Professional mathematicians do not differ from others in analogue magnitude representation: Evidence from prevalence of numerical distance and size effects

Mateusz Hohol*, Klaus Willmes2, Edward Nęcka3, Bartosz Brożek1, Hans-Christoph Nuerk4,5 & Krzysztof Cipora4,5**

1 Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Krakow, Poland
2 Clinic of Neurology, University Hospital, RWTH Aachen University, Aachen, Germany
3 Institute of Psychology, Jagiellonian University, Krakow, Poland
4 Department of Psychology, University of Tuebingen, Tuebingen, Germany
5 LEAD Graduate School & Research Network, University of Tuebingen, Tuebingen, Germany

* corresponding author: mateusz.hohol@uj.edu.pl
** corresponding author: krzysztof.cipora@gmail.com

Abstract: 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 the analogue number magnitude representation. 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, nothing is known about the prevalence of these effects at the individual level. Here we compared professional mathematicians, engineers, social scientists, and a reference sample using the classical magnitude classification task. The groups did not differ with respect to numerical distance and size effects in frequentist or Bayesian analysis. 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 prevalent in mathematicians, suggesting that extremely high mathematical skills neither rely on nor have special consequences on analogue number magnitude processing.

Running head: Analogue numerical processing in professional mathematicians
Introduction

Numerical knowledge is encoded in multiple formats serving specific functions\(^1\)–\(^3\). 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, and other continua including line length\(^4\), physical object size\(^5,6\), luminance\(^7,8\), and non-directly perceptible properties like intelligence\(^9,10\). Walsh\(^11\) 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, both involving numerical and non-numerical instances, are often assumed to be driven by Weber’s law\(^12\), however, other approaches have also been proposed\(^13\). Analogue magnitude comparisons have been studied in different human cultures, languages and notations\(^14\)–\(^16\), age groups\(^17,18\), as well as various animal species\(^19,20\). The numerical distance effect (NDE) is one of the fundamental characteristics of analogue magnitude processing\(^21,22\). In the 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\(^23\). 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 characteristics of basic analogue processing have been traditionally elucidated by Restle’s\(^24\) 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)\(^16,25\) or 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\(^26\). Magnitude can be activated independently from its spatial association\(^27\).
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. 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. 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 representation 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 achievements, but other tasks reveal no correlation or somewhat mixed evidence.

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. Conversely, a larger NDE seems to be associated with a more imprecise analogue magnitude representation.

Although the analogue numerical magnitude representation and the NDE are considered to be, respectively, a fundamental representation and effect, 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. 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 sizes are considered rather than raw RTs. Nevertheless, despite size changes with age, the NDE remains robust in adulthood. The results of several studies on groups with typical mathematical skill levels did not provide straightforward results. Typically the size of the NDE does not explain a substantial amount of variance in mathematical skills. On the other hand, mathematical skill correlates at a moderate level with overall RTs (i.e., not the
NDE) in the magnitude classification task\textsuperscript{36}, 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\textsuperscript{37,38} through no differences in the NDE\textsuperscript{39,40} to a smaller NDE\textsuperscript{41} when compared to groups without mathematical difficulties. There are also case reports of a reverse NDE in dyscalculic individuals\textsuperscript{41,42}. The NDE was also observed in a calculation prodigy\textsuperscript{43}.

It seems that there is no genuinely strong and consistent relationship between NDE and mathematical skill 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, the groups displaying extremely high mathematical skill levels are largely understudied (i.e., to the best of our knowledge, not systematically tested as concerns the NDE). Professional mathematicians may differ from other groups in terms of NDE. For instance, they differ from controls in their SNARC effect\textsuperscript{44} and positive number mapping\textsuperscript{45}; thus, differences on 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.

