude. *J. Cogn. Neurosci.* **15**, 47–56 (2003).
11. Moyer, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. *Percept. Psychophys.* **13**, 180–184 (1973).
12. Maloney, E. A., Barr, N., Risko, E. F. & Fugelsang, J. A. Working memory load dissociates common indices of the numerical distance effect: Implications for the study of numerical cognition. *J. Numer. Cogn.* **5**, 337–357 (2019).
13. Pinel, P., Piazza, M., Le Bihan, D. & Dehaene, S. Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* **41**, 983–993 (2004).
14. Cohen Kadosh, R., Cohen Kadosh, K. & Henik, A. When brightness counts: The neuronal correlate of numerical-luminance interference. *Cereb. Cortex* **18**, 337–343 (2008).
15. Banks, W. P. & Flora, J. Semantic and perceptual processes in symbolic comparisons. *J. Exp. Psychol. Hum. Percept. Perform.* **3**, 278–290 (1977).
16. Cohen Kadosh, R., Lammertyn, J. & Izard, V. Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Prog. Neurobiol.* **84**, 132–147 (2008).
17. Walsh, V. A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* **7**, 483–488 (2003).
18. Pica, P., Lemer, C., Izard, V. & Dehaene, S. Exact and approximate arithmetic in an Amazonian indigene group. *Science* **306**, 499–503 (2004).
19. Dehaene, S. *The Number Sense* (Oxford University Press, Oxford, 2011).
20. Ansari, D., Garcia, N., Lucas, E., Hamon, K. & Dhital, B. Neural correlates of symbolic number processing in children and adults. *NeuroReport* **16**, 1769–1773 (2005).
21. Huntley-Fenner, G. & Cannon, E. Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychol. Sci.* **11**, 147–152 (2000).
22. Dehaene, S., Dehaene-Lambertz, G. & Cohen, L. Abstract representations of numbers in the animal and human brain. *Trends Neurosci.* **21**, 355–361 (1998).
23. Vallortigara, G. An animal's sense of number. In *The nature and Development of Mathematics: Cross Disciplinary Perspectives on Cognition, Learning and Culture* (eds Adams, J. W. et al.) 43–66 (Routledge, Abingdon, 2017).
24. Nieder, A. The adaptive value of numerical competence. *Trends Ecol. Evol.* **35**, 605–617 (2020).
25. Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D. & Dyer, A. G. Numerical ordering of zero in honey bees. *Science* **360**, 1124–1126 (2018).
26. Agrillo, C., Dadda, M., Serena, G. & Bisazza, A. Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Anim. Cogn.* **11**, 495–503 (2008).
27. Stancher, G., Rugani, R., Regolin, L. & Vallortigara, G. Numerical discrimination by frogs (*Bombina orientalis*). *Anim. Cogn.* **18**, 219–229 (2015).
28. Scarf, D., Hayne, H. & Colombo, M. Pigeons on par with primates in numerical competence. *Science* **334**, 1664 (2011).
29. Cantlon, J. F. & Brannon, E. M. Basic math in monkeys and college students. *PLoS Biol.* **5**, 2912–2919 (2007).
30. Tomonaga, M. & Matsuzawa, T. Sequential responding to arabic numerals with wild cards by the chimpanzee (*Pan troglodytes*). *Anim. Cogn.* **3**, 1–11 (2000).
31. Vallortigara, G. Core knowledge of object, number, and geometry: A comparative and neural approach. *Cogn. Neuropsychol.* **29**, 213–236 (2012).
32. Feigenson, L., Dehaene, S. & Spelke, E. S. Core systems of number. *Trends Cogn. Sci.* **8**, 307–314 (2004).
33. Sasanguie, D., Smedt, B. D. & Reynvoet, B. Evidence for distinct magnitude systems for symbolic and non-symbolic number the existence of two different systems for processing. *Psychol. Res.* **81**, 231–242 (2017).
34. Krajcsi, A., Lengyel, G. & Kojouharova, P. The source of the symbolic numerical distance and size effects. *Front. Psychol.* **7**, 1–16 (2016).
