Journal of Experimental Child Psychology 262 (2026) 106393

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Early development of spatial-numerical associations: consistency and variability of the SNARC effect in kindergarten children

Tânia Ramos *, Carrie Georges, Christine Schiltz

Institute of Cognitive Science and Assessment, Department of Behavioural and Cognitive Sciences, Faculty of Humanities, Education and Social Sciences, University of Luxembourg, Esch-Belval, Luxembourg

ARTICLE INFO

Keywords:
SNARC effect
Individual SNARC Consistency
Magnitude Judgment
Development
Kindergarten
Number-Space Associations

ABSTRACT

Numbers and space are associated in the human brain. One of the most-studied spatial-numerical associations (SNAs) is the SNARC effect (Spatial Numerical Association of Response Codes), for which robust group-level effects are reported across adult studies. Despite well-replicated group-level effects, recent individual-level analysis in adults indicate that only a minority of individuals consistently map numbers onto space (Cipora et al., 2019). To date, SNARC studies in children remain generally scarce with inconclusive results. And none have explored the consistency of individual effects at earlier developmental stages. In the present study, we therefore tested 135 kindergarten children performing magnitude judgments to assess not only group-level SNARC effects but also the prevalence of individual consistency using the same methodology recently applied in adults (Cipora et al., 2019). Our findings reveal a significant magnitude SNARC effect at the group-level. However, similarly to adults, only 37% of the children consistently associated numbers with space in a left-to-right direction when considering CIs around observed effects. While these findings suggest that SNAs on average emerge earlier in life, they also point towards considerable heterogeneity across individuals in that respect. How this can help us understand the conflicting results in the literature regarding significant group-level SNARC effects in children, and guide future research on the potential relation between individual SNARC effects and educational measures in math will be discussed.

Introduction

Numbers and space are intimately related in the human brain (Cutini et al., 2014; Wood et al., 2008; Hawes & Ansari, 2020). Behavioural evidence for this link is provided by the so-called SNARC effect (an acronym for the Spatial Numerical Association of Response Codes; Dehaene et al., 1990; Dehaene et al., 1993), characterised by the tendency of faster left/right responses to small/large numbers, respectively, in binary-choice reaction time (RT) tasks. While the SNARC effect is classically observed during parity or magnitude judgments on single Arabic digits using a bimanual response selection format, it has been documented across various number notations and sensory modalities (Nuerk et al., 2005), as well as with different response effectors (Fischer, 2003; Gevers et al., 2010; Hesse et al., 2016; Schwarz & Keus, 2004).

---

* Corresponding author at: University of Luxembourg, Faculty of Humanities, Social and Educational Sciences, Department of Behavioural and Cognitive Sciences, Institute of Cognitive Science and Assessment, CAMPUS BELVAL, Maison des Sciences Humaines11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg.

E-mail address: tania.ramos@uni.lu (T. Ramos).

https://doi.org/10.1016/j.jecp.2025.106393
Received 13 December 2024; Received in revised form 19 September 2025;

The SNARC effect is typically attributed to the spatial mental representation of numerical magnitude along a horizontally oriented axis in long-term memory, commonly referred to as the mental number line (MNL; Dehaene, 1997; Hubbard et al., 2005). However, alternative accounts suggest that the effect may arise from categorical verbal-spatial coding (the polarity coding account; Proctor & Cho, 2006) or the spatial encoding of numerical magnitude in working memory (van Dijck & Fias, 2011). Other perspectives propose that the SNARC effect emerges later in the stimulus–response process (Keus & Schwarz, 2005) and is tied to response rather than early stimulus representations (Gevers et al., 2006). Basso Moro and colleagues (2018; see also Nan et al., 2022) recently reconciled these perspectives, proposing that the SNARC effect is not a unitary phenomenon. Instead, its underlying cause is likely rooted between spatial-numerical representations and response-related processes, which are both thought to be influenced by a range of independent factors (see Cipora et al., 2020; Zhang et al., 2022). Specifically, the extent to which numerical magnitude is associated with space at the *representational stage* is thought to depend on several elements, which are selective attention, the abstractness and/or flexibility with which numerical information is processed, the nature of the spatial code (see above for different spatial accounts), embodied influences such as finger counting habits (Fischer, 2008), working memory availability (e.g., Guida et al., 2020; Herrera et al., 2008; van Dijck et al., 2009), and spatial skills (Viarouge et al., 2014). The manifestation of the SNARC effect at the *response selection stage* then depends on information-processing efficiency and cognitive control (Hoffmann et al., 2014; Zhang et al., 2022). The way these variables influence the size of the observed SNARC effect is believed to depend further on factors such as the task used to assess it (explicit versus implicit magnitude processing, e.g., Georges et al., 2017; Gevers et al., 2010; van Dijck et al., 2009), the numerical notation (Nuerk et al., 2004), the modality in which numbers are presented (visual versus auditory; (Hesse & Bremmer, 2017; see also, Nuerk et al., 2005), and the effector involved in response selection (e.g., finger, eyes, arm; Hesse et al., 2016).

SNARC effect in early childhood

While the aforementioned findings represent part of the extensive body of research on the SNARC effect accumulated over the past three decades, the majority of these studies have focused on adults. In contrast, research involving (younger) children remains relatively scarce and yields inconclusive evidence regarding the presence or absence of significant group-level effects. While group-level SNARC-like effects have been observed in younger children, the significance and strength of these spatial-numerical mappings vary with age and task demands. For instance, Zhang et al. (2024) reported significant group-level SNARC effects using a non-symbolic parity judgment task in 9-year-olds. Using a dot comparison task, Aulet and Lourenco (2018) as well as Ebersbach et al. (2014) also observed significant SNARC-like effects in slightly younger children aged 5–7 years. Patro and Haman (2012) even reported a group-level spatial–numerical congruity effect during a non-symbolic numerosity comparison task in children as young as 4 years. In a follow-up study, Patro et al. (2016) demonstrated that the direction of the effect in similarly aged children depended, however, on visuomotor activities, with reversed SNARC-like mappings following a right-to-left training paradigm. On the other hand, Chan and Wong (2016) found that while kindergarteners oriented small numbers to the left and large numbers to the right when processing the ordinal information of numbers in a month judgment task, no spatial mappings could be observed when processing their magnitude in a dot comparison task. Their findings thus suggest that left-to-right spatial biases at this age may be linked primarily to ordinal rather than magnitude properties of numbers. Collectively, these studies highlight the heterogeneity of SNARC-like effects in children and underscore the influence of developmental stage and/or task design on observed spatial-numerical associations. However, even among studies using comparable methodologies—specifically Arabic digits combined with a lateralized bimanual response setup as in the classical SNARC paradigm—group-level findings in children remain inconsistent. For instance, Hoffmann et al. (2013) used a digit colour as well as a magnitude judgment task to measure SNARC effects in Luxembourgish children aged 5.5 ($n = 36$) and 5.8 years old ($n = 48$). While group-level SNARC effects were observed in both age groups following numerical magnitude-irrelevant colour judgments, a tendency for significant magnitude SNARC effects was reported only in the older children's group. This developmentally early presence of the SNARC effect was corroborated by Yang et al. (2014) using parity judgements, documenting group-level SNARC effects in 44 Chinese children aged 5.8 years-old. Similarly, Cooney and colleagues (2021) observed significant SNARC effects using explicit magnitude judgments in 33 Irish children aged 5–7 years. Conversely, Chan and Wong (2016) failed to report group-level magnitude SNARC effects in 6-year-old Chinese-speaking children ($n = 66$). Similarly, no such effect was observed in the same age group by Gibson and Maurer (2016). These authors, however, reported significant effects in slightly older Canadian children aged 7 and 8 years. Likewise, a recent study by Jiang et al. (2024) assessed the SNARC effect in 6–7-, 7–8-, and 8–9-year-old Chinese children using both magnitude ($n = 164$) and parity judgement tasks ($n = 222$). Neither of the two effects could be significantly documented at the group-level in the youngest age group. They emerged, however, in older children, showing magnitude and parity SNARC effects from 7 to 8 and from 8 to 9-years-old, respectively.