\textit{Numerical size effect and mathematical skills}

Compared to the NDE, our knowledge on the relationship between the size of the NSE and mathematical skill is considerably weaker. As Rousselle and Noel\textsuperscript{41} 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\textsuperscript{46} found that highly math-anxious individuals revealed larger NSE (the same is true regarding 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 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. 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 they should both be dominant, i.e., present in virtually all individuals. To the best of our knowledge, no one has ever studied these phenomena in respect to individual prevalence.

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 skill levels 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 general population reference sample using the 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 proved a somewhat inconsistent picture of relationships between the analogue numerical magnitude representations and mathematical skills, as we 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 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 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 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 the individual prevalence, we expect to find that the NDE and NSE go hand-in-hand as dominant phenomena, since they both characterize the same aspect of basic numerical processing, namely 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 participants. 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) 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 advanced in doctoral studies (the exact dissertation topic approved by the departmental council). Although the educational background in the reference group varied, nobody met the inclusion criteria for M, E, and S,
nor was a PhD student. The members of the groups M, E and S are the same participants as described in Cipora et al. 44

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 here is reported based on this criterion as well. However, these participants also answered Oldfield’s handedness questionnaire 48, which allows determination of handedness in a more fine-grained manner. The Oldfield’s questionnaire score for each participant in the R group are reported in the shared data file.

Materials

We used a computerized magnitude classification task. The participant task was to decide whether a visually presented one-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 reversed response key assignment. In each block each number (1, 2, 3, 4, 6, 7, 8, 9) was presented 30 times. 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 after each trial, and information about response mapping was present on the bottom line of the screen. The order of blocks was counterbalanced among participants. In experimental blocks, no feedback and information about the response key assignment was present. Each trial started with an eye fixation cross presented for 300ms. Subsequently, the target number was presented until the participant responded or for a maximum duration of 2s. The next trial started after 500ms of blank screen presentation. A standard, MS Windows compatible computer running DMDX 49 was used to present stimuli and collect responses.

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 44,50. 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 minutes. 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 R language\textsuperscript{51}. Both the data and analysis script are freely available at Open Science Framework (https://osf.io/msdnr/; DOI: 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 Supplementary Material\textsuperscript{52}.

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 regressions 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 perfectly orthogonal, so there is no collinearity problem. Slopes corresponding to these predictors were measures of the NDE and the NSE, respectively. 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 1). In the case of numerical distance, negative slopes correspond to the typical NDE, the more negative they are, the stronger the numerical distance effect. 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 NDE and positive values for NSE). Both frequentist and Bayesian t-tests were used. Group comparisons were conducted by means of unianova and the BF were computed with the anovaBF function of the R package BayesFactor\textsuperscript{53}.

In the following step, we aimed to investigate the presence of the NDE and NSE at the individual level. Specifically, the regression method does not allow for the making of inferences about the presence of effects of interest at the individual level. This is possible with a bootstrapping approach\textsuperscript{52}. Here we adapted a H0 bootstrapping approach proposed by Cipora et al.\textsuperscript{52}. 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 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 5000 times. The slopes from these bootstrap based regressions were considered as possible outcomes of the analysis if there is no NDE and 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 the case of NDE, if the empirical slope was < 0 and it was outside the 90% CI, a participant was considered as having a reliable NDE. If the slope was positive and it was outside the 90% CI, the participant was considered as having a reliable reverse NDE. If the slope was within the 90% CI, the participant was considered as not having 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 in regard to proportions of participants displaying reliable typical, non-reliable, reliable reverse effects. (In Supplementary Material 2 we present an exhaustive 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. The results remained unchanged, so we do not report them in detail, however, they are available for inspection along with other shared analyses (https://osf.io/tw843/; DOI: 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%. An ANOVA on transformed accuracy data \(2\times\arcsin(\sqrt{\text{proportion correct}})\) revealed significant between group differences \(F(3, 94) = 3.30, p = .024, \eta^2_p = .095, BF = 2.21\). Post hoc analysis (HSD corrected) revealed that the E group had significantly higher performances than the R group.
(p = .047). However, due to very high overall performance, errors were not further analysed. Subsequently, the RT data were filtered. First, correct responses with RTs < 200ms (0.05% of all trials) were treated as anticipations and not further analysed. Eventually, a sequential trimming method was applied: for each participant RTs outside ±3SD 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 504ms (SD = 79). There was no between group differences in mean RT F(3, 94) = 1.41, p = .246, \( \eta^2_p = .05, \) BF = 0.31. Overall mean intraindividual variability in RT was 90ms (SD = 36). The intraindividual did not differ between groups, F(3,94) = 0.52, p = .669, \( \eta^2_p = .02, \) BF = 0.13.