35. Gevers, W., Verguts, T., Reynvoet, B., Caessens, B. & Fias, W. Numbers and space: A computational model of the SNARC effect. *J. Exp. Psychol. Hum. Percept. Perform.* **32**, 32–44 (2006).
36. Parkman, J. M. Temporal aspects of digit and letter inequality judgments. *J. Exp. Psychol.* **91**, 191–205 (1971).
37. Brysbaert, M. Arabic number reading: On the nature of the numerical scale and the origin of phonological recoding. *J. Exp. Psychol. Gen.* **124**, 434–452 (1995).
38. Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K. & Fischer, M. H. Extending the mental number line: A review of multi-digit number processing. *Z. Psychol.* **219**, 3–22 (2011).
39. Knops, A., Nuerk, H.-C. & Göbel, S. M. Domain-general factors influencing numerical and arithmetic processing. *J. Numer. Cogn.* **3**, 112–132 (2017).
40. Hohol, M., Cipora, K., Willmes, K. & Nuerk, H.-C. Bringing back the balance: Domain-general processes are also important in numerical cognition. *Front. Psychol.* **8**, 499 (2017).
41. Bahnueller, J., Nuerk, H. C. & Moeller, K. A taxonomy proposal for types of interactions of language and place-value processing in multi-digit numbers. *Front. Psychol.* **9**, 1024 (2018).
42. Bonato, M., Zorzi, M. & Umiltà, C. When time is space: Evidence for a mental time line. *Neurosci. Biobehav. Rev.* **36**, 2257–2273 (2012).
43. Nuerk, H.-C., Bauer, F., Krummenacher, J., Heller, D. & Willmes, K. The power of the mental number line: How the magnitude of unattended numbers affects performance in an Eriksen task. *Psychol. Sci.* **47**, 34–50 (2005).
44. Dietrich, J. F., Huber, S. & Nuerk, H. C. Methodological aspects to be considered when measuring the approximate number system (ANS)—A research review. *Front. Psychol.* **6**, 295 (2015).
45. Lyons, I. M., Nuerk, H.-C. & Ansari, D. Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *J. Exp. Psychol. Gen.* **144**, 1021–1035 (2015).
46. Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).
47. Lakoff, G. & Núñez, R. E. *Where Mathematics Comes from* (Basic Books, New York, 2000).
48. Feigenson, L., Libertus, M. E. & Halberda, J. Links between the intuitive sense of number and formal mathematics ability. *Child Dev. Perspect.* **7**, 74–79 (2013).
49. Cipora, K., He, Y. & Nuerk, H.-C. The spatial-numerical association of response codes effect and math skills: Why related? *Ann. N. Y. Acad. Sci.* <https://doi.org/10.1111/nyas.14355> (2020)
50. Cipora, K., Patro, K. & Nuerk, H.-C. Are spatial-numerical associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests no. *Mind Brain Educ.* **9**, 190–206 (2015).
51. Cipora, K., Schroeder, P. A., Soltanlou, M. & Nuerk, H.-C. More space, better mathematics: Is space a powerful tool or a cornerstone for understanding arithmetic? In *Visualizing Mathematics: The Role of Spatial Reasoning in Mathematical Thought* (eds Mix, K. S. & Battista, M. T.) 77–116 (Springer, Berlin, 2018).
52. De Smedt, B., Noël, M.-P., Gilmore, C. K. & Ansari, D. 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 Neurosci. Educ.* **2**, 48–55 (2013).
53. Schneider, M. et al. Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev. Sci.* **20**, 1–16 (2017).

54. Hawes, Z., Nosworthy, N., Archibald, L. & Ansari, D. Kindergarten children's symbolic number comparison skills predict 1st grade mathematics achievement: Evidence from a two-minute paper-and-pencil test. *Learn. Instr.* **59**, 21–33 (2019).
55. Szklarek, E. & Brannon, E. M. Does the approximate number system serve as a foundation for symbolic mathematics?. *Lang. Learn. Dev.* **13**, 171–190 (2017).