In sum, previous studies suggest that young children already display SNARC effects before the start of formal school and math education. However, results are not consistent across studies, despite comparable methodologies and relatively similar participant samples—particularly with respect to age and nationality (e.g., Yang et al., 2014 versus Jiang et al., 2024). More work on the SNARC effect in young children is thus required to help elucidate the presence of the SNARC effect before formal schooling.

Measuring the SNARC effect at the group level and the individual level

Group-level SNARC effects are usually quantified using linear regression analysis on averaged data, where numerical magnitude is regressed on RT differences between right- and left-hand responses. The presence of the SNARC effect within any given sample is then assessed by testing the resulting linear regression slope against zero using a one-sample *t*-test (Fias et al., 1996; Lorch & Myers, 1990), with significantly negative regression coefficients reflecting reliable left-to-right spatial-numerical associations.

In experimental psychology, such group-level analyses are common practice and remain, without a doubt, a valuable contribution to the field. Their outcomes should, however, be interpreted with caution as important information stemming from inter-individual variation might simply be averaged out (see, e.g., Siegler, 1987). Findings drawn from averaged data should thus only be regarded as a starting point rather than a definitive conclusion (Rouder & Haaf, 2021). While it is commonly assumed in experimental psychology that the group-level presence of any effect provides evidence that the effect is equally present in every participant within a sample, this may not necessarily always be the case (e.g., Hohol et al., 2022; Cipora et al., 2019). Extrapolating from the group to the individual level thus seems inappropriate (Fisher et al., 2018; Robinson, 1950, for the ecological fallacy; see also discussion on the problem of generalising from group to person by McManus et al., 2023) as it might result in the loss of valuable information with regards to inter-individual variance in cognitive ability. The inherent variability between individuals should therefore be considered as data rather than noise to fully characterise the cognitive processes underlying human behaviour (Borsboom & Haslbeck, 2024; Sauce & Matzel, 2013; Vogel & Awh, 2008). Focusing on individual data to complement group data result will not only allow researchers to better understand group-level discrepancies in cognitive effects between studies depending on sample composition but also offer the possibility to relate these effects to individual brain structure and function. This seems especially relevant with regards to research on the SNARC effect in children, which is characterised by discrepancies across studies in terms of the group-level presence of the effect and aims to assess potential relations of SNAs with educational success, notably math skills.

As mentioned above, the SNARC effect is not a unitary construct (Basso Moro et al., 2018; see also Nan et al., 2022), and its behavioural manifestation in response time patterns is influenced by several independent factors at both the semantic representation and response selection stages (see Cipora et al., 2020; Zhang et al., 2022), in a task-, notation-, modality-, and effector-specific manner. Some of these variables may thus introduce noise into individual estimates—thereby affecting not only their observed strengths, but also the reliability with which they are measured. Inter-individual differences in the size of the effect and, critically, also in associated measurement noise could consequently impact the extent to which group-level effects are observed. Therefore, relying solely on observed SNARC regression slopes when quantifying the effect at the individual level, as is typically done, may not provide the most informative index, as it neglects potentially important measurement noise.

To quantify the SNARC effect at the individual level, (Cipora et al., 2019) have recently proposed a new methodology that calculates a confidence interval (CI) around the observed SNARC effect using a bootstrapping approach. Namely, a random sample of trials by any given participant is selected to calculate the SNARC effect using the aforementioned linear regression analysis. This procedure of random sampling (with replacement) is repeated many times to obtain a distribution of estimates for the effect. The range in which a certain proportion (e.g., 90 %) of these estimates is located is then taken to be the participant's CI. If the CI around the observed SNARC regression slope does not contain zero and the slope is negative, one may claim with predefined confidence that the participant reveals a consistent left-to-right SNARC effect. Alternatively, if the CI does not contain zero and the observed slope is positive, one may claim that the participant reveals a consistent right-to-left SNARC effect. Finally, if the CI contains zero, the participant should be classified as showing an inconsistent SNARC effect.

Cipora et al. (2019) applied this methodology to 18 adult datasets and observed that despite the robustness of the SNARC effect at the group-level across these studies, the proportion of participants consistently revealing the effect in a left-to-right manner only corresponded to 35 % to 45 % of individuals. 5 % to 10 % showed a consistent right-to-left SNARC effect with then almost more than half of the participants not showing any consistent spatial-numerical associations. These authors then concluded that despite robust group-level SNARC effects, only a minority of individuals consistently showed the effect. To date, to the best of our knowledge, this methodological approach has been limited to adult SNARC data (Cipora et al., 2019), with the prevalence of consistent SNARC effects never been studied across development.

The current study

Despite extensive research on the SNARC effect in adults over the past three decades, the existing literature concerning (younger) children remains limited and inconclusive. While the exact reasons behind inconsistent group-level findings in this population remain largely unclear, one possible explanation is that, since the SNARC effect depends on numerous variables (as previously discussed), these factors may introduce varying levels of measurement noise depending on individual traits. This variability at the individual level could, in turn, affect the overall strength of group-level effects, contributing to their presence or absence depending on the composition of the sample. To gain a clearer understanding of how such potential sample-related differences could account for discrepancies in group-level outcomes, further research on the SNARC effect is needed, particularly at the individual level. We therefore aimed to test a relatively large sample of 135 kindergarten children to: 1) further assess the significant presence of the magnitude SNARC effect at the group level, and 2) explore a new methodology that calculates CIs around observed individual SNARC estimates through bootstrapping (previously applied only to adult SNARC data, Cipora et al., 2019). This approach offers a more comprehensive analysis of the SNARC effect at the individual level than merely considering regression slopes. Namely, analysing CIs around observed SNARC effects can help account for noise in the data, thereby enabling the identification of individuals who exhibit consistent effects that are reliably stronger than the noise, as indicated by CIs not containing zero.

Given that knowledge about parity status is limited in kindergarten children, we administered a bimanual explicit numerical magnitude classification task. This response format is commonly used across SNARC studies in children (e.g., Chan & Wong, 2016; Cooney et al., 2021; Hoffmann et al., 2013), and maximises comparability of our results with adult studies, typically measuring SNARC effects with bimanual parity judgments.

With regards to a priori hypotheses, the present study is mostly explorative considering that (1) previous studies have yielded contradictive results with respect to significant group-level SNARC effects and (2) to the best of our knowledge, none has looked at individual SNARC effect consistency in younger children.

Although this study may not pinpoint the specific factors underlying individual noise and as such in/consistency in SNARC effects, we believe that incorporating the consistency of measurements, alongside the strength of the observed effect, when examining the SNARC effect at the individual level, could be an important first step in addressing discrepancies in group-level findings and guiding future research on the relation between individual SNARC effects and educational outcomes.

Method

Participants

A total of 173 children at the end of their final year in kindergarten participated in the present study. The participants were recruited from eight different public schools in Luxembourg. Prior to the start of the study, written informed consent was obtained from both parents and children gave their verbal assent to participate. Approval from the Ethics Review Panel of the University of Luxembourg was obtained.

From the initial sample, thirty-eight children were discarded from the analyses for the following reasons: 28 children did not complete all the testing sessions, 6 showed a lack of adequate understanding of the task (reflected in accuracy below 60 %), and 4 were not retained by the bootstrapping analysis due to empty response side x number cells. This resulted in a final sample of 135 children (74 girls and 61 boys).

Parents were invited to answer some questions regarding their socio-economic status (SES) and language background when giving their consent. More concretely, SES was measured by a parental questionnaire asking about the main professional occupation and qualification/skill level through 10 response alternatives (e.g., unemployed, agricultural worker, scientist, or related profession, etc.). Parents' occupational level was established according to the International Standard Classification of Occupations (ISCO-08) and the corresponding Socio-Economic Index of Occupational Status (ISEI, Ganzeboom et al., 1992; Ganzeboom, 2010). Both parents' responses were collected, however, for further analyses, a family index was computed by considering the highest value out of the two (Highest Socio-Economic Index, HISEI).