It is worth noting that the reliability of the NDE and NSE was very high. In the case of the NDE it was .96 and in case of the NSE it was .91.
Table 1. The numerical distance and the numerical size effects.

| Group | Mean (SD) | t-test* | d | BF10 | Proportion | Mean (SD) | t-test* | d | BF10 | Proportion |
|-------|-----------|---------|---|------|------------|-----------|---------|---|------|------------|
|       |           |         |   |      |             |           |         |   |      |             |
|       |           |         |   |      |             |           |         |   |      |             |
| Overall | -13.08 (7.90) | t(97) = -16.39, p < .001** | 1.66 | > 10^{26} | 91% | 0% | 9% | 1.20 (3.16) | t(97) = 3.77, p < .001 | 0.38 | 142.00 | 29% | 11% | 60% |
| M      | -13.19 (9.41) | t(12) = -5.06, p < .001 | 1.40 | 234.00 | 92% | 0% | 8% | 0.81 (3.20) | t(12) = 0.92, p = .189 | 0.25 | 0.63 | 38% | 15% | 46% |
| E      | -12.43 (7.60) | t(13) = -6.12, p < .001 | 1.64 | 1391.00 | 93% | 0% | 7% | 1.68 (3.19) | t(13) = 1.97, p = .035 | 0.53 | 2.31 | 29% | 14% | 57% |
| S      | -10.21 (6.47) | t(14) = -6.12, p < .001 | 1.58 | 1846.00 | 73% | 0% | 27% | 0.37 (2.06) | t(14) = 0.69, p = .250 | 0.18 | 0.48 | 27% | 20% | 53% |
| R      | -13.98 (7.96) | t(55) = -13.15, p < .001 | 1.76 | > 10^{15} | 95% | 0% | 5% | 1.40 (3.39) | t(55) = 3.08, p = .002 | 0.41 | 19.10 | 27% | 7% | 66% |

Notes. * one sample t-test against zero (one sided); ** Significant results are marked with a bold font.
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 effect did not differ between groups $F(3, 94) = 0.93, p = .428, \eta_p^2 = .03, BF = 0.20$. The results of the bootstrapping analysis are summarized in Figure 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 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, $p = .085$).

Figure 1. Summary of the NDE results. Black dots represent group means. Colourful points (horizontally jittered) depict individual participants. Colour represents whether a given participant had a reliable NDE or not (as calculated using the bootstrapping method).
**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 M, E, and S groups Bayesian evidence was highly inconclusive. Again, groups did not differ regarding the NSE, $F(3, 94) = 0.59, p = .626, \eta_p^2 = .02, BF = 0.13$. The results of the bootstrapping analysis are summarized in Figure 2. Interestingly, only 29% of participants displayed a reliable NSE. On the other hand, 11% of 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 $p = .629$).

![Figure 2](image.png)

*Figure 2.* Summary of the NSE results. 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).
Discussion

Overview

This study aimed to investigate how two hallmarks of the analogue representation of 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 skill operationalized in terms of formal education. For this purpose, we recruited four groups of participants: professional mathematicians, engineers, social scientists, and a reference sample. We administered a magnitude classification task. Secondly, we checked the individual prevalence of the phenomena of interest, i.e., how many individuals reveal a reliable NDE and NSE.

Despite replicating a robust group level NDE both at the whole sample level, and in each group separately, we did not find between group differences. 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 sample and in the engineers 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.