56. De Smedt, B., Verschaffel, L. & Ghesquière, P. The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *J. Exp. Child Psychol.* **103**, 469–479 (2009).
57. Sekuler, R. & Mierkiewicz, D. Children's judgments of numerical inequality. *Child Dev.* **48**, 630 (1977).
58. Bugden, S. & Ansari, D. Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition* **118**, 32–44 (2011).
59. Gebuis, T. & Reynvoet, B. Number representations and their relation with mathematical ability. In *The Oxford Handbook of Numerical Cognition* (eds Cohen Kadosh, R. & Dowker, A.) 331–344 (Oxford University Press, Oxford, 2015).
60. Ashkenazi, S., Mark-Zigdon, N. & Henik, A. Numerical distance effect in developmental dyscalculia. *Cogn. Dev.* **24**, 387–400 (2009).
61. Mussolin, C., Mejias, S. & Noël, M.-P. Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition* **115**, 10–25 (2010).
62. Maloney, E. A., Ansari, D. & Fugelsang, J. A. The effect of mathematics anxiety on the processing of numerical magnitude. *Q. J. Exp. Psychol.* **64**, 10–16 (2011).
63. Soltész, F., Szucs, D., Dékány, J., Márkus, A. & Csépe, V. A combined event-related potential and neuropsychological investigation of developmental dyscalculia. *Neurosci. Lett.* **417**, 181–186 (2007).
64. Rousselle, L. & Noël, M.-P. Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition* **102**, 361–395 (2007).
65. Butterworth, B. *The Mathematical Brain* (Macmillan, New York, 1999).
66. Pesenti, M., Seron, X., Samson, D. & Duroux, B. Basic and exceptional calculation abilities in a calculating prodigy: A case study. *Math. Cogn.* **5**, 97–148 (1999).
67. Cipora, K. et al. Professional mathematicians differ from controls in their spatial-numerical associations. *Psychol. Res.* **80**, 710–726 (2016).
68. Sella, F., Sader, E., Lollot, S. & Cohen Kadosh, R. Basic and advanced numerical performances relate to mathematical expertise but are fully mediated by visuospatial skills. *J. Exp. Psychol. Learn. Mem. Cogn.* **42**, 1458–1472 (2016).
69. Núñez-Peña, M. I. & Suárez-Pellicioni, M. Less precise representation of numerical magnitude in high math-anxious individuals: An ERP study of the size and distance effects. *Biol. Psychol.* **103**, 176–183 (2014).
70. Rouder, J. N. & Haaf, J. M. Power, dominance, and constraint: A note on the appeal of different design traditions. *Adv. Methods Pract. Psychol. Sci.* **1**, 19–26 (2018).
71. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
72. Forster, K. I. & Forster, J. C. DMDX: A Windows display program with millisecond accuracy. *Behav. Res. Methods Instrum. Comput.* **35**, 116–124 (2003).
73. Cipora, K. Between task consistency of the SNARC effect. in *32 European Workshop on Cognitive Neuropsychology* (2014).
74. R Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, 2013).
75. Cipora, K. et al. A minority pulls the sample mean: on the individual prevalence of robust group-level cognitive phenomena—The instance of the SNARC effect. *PsyArXiv Prepr.* <https://doi.org/10.31234/osf.io/bwyr3> (2019).
76. Morey, R. D. et al. *BayesFactor* (2018).
77. Cipora, K. & Nuerk, H.-C. Is the SNARC effect related to the level of mathematics? No systematic relationship observed despite more power, more repetitions, and more direct assessment of arithmetic skill. *Q. J. Exp. Psychol.* **66**, 1974–1991 (2013).
78. Preacher, K. J., Rucker, D. D., MacCallum, R. C. & Nicewander, W. A. Use of the extreme groups approach: A Critical reexamination and new recommendations. *Psychol. Methods* **10**, 178–192 (2005).
79. Göbel, S. M. Number processing and arithmetic in children and adults with reading difficulties. In *The Oxford Handbook of Numerical Cognition* (eds Cohen Kadosh, R. & Dowker, A.) 696–720 (Oxford University Press, Oxford, 2015).