Language background was obtained through a questionnaire asking about which language each parent predominantly spoke with the child. One hundred and thirty-three children spoke a home language with reading and writing direction from left-to-right (e.g., Portuguese, French, Luxembourgish), and two children spoke a right-to-left (e.g., Arabic). Seven parents did not provide information regarding their SES and language background. All descriptive information can be found in Table 1.

Procedure

The current study was conducted in the context of a larger project comprising multiple tasks spread across different testing sessions. Concretely, the data used in this study were drawn from two sessions, which took place in June-July 2022 and October-November 2022, respectively. The interval between sessions depended on the teachers' and children's availability and was on average 3.13 months ($SD = 0.35$). The first of these sessions lasted approximately 35 min, and the second session lasted on average 10 min. Both sessions were conducted by trained research assistants in a quiet area near the children's classrooms. All children performed the tasks in the same fixed order as follows[1]: session 1: magnitude judgement task, children's mental transformation task (CMTT), symbolic ordinal judgment task, spatial language task and phonological awareness task; session 2: magnitude judgement task. The magnitude judgement task, the focus of the present study, will be described in more detail below.

Magnitude judgement task

The magnitude judgement task, adapted from Hoffmann and colleagues (2013) (see also van Galen & Reitsma, 2008), is a computerised task programmed in E-prime (Version 2.0.10.356) and administered using an Acer Spin 3 with a 14-inch colour monitor (1024 x 768 pixels) on a QWERTZ keyboard. This binary classification task required participants to judge whether a centrally presented digit between 1 and 9 (excluding 5) is smaller or larger than 5 by giving left- and right-sided manual responses. Successful task completion thus involves explicit access to numerical magnitude information. Each trial started with an empty, black-bordered square in the centre of a white screen. After one second, one of 8 potential stimuli (1, 2, 3, 4, 6, 7, 8, 9, in font Arial point size 48) appeared in the centre of the square and remained until response, or until five seconds had elapsed. This was followed by an empty screen intertrial interval of 1 s. The stimuli were presented in pseudo-random order, no number appeared twice in a row, and the correct response was

---

[1] For a detailed explanation of all the tasks administered in session 1, please see Georges et al., 2023.

Table 1
Children's descriptive information regarding their age and HISEI.

Variable	N	Mean	SD	Min.-Max.
Age (in years)	135	6.35	.32	5.69 – 7.12
HISEI	128	52.25	15.94	20.82 – 69.90

on the same side no more than two times consecutively. In the congruent mapping A, the left response key ("A") was associated with smaller numbers (1,2,3,4), and the right response key ("L") with larger numbers (6,7,8,9); in the incongruent mapping B, these associations were reversed. Each child completed both congruent (A) and incongruent (B) number-space mappings. In session 1, mapping A was followed by mapping B, while in session 2, mapping B was followed by mapping A, for every child.[2] Each number appeared a total of 20 times, divided equally across the task (i.e., 10 times per session x 2 response sides). Given the age of the children, response keys were highlighted with two coloured dots placed on the "A" and "L" buttons, in addition to drawings of a mouse and elephant placed above the response keys to indicate "smaller" and "larger", respectively.

Before starting the experiment, children were positioned centrally in front of the screen and instructions were provided Luxembourgish. Each block started with 8 training trials, of which at least 80 % had to be correct; otherwise, an additional verbal explanation was provided, followed by a new set of 8 training trials. A maximum of two practice rounds per mapping/session was carried out, with the number of training trials thus ranging between 16 and 32 per session, depending on the child's performance. Feedback was only provided during the training using happy or sad smiley faces, depending on accuracy. In total, 160 experimental trials were completed, divided equally across 4 mappings (i.e., 40 trials per A/B mapping, 80 trials per session).

Statistical analysis

Analysis was conducted on RT data for which only correct experimental trials were retained. In a preliminary step, RT data was filtered first for anticipations and then for outliers. More concretely, RTs below 200 ms were considered as anticipations and discarded (for comparability reasons with previous studies in children, e.g., Cooney et al., 2021; He et al., 2021, and adults, e.g., Cipora et al., 2019). After their removal, a sequential filter was applied that discarded individual RTs more or less than 3SD from the individual mean RT until no more changes in mean RT occurred or for a maximum of 20 repetitions (please see R-Script in: https://osf.io/fn36p, provided by Cipora et al., 2018).

To analyse the presence of the SNARC effect at the *group level*, the average of the RTs was computed for each digit and response side (left and right side), separately for each child, then a subtraction between right-handed RTs and left-handed RTs was calculated for each digit to retrieve the individual differences in reaction time (dRT). Afterwards, these dRTs were entered in a linear regression as a dependent variable and the digits as a predictor, as suggested by Lorch and Myers (1990) and recommended by Fias and colleagues (1996). To determine whether the regression weight of the group deviated statistically from zero, a one-sample *t*-test against 0 was performed. A significantly negative SNARC regression slope reveals the presence of number-space associations in the expected direction, corresponding to faster RTs for small/large digits on the left/right side, respectively. Additionally, internal consistency was calculated through the "split-half" method, in which valid trials were divided into two parts based on the order of appearance with the "odd–even" method. Then, for each part (odd and even) and participant, regression slopes were calculated and subsequently correlated with each other. The higher the correlation between these two parts (odd and even), the greater the reliability of the task. The Spearman-Brown correction was applied to adjust for task length.

To evaluate the consistency of the SNARC effect at the individual level, we applied a bootstrapping approach following the procedure outlined by Cipora et al. (2019; R script available at https://osf.io/d7w2k). This technique involves generating a confidence interval (CI) around each participant's observed SNARC regression slope. The rationale behind this method is that each observed measurement consists of a true underlying effect combined with random measurement error.[3] By resampling the data with replacement, bootstrapping provides an estimate of the variability in the slope and allows us to determine the range within which the true effect is likely to fall. More precisely, in this technique, for each participant and experimental cell (defined by number × response hand), a random sample of valid trials was selected with replacement. The number of trials drawn was equal to the total number of repetitions for each digit across the experiment—in our case, 20 trials per digit (10 with a left-hand response and 10 with a right-hand

---

[2] As the extent to which mapping order introduces noise in the measurement of spatial-numerical associations (SNAs) remains unclear, we implemented a fixed block order across participants to ensure that any potential order effects did not differentially impact individuals and thus could not account for inter-individual differences in SNARC effect consistency. Using a fixed order is standard practice and recommended in individual differences research (Carlson & Moses, 2001; see also Georges et al., 2016, for not counterbalancing in SNARC studies). To reduce bias toward a specific order and provide an unbiased estimate of individual and group-level SNARC effects, each participant completed both orders (A/B and B/A), and results were averaged across them. The splitting of the two orders across sessions was due to a procedural error; originally, both orders were intended to be administered within the same session. Intra-individual stability between sessions is reported in Appendix D.

[3] This random measurement error refers to trial-to-trial variability in responses due to non-systematic influences (i.e., noise). It differs from the residual error term in the linear regression used to calculate the SNARC slope, which captures how closely the observed data points fit the predicted regression line. While regression residuals pertain to model fit, the measurement error addressed through bootstrapping relates to the stability of the SNARC effect across repeated samples.

response). This resampling process was repeated 5000 times. For each bootstrap sample, the SNARC effect was calculated using the linear regression method described above, producing a distribution of slope estimates. From this distribution, an 80 % CI was computed based on previous studies using the same approach in young children (van Dijck et al., 2020). If the resulting CI does not contain zero, the SNARC effect is considered consistently present for that individual. In contrast, if the CI includes zero, the presence of the effect is not consistent, suggesting that it may be absent or unreliable. From this, the following three consistency groups were thus deduced: inconsistent mapping behaviour, consistent left-to-right mapping behaviour (i.e., negative CI not containing zero), and consistent right-to-left mapping behaviour (i.e., positive CI not containing zero).[4]

Two separate ANOVAs were conducted to examine whether consistency group influenced (1) the size of the SNARC regression slope and (2) the width of the bootstrap CI. Due to significant differences in group variances and the violation of the homogeneity of variance assumption, Welch's ANOVA was used. Post-hoc comparisons were performed using the Games-Howell procedure.