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 reversed NDE. The results pattern for the NSE was different. There were several individuals who had a reliable reverse NSE, but more than 60% of participants did not display either a reverse or a typical NSE.

Analogue magnitude processing in mathematicians: Numerical distance and size effects

Taking into account that due to intense training and daily exposure to 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\textsuperscript{17,18}, and it presumably does not depend on the mathematical skill level\textsuperscript{15}, 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, existing pieces of evidence together suggest 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 found that mathematical learning problems go hand in hand with a reduced NSE, while Núñez-Peña and Suárez-Pellicioni 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. 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. Studies performed on groups displaying high mathematical skill 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. Nevertheless, their skills are usually limited to a set of arithmetic problems and originate mostly from extended drill. 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, which may doubly dissociate from calculation proficiency.

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. This means that previous research showing the NDE in a calculation prodigy 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, 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, arithmetic operations skills, and spatial-numerical mappings. Analogue magnitude processing 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.

The individual prevalence of numerical distance and size effects

We found a reliable NDE for almost all participants. A reversed reliable NDE occurred for about 11% of individuals. On the other hand, prevalence of the NSE was much lower. Only 29% of participants displayed a reliable NSE, while 60% did not have a reliable NSE.

The difference in the presence of reliable effects is surprising, although both the NDE and the NSE are commonly considered as manifestations of analogue magnitude representation. We used a magnitude classification task with single-digit numbers, however differences in single-digit and multi-digit number processing have been reported. The dissociation we found calls for future research, because it seems to be at odds with the most common interpretation of NDE and NSE. Using the recent distinction of psychological phenomena into dominant and indominant ones, 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.
Conclusions

Analogue magnitude one of multiple mental representations of 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 skill. Furthermore, so far, studies on the NDE and the NSE have been at the group level, with nothing known about their individual prevalence. Testing professional mathematicians, engineers, social scientists, and the reference sample, 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 NDE was just weaker than NDE (e.g., because we used single-digit numbers only): when controlled for the mean RT, NDE and NSE did not correlate with each other (see Table S2 in Supplementary Material 2).

Contributions

MH and KC wrote the main manuscript. All authors conceived and designed the experiment. MH and KC performed the experiment. KC analysed the data and prepared the figures and tables. All authors substantially contributed to the interpretation of data. All authors reviewed the manuscript.

Competing interests statement

The authors declare that they have no competing interests.

Data availability

Both the data and analysis script are openly available at the Open Science Framework: https://osf.io/msdnr/ (DOI: 10.17605/OSF.IO/MSDNR).
Funding

Mateusz Hohol was supported by the research grant 2015/19/B/HS1/03310 funded by the National Science Centre, Poland (NCN). Hans-Christoph Nuerk and Krzysztof Cipora were supported by a grant from the Deutche 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).

References

1. Dehaene, S., Piazza, M., Pinel, P. & Cohen, L. Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506 (2003).

2. 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).

3. Schmithorst, V. J. & Brown, R. D. Empirical validation of the triple-code model of numerical processing for complex math operations using functional MRI and group Independent Component Analysis of the mental addition and subtraction of fractions. *Neuroimage* 22, 1414–1420 (2004).

4. Johnson, D. M. Confidence and speed in the two-category judgment. *Arch. Psychol.* 241, 1–52 (1939).

5. Moyer, R. S. Comparing objects in memory: Evidence suggesting an internal psychophysics. *Percept. Psychophys.* 13, 180–184 (1973).

6. 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).

7. 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).

8. Cohen Kadosh, R., Kadosh, K. C. & Henik, A. When brightness counts: The neuronal correlate of numerical-luminance interference. *Cereb. Cortex* 18, 337–343 (2008).

9. Banks, W. P. & Flora, J. Semantic and perceptual processes in symbolic comparisons.
10. 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).

11. Walsh, V. A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* **7**, 483–488 (2003).

12. 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).