80. Kucian, K., Kaufmann, L. & von Aster, M. Brain correlates of numerical disabilities. In *The Oxford Handbook of Numerical Cognition* (eds Cohen Kadosh, R. & Dowker, A.) 732–744 (Oxford University Press, Oxford, 2015).
81. Hoard, M. H. Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In *Handbook of Mathematical Cognition* (ed. Campbell, J. I. D.) 253–259 (Psychology Press, New York, 2005).
82. Fehr, T., Weber, J., Willmes, K. & Herrmann, M. Neural correlates in exceptional mental arithmetic—About the neural architecture of prodigious skills. *Neuropsychologia* **48**, 1407–1416 (2010).
83. Hardy, G. H. *A Mathematician's Apology* (Cambridge University Press, Cambridge, 2012).
84. Sriraman, B. The characteristics of mathematical creativity. *Math. Educ.* **14**, 19–34 (2004).
85. Amalric, M. & Dehaene, S. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc. Natl. Acad. Sci.* **113**, 4909–4917 (2016).
86. Pesenti, M. Calculation abilities in expert calculators. In *Handbook of Mathematical Cognition* (ed. Campbell, J. I. D.) 413–430 (Psychology Press, New York, 2005).
87. Dowker, A. Computational estimation strategies of professional mathematicians. *J. Res. Math. Educ.* **23**, 45–55 (1992).
88. Dowker, A., Flood, A., Griffiths, H., Harriss, L. & Hook, L. Estimation strategies of four groups. *Math. Cogn.* **2**, 113–135 (1996).
89. Obersteiner, A., Van Dooren, W., Van Hoof, J. & Verschaffel, L. The natural number bias and magnitude representation in fraction comparison by expert mathematicians. *Learn. Instr.* **28**, 64–72 (2013).
90. Bonato, M., Fabbri, S., Umiltà, C. & Zorzi, M. The mental representation of numerical fractions: Real or integer?. *J. Exp. Psychol. Hum. Percept. Perform.* **33**, 1410–1419 (2007).
91. Krajcsi, A. Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychon. Bull. Rev.* **24**, 927–934 (2017).
92. Krajcsi, A., Kojouharova, P. & Lengyel, G. Processing symbolic numbers: The example of distance and size effects. *PsyArXiv Prepr.* <https://doi.org/10.31234/osf.io/5wzcx> (2020).
93. Dehaene, S. & Mehler, J. Cross-linguistic regularities in the frequency of number words. *Cognition* **43**, 1–29 (1992).
94. Kallai, A. Y. & Tzelgov, J. The place-value of a digit in multi-digit numbers is processed automatically. *J. Exp. Psychol. Learn. Mem. Cogn.* **38**, 1221–1233 (2012).
95. Halberda, J., Mazocco, M. M. & Feigenson, L. Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* **455**, 665–668 (2008).
96. Maloney, E. A., Risko, E. F., Preston, F., Ansari, D. & Fugelsang, J. A. Challenging the reliability and validity of cognitive measures: The case of the numerical distance effect. *Acta Psychol.* **134**, 154–161 (2010).

## Acknowledgements

Hans-Christoph Nuerk and Krzysztof Cipora were supported by a grant from the Deutsche Forschungsgemeinschaft (German Research Foundation) to Hans-Christoph Nuerk (NU 265/3-1). Edward Nęcka was supported by the research grant 2013/08/A/HS6/00045 funded by the National Science Centre, Poland (NCN).

### Author contributions

M.H. and K.C. wrote the draft of the main manuscript. M.H., K.W., E.N., B.B., H.C.N. and K.C. conceived and designed the experiment. M.H. and K.C. performed the experiment. KC analysed the data and prepared the figures and tables. M.H., K.W., E.N., B.B., H.C.N. and K.C. substantially contributed to the interpretation of data. M.H., K.W., E.N., B.B., H.C.N. and K.C. reviewed and accepted the final version of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-68202-z>.

**Correspondence** and requests for materials should be addressed to M.H. or K.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020