All analyses were conducted in JAMOVI (version 2.3.21) and R software.

Results

Preliminary analyses

The children considered for analysis ($n = 135$) responded on average correctly on 147.27 out of 160 trials (i.e., 92.04 %), with a mean RT of 1621 ms ($SD = 502$ ms). From those trials, on average, 0.34 trials (i.e., 0.23 %) were removed as anticipations (i.e., RTs below 200 ms), leading to a mean RT of 1626 ms ($SD = 506$ ms) after this removal. Subsequently, an additional average of 9.03 trials (i.e., 6.15 %) were discarded as outliers by the sequential filter. Analyses were thus based on an average of 137.90 trials ($SD = 14.89$) with a mean RT of 1392 ms ($SD = 352$ ms).

There was no evidence of a speed-accuracy trade-off, as accuracy and RT were negatively correlated ($r = -.21$, $p = .014$). This indicates that faster responses (i.e., lower RTs) were associated with higher accuracy, suggesting that both measures tap into the same underlying construct. Therefore, it is reasonable to focus on RT alone as an indicator of performance in magnitude tasks, as is classically done in SNARC studies.

The Spearman-Brown corrected split-half reliability estimate based on the odd–even method was.77.

Group-level SNARC effect

Number-space associations were found at the group level, with a mean regression slope of −19.21 ($SD = 58.1$). A one-tailed t-test confirmed the presence of a significant SNARC effect, revealing a significantly negative dRT slope regression weight, $t\,(134) = -3.84$, $p = . < 001$ (see Fig. 1). Complementary bootstrap analyses, yielding an 80 % CI [-25.58, −13.05], based on 5000 resamples, confirmed the robustness of the negative group-level effect.

Individual-level SNARC effect distribution

Descriptively, 68.9 % of the children displayed a negative SNARC regression slope (n = 93), while the slope was positive in 31.1 % (n = 42). Analysis of the distribution of individual SNARC regression slopes (see Fig. 2) revealed an approximately normal shape, exhibiting mild leptokurtosis (kurtosis = 3.86; $SE = .414$) and slight left skewness (skewness = -.088; $SE = .209$).

Individual-level SNARC effect consistency

The bootstrapping approach based on the 80 % CI revealed that in 56 % of the children (75 out of 135), the SNARC effect was consistently observed (i.e., CI did not contain zero). Specifically, 37 % (49 children) consistently mapped numbers from left-to-right, while 19 % (26 children) showed a consistent right-to-left mapping. The remaining 44 % had inconsistent behaviour (see Fig. 3). Percentages for each consistency group based on alternative CI thresholds are provided in Appendix A. Welch's ANOVA revealed a significant effect of consistency group on the SNARC regression slope, $F\,(2, 57.4) = 101$, $p = <.001$. Games-Howell post hoc test indicated that all groups differed significantly. More precisely, children with consistent left-to-right mapping behaviour ($M = -72.60$, $SD = 49.98$) revealed stronger SNARC effects (i.e., more negative regression slopes) than children with consistent right-to-left mapping behaviour ($M = 53.85$, $SD = 28.81$, $p = <.001$) and inconsistent behaviour ($M = -6.16$, $SD = 19.45$, $p = <.001$). The slope of inconsistent mappers also differed significantly from those with consistent right-to-left behaviour ($p = <.001$). The distribution of SNARC regression slopes by consistency group is available in Appendix B. In contrast, no significant differences were observed between consistency groups in terms of the width of the CI, $F = (2, 61.3) = .120$, $p = 0.887$ (consistent left-to-right mappers: $M = 70.5$, $SD = 41.0$, consistent right-to-left mappers: $M = 69.8$, $SD = 31.6$, inconsistent mappers: $M = 65.6$, $SD = 43.9$). This suggests that individual variability in mapping behaviour was comparable across the different bootstrapping classifications. Correlations between the widths of

---

[4] Using an 80% confidence interval as the criterion for classifying consistent SNARC effects is illustrative rather than definitive. This threshold reflects the level of measurement noise that we were willing to consider for the purposes of the present analysis. However, we acknowledge that applying different certainty levels (e.g., a 75% CI) would naturally lead to changes in the proportion of individuals classified as e.g., consistent left-to-right mappers (see also Appendix A).

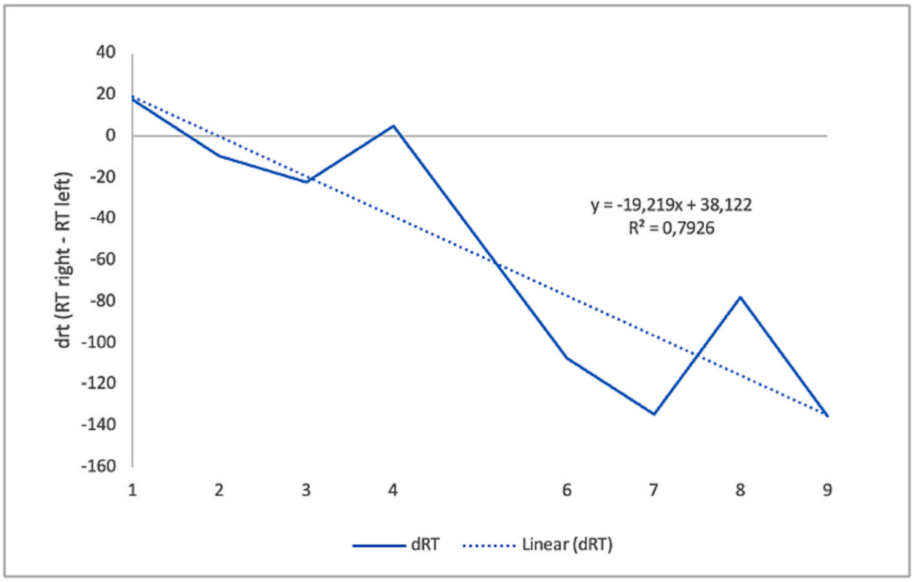

Fig. 1. Regression analysis of the difference between right- and left-hand reaction times (dRT) using number as a predictor variable. The resulting linear regression slope is significantly negative; hence, the SNARC effect was present at the group level.