13. Krajcsi, A., Lengyel, G. & Kojouharova, P. The source of the symbolic numerical distance and size effects. *Front. Psychol.* **7**, 1–16 (2016).

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

15. Buckley, P. B. & Gillman, C. B. Comparisons of digits and dot patterns. *J. Exp. Psychol.* **103**, 1131–1136 (1974).

16. Dehaene, S. *The number sense*. (Oxford University Press, 2011).

17. Ansari, D., Garcia, N., Lucas, E., Hamon, K. & Dhital, B. Neural correlates of symbolic number processing in children and adults. *Neuoreport* **16**, 1769–1773 (2005).

18. Huntley-Fenner, G. & Cannon, E. Preschoolers’ magnitude comparisons are mediated by a preverbal analog mechanism. *Psychol. Sci.* **11**, 147–152 (2000).

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

20. Nieder, A. & Miller, E. K. Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* **37**, 149–157 (2003).

21. Moyer, R. S. & Landauer, T. K. Time required for judgements of numerical inequality. *Nature* **215**, 1519–1520 (1967).

22. Parkman, J. M. Temporal aspects of digit and letter inequality judgments. *J. Exp. Psychol. Psychol.* **91**, 191–205 (1971).
23. 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).

24. Restle, F. Speed of adding and comparing numbers. *J. Exp. Psychol.* **83**, 274–278 (1970).

25. Feigenson, L., Dehaene, S. & Spelke, E. S. Core systems of number. *Trends Cogn. Sci.* **8**, 307–314 (2004).

26. Bonato, M., Zorzi, M. & Umiltà, C. When time is space: Evidence for a mental time line. *Neurosci. Biobehav. Rev.* **36**, 2257–2273 (2012).

27. 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).

28. Dehaene, S. & Cohen, L. Cultural recycling of cortical maps. *Neuron* **56**, 384–398 (2007).

29. Lakoff, G. & Núñez, R. E. *Where mathematics comes from.* (Basic Books, 2000).

30. 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).

31. 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, 2018).

32. De Smedt, B., Noël, M. P., Gilmore, C. & 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).

33. Sekuler, R. & Mierkiewicz, D. Children’s judgments of numerical inequality. *Child Dev.* **48**, 630 (1977).

34. 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).

35. 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).

36. 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, 2015).

37. Ashkenazi, S., Henik, A., Ifergane, G. & Shelef, I. Basic numerical processing in left intraparietal sulcus (IPS) acalculia. *Cortex** 44*, 439–448 (2008).

38. Mussolin, C., Mejias, S. & Noël, M.-P. Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition** 115*, 10–25 (2010).

39. 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).

40. 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).

41. 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).

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

43. 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).

44. Cipora, K. *et al.* Professional mathematicians differ from controls in their spatial-numerical associations. *Psychol. Res.* **80**, 710–726 (2016).

45. Sella, F., Sader, E., Lolliot, 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).

46. 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).

47. 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).

48. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).

49. Forster, K. I. & Forster, J. C. DMDX: A Windows display program with millisecond accuracy. *Behav. Res. Methods, Instruments, Comput.* **35**, 116–124 (2003).

50. Cipora, K. Between task consistency of the SNARC effect. in *32. European Workshop on Cognitive Neuropsychology* (2014).

51. R Core Team. *R: A language and environment for statistical computing.* (R Foundation for Statistical Computing, 2013).

52. 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.* 1–53 (2019). doi:10.31234/osf.io/bwyr3

53. Morey, R. D. *et al.* *BayesFactor.* (2018). ftp://alvarestech.com/pub/plan/R/web/packages/BayesFactor/BayesFactor.pdf

54. 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).

55. 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).

56. Göbel, S. Number processing and arithmetric 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, 2015).

57. 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, 2015).
58. 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, 2005).

59. 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).