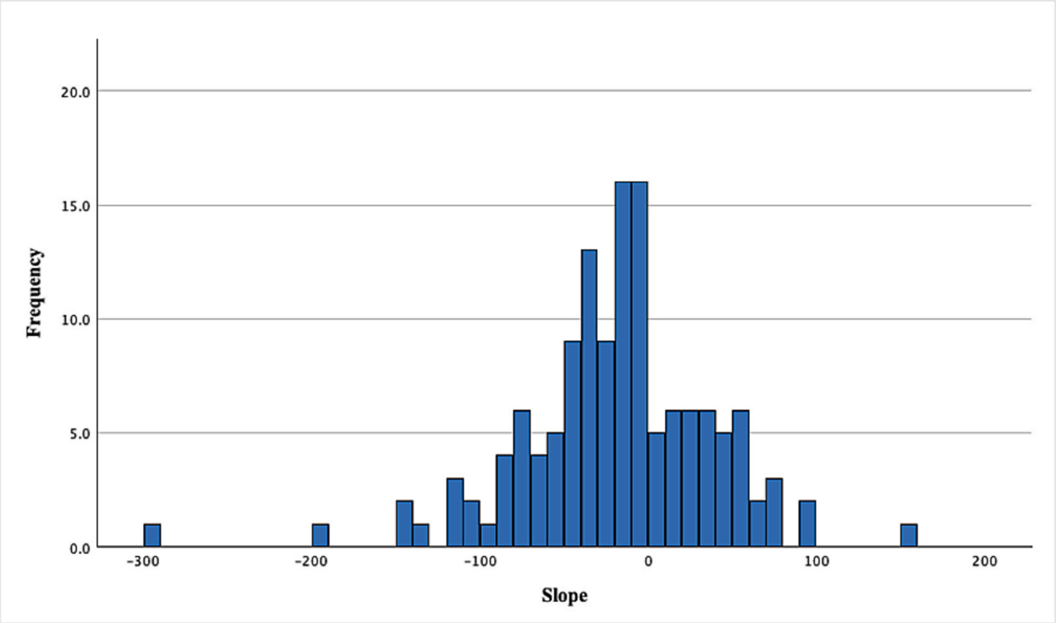

Fig. 2. Distribution of individual SNARC effects. The histogram displays the distribution of the participants' SNARC regression slopes, with the frequency (number of participants) on the vertical axis and the bin of the regression slope values on the horizontal axis.

the CI and absolute SNARC regression slopes for each consistency group can be found in Appendix C.

Discussion

Mapping numbers to space is a key feature of learning and doing mathematics (e.g., Cutini et al., 2014; Wood et al., 2008; Hawes & Ansari, 2020). Amongst the large variety of SNAs supporting the existence of a spatial number representation (for a review and taxonomy proposal, see, e.g., Cipora et al., 2018), the SNARC effect is the most well-known and studied.

Despite this, the number of studies on the SNARC effect in children is still limited and the few available studies show inconsistent SNARC effects at the group-level. Therefore, the present study examined a large group of kindergarten children to assess the presence

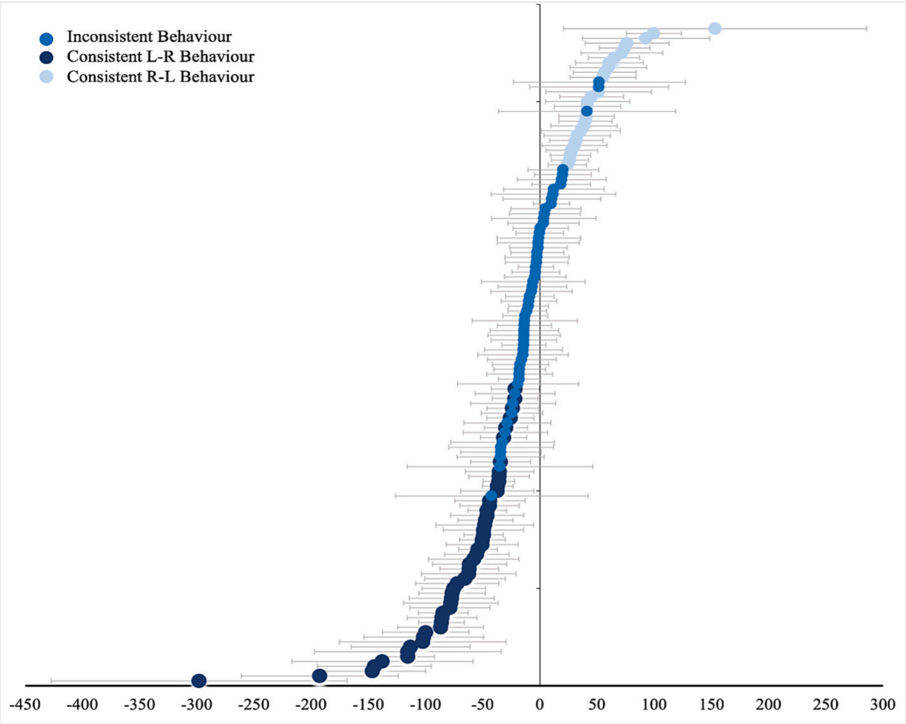

Fig. 3. Bootstrapping results for individual SNARC consistency. Each dot represents a single participant, with grey lines indicating the 80% confidence interval (CI) around their mean slope. Dark blue denotes consistent left-to-right mappings, light blue indicates consistent right-to-left mappings, and medium blue reflects inconsistent behaviour. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the magnitude SNARC effect 1) at the group level and 2) at the individual level using a bootstrapping method, previously applied only to adult SNARC data (Cipora, Van Dijck, et al., 2019). Analysing CIs around observed individual estimates accounts for noise in their measurements. It thereby allows to identify those participants who show consistent SNARC effects reliably exceeding noise, as indicated by CIs not containing zero. The relative proportions of consistent versus inconsistent mappers may influence the overall significance of group-level SNARC effects, as a relatively high prevalence of consistent left-to-right mappers could be sufficient to drive such effects. Nonetheless, this is not strictly necessary, since even when inconsistent participants are more numerous, a significant group-level effect may still emerge if many of them display negative SNARC regression slopes. Therefore, considering the consistency of individual SNARC measurements in addition to their directional strengths could serve as an important first step towards better understanding discrepancies in group-level findings. Furthermore, taking into account individual measurement uncertainty should inform future research on the relation between individual SNARC effects and educational outcomes, particularly math skills.

The magnitude SNARC effect was observed at the group level in young children aged on average 6.35 ± 0.32 years. These results suggest that on average SNAs, as indexed by the magnitude SNARC effect, are already present in relatively young children attending kindergarten. This finding is in line with previous studies observing SNAs in infants (Bulf et al., 2016; de Hevia et al., 2014; de Hevia et al., 2017), toddlers (McCrink et al., 2017), and preschoolers (de Hevia & Spelke, 2009; Opfer et al., 2010; Patro & Haman, 2012; Shaki et al., 2012). It also further confirms the relatively early presence of SNARC effects reported by several authors (Cooney et al., 2021; Hoffmann et al., 2013; Yang et al., 2014). Conversely, Chan and Wong (2016) and Gibson and Maurer (2016), as well as Jiang et al. (2024) did not observe significant group-level magnitude SNARC effects in similar age groups.

It remains poorly understood why some studies, including this one, observed SNARC effects in children aged 6–8 years, while others failed to report significant group-level effects. Given that the SNARC effect is not a unitary construct (Basso Moro et al., 2018; see also Nan et al., 2022), and that its behavioural manifestation in response time patterns is influenced by several independent factors at both the semantic representation and response selection stages (see Cipora et al., 2020; Zhang et al., 2022) it is difficult to measure at the individual level. Moreover, the representational and response-related stages are influenced in a task-, notation-, modality-, and effector-specific manner, which introduces varying levels of noise into individual estimates and thereby affects their reliability and observed strength. Inter-individual differences in the size of the effect and, critically, also in associated measurement noise could then, more generally, impact the extent to which group-level effects are observed.

In a second step, we therefore investigated the presence of individual SNARC effects and their consistency in the present kindergarten children using the same bootstrapping approach recently described in adults (Cipora et al., 2019). To the best of our knowledge, no one has yet applied this methodology to SNARC data in children (see van Dijck et al., 2020, for an application of this method in children concerning the ordinal position effect).

Individual-level analysis showed that the significant group-level SNARC effect was driven by approximately 70 % of children (n = 93) showing negative SNARC regression slopes. Of those, only slightly more than half (n = 49) exhibited consistent left-to-right effects, as indicated by CIs around the regression slopes that did not contain zero. Overall, 37 % of the children consistently associated numbers with space in a left-to-right direction during explicit magnitude judgments. The remaining children either showed a consistent right-to-left organisation (i.e., 19 %) or did not consistently associate numbers with space during magnitude judgments, based on the bootstrapping outcomes (i.e., CIs around observed effects contained zero). While consistent mappers significantly differed from inconsistent ones in terms of observed regression slopes—showing stronger negative and positive effects for consistent left-to-right and right-to-left mappers, respectively—the level of noise was, on average, comparable across consistency groups.

Like in adults, the majority of children consistently mapping numbers to space did so in a left-to-right direction. One potential explanation for the relatively higher prevalence of consistent left-to-right over right-to-left mapping, aligning with the direction of Western scripts, is the 'implicit instruction account' proposed by Patro and colleagues (2016). Accordingly, early preferences for a spatial arrangement of numbers or quantities arises as an effect of implicit learning about culture-specific exploration that children receive from observation and interaction with parents or teachers (McCrink et al., 2018). For instance, during shared readings, the implicit directional cues perceived during this procedure may activate or consolidate numerical-spatial representations according to the direction of reading, even before the language is mastered as a script (Göbel et al., 2018; Shaki et al., 2012). Visual-motor activities (e.g., manual and tablet games), scanning activities, and gestures have also been demonstrated to contribute to forming a specific mapping between numbers and space (Patro et al., 2016; Shaki et al., 2012). Children's experiences within their social and cultural context may thus result in the internalisation of a cultural script, which guides the children's actions in accordance with the norms and expectations of their culture.

The current proportion of consistent left-to-right mappers was also similar to that reported in adults (Cipora, Van Dijck, et al., 2019). Nonetheless, a higher percentage of children (i.e., 19 %) than adults (i.e., 5–10 %) exhibited consistent right-to-left SNARC effects in the magnitude judgment task. This finding is not surprising, as a study by McCrink and colleagues (2017) on Westernised toddlers found that, in a counting task, only 54 % of participants consistently counted from left-to-right, while 31 % preferred right-to-left counting, and 11 % displayed non-directional or random counting patterns. Although linguistic background has been suggested to influence the direction of number-space mappings in adults (Shaki & Fischer, 2008; Shaki et al., 2009) and preschoolers (Shaki et al., 2012), script direction itself cannot explain the higher proportion of consistent right-to-left mappers in the current sample compared to adults, as most children in this study were exposed to languages using a Western script. It is possible, therefore that, given toddlerhood is still a period of cognitive flexibility, children may not yet have fully internalised the norms of their social environment, leading to consistent number-space associations that deviate from expected patterns.

Future research should explore how the relative proportions of consistent left-to-right versus right-to-left mappers evolve over development. If schooling and the acquisition of reading/writing skills contribute to directional preferences in SNAs according to the cultural norms, one might expect to observe fewer right-to-left consistent mappers in first grade and beyond. This assumption finds support in the relatively fewer individuals with consistent right-to-left mappings in previous adult studies (Cipora et al., 2019). However, the direction of SNAs may be less relevant when exploring their potential links to educational outcomes like math abilities (see Treccani & Umiltà, 2011), especially in younger children. In fact, Fischer and colleagues (2010) observed situational changes in the direction of the SNARC effect in adults. Furthermore, Sella et al. (2019) found that, in preschool children, spatial-numerical mappings—regardless of direction—were associated with performance on a digit comparison task.

Interestingly, although children with consistent SNARC effects in the culturally typical left-to-right direction showed, on average, stronger slopes and comparable levels of noise relative to inconsistent mappers, a closer examination reveals that this group is far from homogeneous. Aside from consistency in measurement, no two left-to-right consistent mappers are identical in terms of the strength of their observed regression slope and the uncertainty surrounding that estimate (i.e., the width of the CI). Notably, we observed scalar variance among those mappers, with more negative observed effects tending to be associated with larger CIs (see Appendix C). Consistency in behaviour, therefore, resulted from either strong effects despite considerable noise, or minimal noise despite weaker effects. What ultimately mattered in that group was the relative strength of the SNA in relation to the measurement noise. Accordingly, it is essential to consider both the size of the observed effect and the level of measurement noise when interpreting individual consistency.

In line with previous findings in adults, a substantial proportion of children did not exhibit consistent number–space mappings, as indicated by CIs around their SNARC effects that included zero. Nonetheless, some of these children displayed negative observed estimates and may therefore have contributed to the overall significance of the group-level effect. While it is possible that some children exhibiting inconsistent behaviour simply do not yet associate numbers with space in the current setup, this is likely not the sole explanation for all instances of inconsistent behaviour. This is especially true considering that the mean of observed SNARC slopes across inconsistent mappers is negative, with many slopes overlapping with those of consistent mappers in the negative range near zero (see also Appendix B). Consequently, rather than the complete absence of SNAs, inconsistency in their behaviour might be explained by excessive noise interfering with the measurement of (still) relatively weak mappings. Some of these noise-inducing factors might even have had a systematic and strong impact, as seen in the significant deviation of RTs for numbers 4 and 8 from the general linear trend.

This suggests that the SNAs of some children currently classified as inconsistent mappers—particularly those with negative observed estimates intermingled among consistent mappers in the slope distribution—may, in fact, resemble those of a subset of consistent left-to-right mappers who displayed relatively weak but reliable SNARC effects. The key difference here lies in the level of measurement precision: weaker slopes can be associated with consistent behaviour if noise is low, whereas similar slopes may appear inconsistent if noise is higher. In other words, some inconsistent mappers may exhibit directional effects similar to those of consistent mappers, but greater variability explains why their behaviour is classified as inconsistent. As these children grow older and their measurement precision improves (as a result of reduced noise), some may be reclassified as consistent simply because their CIs become narrower and consequently no longer contain zero (see also Appendix A for how the prevalence of consistent left-to-right mappers varies with the certainty level applied). However, it should be noted that the overall percentage of consistent left-to-right mappers appears to remain relatively stable over the lifespan, as evidenced by the comparable proportions of adults and children exhibiting consistent left-to-right SNAs (i.e., ±40 %). The assumption that consistent and inconsistent mappers may not differ categorically in their SNAs is supported by the approximately normal distribution of SNARC regression slopes[5] (see Fig. 2), which cluster around a single peak slightly left of zero, a pattern also observed in adults (Cipora et al., 2019, see Fig. 2; see also Bulut et al., 2025) with the majority of inconsistent mappers displaying negative SNARC effects (Cipora et al., 2019, see Appendix E). If distinct SNA categories existed, a bi- or trimodal distribution would be expected. Moreover, the substantial overlap in the distribution of SNARC effects across consistency groups further suggests that group distinctions are likely continuous rather than discrete (see Appendix B). Consequently, while the current approach allows us to estimate the prevalence of individuals who consistently exhibit SNAs, it does not permit to conclude that consistent and inconsistent mappers differ categorically in terms of their SNAs.

Given the relatively high proportion of inconsistent mappers in the current study (and among adults, see Cipora et al., 2019), future research should focus on identifying the factors that contribute to high measurement noise, particularly when it co-occurs with weak observed estimates. Moreover, it will be important to understand why these factors affect some individuals more than others, and whether this variability reflects genuine qualitative differences in SNAs or merely differences in measurement reliability.

Amongst the factors generally contributing to noise in measurement could be the relatively high executive load entailed by the current bimanual explicit magnitude judgment task, particularly in young children. Namely, task instructions with respect to the stimulus–response mapping switched halfway through the experiment. In addition, the bimanual response set-up required individuals to constantly override preferences for the dominant, mostly right, hand. These factors might have unsystematically interfered with response-related RTs, thereby distorting outcomes. It therefore cannot be excluded that the present findings underestimated the prevalence of kindergarten children displaying consistent SNAs. Future research in children should therefore administer tasks involving unimanual response selection to determine how this changes variance in RTs and as a result overall consistency in SNAs across participants. In this context, one could also envisage the use of the colour judgment task, proposed by Hoffmann et al. (2013). Despite its bimanual response paradigm, it does not require explicit numerical magnitude processing, similarly to the parity judgment task usually administered in adult studies. Task instructions also do not change throughout the experiment, thereby minimising executive load. If proportions of consistent versus inconsistent mappers remain similar across different SNA tasks, this would indicate that task-related factors were not primarily responsible for the high prevalence of inconsistent behaviour.