60. Hardy, G. H. *A mathematician’s apology*. (Cambridge University Press, 2012).

61. Sriraman, B. The characteristics of mathematical creativity. *Math. Educ.* **14**, 19–34 (2004).

62. Amalric, M. & Dehaene, S. Origins of the brain networks for advanced mathematics in expert mathematicians. *Proc. Natl. Acad. Sci.* **113**, 4909–4917 (2016).

63. Pesenti, M. Calculation abilities in expert calculators. in *Handbook of mathematical cognition* (ed. Campbell, J. I. D.) 413–430 (Psychology Press, 2005).

64. Dowker, A. Computational estimation strategies of professional mathematicians. *J. Res. Math. Educ.* **23**, 45–55 (1992).

65. Dowker, A., Flood, A., Griffiths, H., Harriss, L. & Hook, L. Estimation strategies of four groups. *Math. Cogn.* **2**, 113–135 (1996).

66. 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).

67. 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).

68. 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).

69. Krajcsi, A. Numerical distance and size effects dissociate in Indo-Arabic number comparison. *Psychon. Bull. Rev.* **24**, 927–934 (2017).
Supplementary Materials to:

Professional mathematicians do not differ from others in analogue magnitude representation: Evidence from prevalence of numerical distance and size effects

Mateusz Hohol, Klaus Willmes, Edward Nęcka, Bartosz Brożek, Hans-Christoph Nuerk & Krzysztof Cipora

Supplementary Material 1: The SNARC effect

Analysis

The analysis reflects the approach reported in the main text. To quantify the SNARC effect, we used the individual regression slopes (as it was done in the case of the distance and the size effects described in the main text). Importantly, the SNARC effect in the magnitude classification task, contrary to the parity judgment task, is categorical rather than linear. Therefore, we quantified the SNARC effect as the individual regression slope, where dRTs (RT differences: right hand – left hand) are regressed on a magnitude contrast: numbers smaller than 5 were coded as -1 and numbers larger than 5, as +1. More negative slopes correspond to a stronger SNARC effect. Firstly, we estimated the reliability of the SNARC effect slopes; subsequently, using one-sample t-tests (two-sided) we tested for the presence of the SNARC effect at the whole sample level as well as in each of the four groups separately; we also compared groups by means of ANOVA as regarded the size of the slope. As in the main analysis, both frequentist and Bayesian analyses were used. In the next step, we tested for the individual prevalence of the SNARC effect using the H0 bootstrapping approach. We resampled RTs for each number (with replacement) and randomly allocated them into two subsets, which were subsequently treated as Right-hand and Left-hand responses. As in the analyses reported in the main text, 90% H0 CIs were used, and participants were classified based on whether they displayed a reliable SNARC effect. Ultimately, by means of a Fisher exact test, we tested for differences in the proportion of participants displaying a reliable SNARC effect, a reliable reversed SNARC effect, or no reliable SNARC effect.

Testing for the SNARC effect

The reliability of the SNARC slopes was .89. The SNARC slopes at the whole sample level as well as in each group separately are presented in Table S1.
Table S1. The SNARC effect at the whole sample level and in each group separately.

| Group | Mean (SD) | t-test* | Cohen's d | BF 10 | Proportion |
|-------|-----------|---------|-----------|-------|------------|
|       |           |         |           |       | Reliable SNARC | Reliable reversed SNARC | No reliable SNARC |
| Overall | -8.02 (25.3) | $t(97) = -3.14, p = .002^{**}$ | 0.32 | 10.80 | 44% | 16% | 40% |
| M     | 4.07 (26.3) | $t(12) = 0.56, p = .588$ | 0.16 | 0.32 | 31% | 23% | 46% |
| E     | -12.20 (24.20) | $t(13) = -1.88, p = .082$ | 0.51 | 1.08 | 50% | 21% | 29% |
| S     | -5.32 (25.1) | $t(14) = -0.82, p = .425$ | 0.21 | 0.35 | 47% | 20% | 33% |
| R     | -10.50 (25.1) | $t(55) = -3.13, p = .003$ | 0.42 | 11.0 | 45% | 13% | 43% |

Notes. * one sample t-test against zero (one sided); ** Significant results are marked with a bold font.