Another possibility that more specifically accounts for inconsistent behaviour is that inconsistent mappers might associate numbers with space in (still) a relatively more flexible way, thereby not leading to directionally consistent SNARC effects. Inconsistent mappers might, for instance, create a new mapping each time, without a preferred direction or with only a relatively weak left-to-right preference or use a different, non-horizontal, spatial reference frame (e.g., vertical; Cooney et al., 2021). Such directional switches might occur either at the representational or response selection stage or even both. Trial-by-trial flexibility in spatial-numerical mappings might decrease with age and experience, notably also the acquisition of reading/writing skills based on Western script direction. On the other hand, this flexible behaviour might pertain into adulthood (see also Fischer et al., 2010; Treccani & Umiltà, 2011), thereby potentially also, at least to some extent, underlying the relatively high proportion of inconsistent mappers previously reported in adults (Cipora et al., 2019). As the current psychometric approach does not allow for such distinctions to be made, future research should focus on identifying the factors underlying inconsistent behaviour.

Finally, future studies might also examine the effect of mapping order on individual SNARC effect consistency, as previous work cannot rule out its influence on the strength of group-level effects (see e.g., Cipora et al., 2019) and, consequently, on individual consistency. In the present study, administering both orders to every participant and averaging across them provided an estimate of individual and group-level SNARC effects avoiding bias toward a specific order. However, using a fixed block procedure without counterbalancing across participants prevented direct assessment of between-subject order effects. Moreover, because the two orders were administered across separate sessions, any potential within-subject order effects are confounded with session and cannot be unequivocally interpreted. A related analysis comparing participant classifications between sessions is reported in Appendix D, showing relatively low intra-individual stability in SNARC effect consistency.

Conclusion

In summary, our study offers valuable new insights into the SNARC effect in young children. We found a significant group-level magnitude SNARC effect, with approximately 70 % of children having negative regression slopes, indicating left-to-right SNAs. However, when considering CIs around observed effects, only 37 % of the children consistently associated numbers with space in a left-to-right direction. This prevalence aligns with findings previously reported in adult studies. While the present outcomes suggest that on average SNAs emerge early in life before formal schooling, they also point towards considerable heterogeneity across individuals both in terms of observed estimates and noise surrounding these measurements. Most importantly, given the numerous previous efforts to better understand SNAs and their potential role in math development, the current outcomes underscore the need to incorporate CIs around observed effects as a crucial first step toward: (1) improving our understanding of the conflicting results in the literature

---

[5] Our categorisation is intended as a heuristic tool for exploring variability, not as evidence of stable traits. Given the unimodal distribution, we stress that the approach is a methodological exploration of heterogeneous cognitive effects, not a basis for clear-cut participant categories.

regarding significant group-level SNARC effects, especially in children, and (2) guiding future research on the relation between individual SNARC effects and educational measures, notably math skills.

Declaration of generative AI and AI-assisted technologies in the writing process.

During the preparation of this work, the author(s) used GPT-4o from OpenAI (2024) in order to improve language and readability. After using this tool, the author (s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

CRediT authorship contribution statement

Tânia Ramos: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Carrie Georges:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Christine Schiltz:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We thank the reviewers for their valuable feedback, which has substantially contributed to the quality of this manuscript. We also extend our gratitude to Nicolas Masson for his insightful comments on the methodological aspects of the study. This work was supported by the Luxembourg National Research Fund (FNR) [INTER/FWO/20/14716314 – SpaNuMaDev]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2025.106393.

Data availability

Data will be made available on request.

References

Aulet, L. S., & Lourenco, S. F. (2018). The developing mental number line: does its directionality relate to 5- to 7-year-old children's mathematical abilities? *Frontiers in Psychology, 9*, 1142. https://doi.org/10.3389/fpsyg.2018.01142

Basso Moro, S., Dell'Acqua, R., & Cutini, S. (2018). The SNARC effect is not a unitary phenomenon. *Psychonomic Bulletin & Review, 25*(2), 688–695. https://doi.org/10.3758/s13423-017-1408-3

Borsboom, D., & Haslbeck, J. (2024). Integrating Intra- and Interindividual Phenomena in Psychological Theories. *Multivariate Behavioral Research.* https://doi.org/10.1080/00273171.2024.2336178

Bulf, H., De Hevia, M. D., & Macchi Cassia, V. (2016). Small on the left, large on the right: Numbers orient visual attention onto space in preverbal infants. *Developmental Science, 19*(3), 394–401. https://doi.org/10.1111/desc.12315

Bulut, M., Roth, L., Bahreini, N., Cipora, K., Reips, U. D., & Nuerk, H.-C. (2025). One direction? Cultural aspects of the mental number line beyond reading direction. *Psychological Research Psychologische Forschung, 89*(1), 37. https://doi.org/10.1007/s00426-024-02038-4

Carlson, S. M., & Moses, L. J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Children Development, 72*(4), 1032–1053. https://doi.org/10.1111/1467-8624.00333

Chan, W. W. L., & Wong, T.-T.-Y. (2016). The underlying number–space mapping among kindergarteners and its relation with early numerical abilities. *Journal of Experimental Child Psychology, 148*, 35–50. https://doi.org/10.1016/j.jecp.2016.03.010

Cipora, K., He, Y., & Nuerk, H. (2020). The spatial–numerical association of response codes effect and math skills: Why related? *Annals of the New York Academy of Sciences, 1477*(1), 5–19. https://doi.org/10.1111/nyas.14355

Cipora, K., Masson, N., van Dijck, J.-P., Nuerk, H.-C., Georges, C., Pesenti, M., Goebel, S., Willmes, K., & Schiltz, C. (2018). *The Prevelance of the SNARC effect.*. https://doi.org/10.17605/OSF.IO/N7SZG

Cipora, K., Schroeder, P. A., Soltanlou, M., & Nuerk, H.-C. (2018). More Space, Better Mathematics: Is Space a Powerful Tool or a Cornerstone for Understanding Arithmetic? In K. S. Mix & M. T. Battista (Eds.), *Visualizing Mathematics* (pp. 77–116). Springer International Publishing. doi: 10.1007/978-3-319-98767-5_4.

Cipora, K., Soltanlou, M., Reips, U.-D., & Nuerk, H.-C. (2019). The SNARC and MARC effects measured online: Large-scale assessment methods in flexible cognitive effects. *Behavior Research Methods, 51*(4), 1676–1692. https://doi.org/10.3758/s13428-019-01213-5

Cipora, K., Van Dijck, J.-P., Georges, C., Masson, N., Goebel, S. M., Willmes, K., Pesenti, M., Schiltz, C., & Nuerk, H.-C. (2019). *A Minority pulls the sample mean: On the individual prevalence of robust group-level cognitive phenomena – the instance of the SNARC effect.* doi: 10.31234/osf.io/bwyr3.

Cooney, S. M., Holmes, C. A., & Newell, F. N. (2021). Children's spatial–numerical associations on horizontal, vertical, and sagittal axes. *Journal of Experimental Child Psychology, 209*, Article 105169. https://doi.org/10.1016/j.jecp.2021.105169

Cutini, S., Scarpa, F., Scatturin, P., Dell'Acqua, R., & Zorzi, M. (2014). Number–Space Interactions in the Human Parietal Cortex: Enlightening the SNARC effect with Functional Near-Infrared Spectroscopy. *Cerebral Cortex, 24*(2), 444–451. https://doi.org/10.1093/cercor/bhs321

De Hevia, M. D., Addabbo, M., Nava, E., Croci, E., Girelli, L., & Macchi Cassia, V. (2017). Infants' detection of increasing numerical order comes before detection of decreasing number. *Cognition, 158*, 177–188. https://doi.org/10.1016/j.cognition.2016.10.022

De Hevia, M. D., Girelli, L., Addabbo, M., & Macchi Cassia, V. (2014). Human infants' Preference for Left-to-right Oriented increasing Numerical Sequences. *PLoS One1, 9*(5), Article e96412. https://doi.org/10.1371/journal.pone.0096412

De Hevia, M.-D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition, 110*(2), 198–207. https://doi.org/10.1016/j.cognition.2008.11.003

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

Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General, 122*(3), 371–396.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). *Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison*.