The SNARC effect was robust at the whole sample level as well as in the R group. On the other hand, the effect was not significant either in the M, E, or S group. In all the groups, Bayesian evidence was inconclusive and did not provide much support in favor of either the null or the alternative hypothesis. However, groups did not differ with respect to the SNARC effect as indicated by both frequentist and Bayesian evidence: $F(3, 94) = 1.37, p = .257, \eta^2 = .04, BF = 0.29$.

The results of the bootstrapping analysis are summarized in Figure S1. The distribution of SNARC effect slopes largely overlap between groups. Interestingly, only 43.9% of participants had a reliable SNARC effect. On the other hand, 16.3% of participants displayed a reliable reversed SNARC effect. Most importantly, 39.8% of participants did not have a reliable SNARC effect. The proportions of participants displaying reliable, reliable reversed, and no reliable SNARC effect did not differ between groups (Fisher exact test $p = .792$).
Figure S1. Summary of the SNARC effect results. Black dots represent the respective group means. Colored points (horizontally jittered) depict individual participants. Color represents whether the given participant reveals a reliable SNARC effect, a reliable reversed SNARC effect, or no reliable SNARC effect, as calculated using the bootstrapping method.
## Supplementary Material 2: Correlations

**Table S2.** Zero order correlations between measures used in the study.

| Measure   | Accuracy | RT  | SD(RT) | Distance | Size | SNARC |
|-----------|----------|-----|--------|----------|------|-------|
| Accuracy  | r        | —   | .29    | .06      | -.28 | -.01  |
|           | p        | —   | .004   | .579     | .006 | .937  |
|           | 95% CI Upper | —   | .46    | .25      | -.08 | .19   |
|           | 95% CI Lower | —   | .09    | -.14     | -.45 | -.21  |
| RT        | r        | —   | .90    | -.71     | .41  | .04   |
|           | p        | —   | <.001  | <.001    | <.001| .667  |
|           | 95% CI Upper | —   | .93    | -.60     | .56  | .24   |
|           | 95% CI Lower | —   | .86    | -.80     | .23  | -.16  |
| SD(RT)    | r        | —   | -.71   | .52      | -.01 |       |
|           | p        | —   | <.001  | <.001    | <.001| .943  |
|           | 95% CI Upper | —   | -.59   | .65      | .19  |       |
|           | 95% CI Lower | —   | -.79   | .36      | -.21 |       |
| Distance  | r        | —   | -.25   | -.05     |       |       |
|           | p        | —   | .012   | .602     |       |       |
|           | 95% CI Upper | —   | -.06   | .15      |       |       |
|           | 95% CI Lower | —   | -.43   | -.25     |       |       |
| Size      | r        | —   | -.15   |         |       |       |
|           | p        | —   | .146   |         |       |       |
|           | 95% CI Upper | —   | .05    |         |       |       |
|           | 95% CI Lower | —   | -.34   |         |       |       |
| SNARC     | r        | —   |        |         |       |       |
|           | p        | —   |        |         |       |       |
|           | 95% CI Upper | —   |        |         |       |       |
|           | 95% CI Lower | —   |        |         |       |       |

*Note:* Significant correlations are marked with bold font.

**Table S3.** Partial correlations between the NDE, NSE, and the SNARC effect controlled for mean reaction time.

| Measure   | Distance | Size | SNARC |
|-----------|----------|------|-------|
| Distance  | -        | 0.06 | -0.03 |
| Size      | -        |      | -0.18 |

*Note:* None of the partial correlations gets significant.
References

1. Dehaene, S., Bossini, S. & Giraux, P. The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* **122**, 371–396 (1993).

2. 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).

3. Wood, G., Willmes, K., Nuerk, H.-C. & Fischer, M. H. On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychol. Sci. Q.* **50**, 489–525 (2008).