Ebersbach, M., Luwel, K., & Verschaffel, L. (2014). Further evidence for a Spatial-Numerical Association in Children before Formal Schooling. *Experimental Psychology, 61*(4), 323–329. https://doi.org/10.1027/1618-3169/a000250

Fias, W., Brysbaert, M., Geypens, F., & d'Ydewalle, G. (1996). The Importance of Magnitude Information in Numerical Processing: Evidence from the SNARC effect. *Mathematical Cognition, 2*(1), 95–110. https://doi.org/10.1080/135467996387552

Fischer, M. (2003). Spatial representations in number processing—Evidence from a pointing task. *Visual Cognition, 10*(4), 493–508. https://doi.org/10.1080/13506280244000186

Fischer, M. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex, 44*(4), 386–392. https://doi.org/10.1016/j.cortex.2007.08.004

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial–numerical associations. *Brain and Cognition, 72*(3), 333–336. https://doi.org/10.1016/j.bandc.2009.10.010

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27). https://doi.org/10.1073/pnas.1711978115

Ganzeboom, H. (2010). *A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupational 2008 (ISCO-08) constructed with data from the ISSP 2002-2007*. *1*.

Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*(1), 1–56. https://doi.org/10.1016/0049-089X(92)90017-B

Georges, C., Cornu, V., & Schiltz, C. (2023). The importance of spatial language for early numerical development in preschool: Going beyond verbal number skills. *PLoS One1, 18*(9), Article e0292291. https://doi.org/10.1371/journal.pone.0292291

Georges, C., Hoffmann, D., & Schiltz, C. (2016). How Math Anxiety Relates to Number-Space Associations. *Frontiers in Psychology*. , Article 07. https://doi.org/10.3389/fpsyg.2016.01401

Georges, C., Hoffmann, D., & Schiltz, C. (2017). Mathematical abilities in elementary school: Do they relate to number–space associations? *Journal of Experimental Child Psychology, 161*, 126–147. https://doi.org/10.1016/j.jecp.2017.04.011

Gevers, W., Ratinckx, E., De Baene, W., & Fias, W. (2006). Further evidence that the SNARC effect is Processed along a Dual-Route Architecture: Evidence from the Lateralized Readiness potential. *Experimental Psychology, 53*(1), 58–68. https://doi.org/10.1027/1618-3169.53.1.58

Gevers, W., Santens, S., Dhooge, E., Chen, Q., Van Den Bossche, L., Fias, W., & Verguts, T. (2010). Verbal-spatial and visuospatial coding of number–space interactions. *Journal of Experimental Psychology: General, 139*(1), 180–190. https://doi.org/10.1037/a0017688

Gibson, L. C., & Maurer, D. (2016). Development of SNARC and distance effects and their relation to mathematical and visuospatial abilities. *Journal of Experimental Child Psychology, 150*, 301–313. https://doi.org/10.1016/j.jecp.2016.05.009

Göbel, S. M., McCrink, K., Fischer, M. H., & Shaki, S. (2018). Observation of directional storybook reading influences young children's counting direction. *Journal of Experimental Child Psychology, 166*, 49–66. https://doi.org/10.1016/j.jecp.2017.08.001

Guida, A., Fartoukh, M., & Mathy, F. (2020). The development of working memory spatialization revealed by using the cave paradigm in a two-alternative spatial choice. *Annals of the New York Academy of Sciences, 1477*(1), 54–70. https://doi.org/10.1111/nyas.14433

Hawes, Z., & Ansari, D. (2020). What explains the relationship between spatial and mathematical skills? a review of evidence from brain and behavior. *Psychonomic Bulletin & Review, 27*(3), 465–482. https://doi.org/10.3758/s13423-019-01694-7

He, Y., Nuerk, H.-C., Derksen, A., Shi, J., Zhou, X., & Cipora, K. (2021). A gifted SNARC? Directional spatial–numerical associations in gifted children with high-level math skills do not differ from controls. *Psychological Research Psychologische Forschung, 85*(4), 1645–1661. https://doi.org/10.1007/s00426-020-01354-9

Herrera, A., Macizo, P., & Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychologica, 128*(2), 225–237. https://doi.org/10.1016/j.actpsy.2008.01.002

Hesse, P. N., & Bremmer, F. (2017). The SNARC effect in two dimensions: Evidence for a frontoparallel mental number plane. *Vision Research, 130*, 85–96. https://doi.org/10.1016/j.visres.2016.10.007

Hesse, P. N., Fiehler, K., & Bremmer, F. (2016). *SNARC Effect in Different Effectors. Perception, 45*(1–2), 180–195. https://doi.org/10.1177/0301006615614453

Hoffmann, D., Hornung, C., Martin, R., & Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *Journal of Experimental Child Psychology, 116*(4), 775–791. https://doi.org/10.1016/j.jecp.2013.07.013

Hoffmann, D., Mussolin, C., Martin, R., & Schiltz, C. (2014). The Impact of Mathematical Proficiency on the Number-Space Association. *PLoS One1, 9*(1), Article e85048. https://doi.org/10.1371/journal.pone.0085048

Hohol, M., Wołoszyn, K., & Cipora, K. (2022). No fingers, no SNARC? neither the finger counting starting hand, nor its stability robustly affect the SNARC effect. *Acta Psychologica, 230*, Article 103765. https://doi.org/10.1016/j.actpsy.2022.103765

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience, 6*(6), 435–448. https://doi.org/10.1038/nrn1684

Jiang, J., Yang, X., Qi, Y., Yang, X., Yu, X., Lei, X., Xing, J., & Shao, W. (2024). How do symbolic and non-symbolic spatial-numerical associations develop? evidence from the parity judgment task and the magnitude comparison task. *Current Psychology, 43*(18), 16572–16590. https://doi.org/10.1007/s12144-023-05571-4

Keus, I. M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition, 33*(4), 681–695. https://doi.org/10.3758/BF03195335

Lorch, R. F., & Myers, J. L. (1990). *Regression Analyses of Repeated Measures Data in Cognitive Research, 16*(1), 149–157.

McCrink, K., Caldera, C., & Shaki, S. (2018). The Early Construction of Spatial attention: Culture, Space, and Gesture in Parent–Child Interactions. *Child Development, 89*(4), 1141–1156. https://doi.org/10.1111/cdev.12781

McCrink, K., Perez, J., & Baruch, E. (2017). Number prompts left-to-right spatial mapping in toddlerhood. *Developmental Psychology, 53*(7), 1256–1264. https://doi.org/10.1037/dev0000342

McManus, R. M., Young, L., & Sweetman, J. (2023). Psychology is a Property of Persons, not Averages or Distributions: Confronting the Group-to-Person Generalizability Problem in Experimental Psychology. *Advances in Methods and Practices in Psychological Science, 6*(3), Article 25152459231186615. https://doi.org/10.1177/25152459231186615

Nan, W., Yan, L., Yang, G., Liu, X., & Fu, S. (2022). Two processing stages of the SNARC effect. *Psychological Research Psychologische Forschung, 86*(2), 375–385. https://doi.org/10.1007/s00426-021-01506-5

Nuerk, H., Iversen, W., & Willmes, K. (2004). Notational Modulation of the SNARC and the MARC (Linguistic Markedness of Response Codes) effect. *The Quarterly Journal of Experimental Psychology Section A, 57*(5), 835–863. https://doi.org/10.1080/02724980343000512

Nuerk, H.-C., Wood, G., & Willmes, K. (2005). The Universal SNARC effect: The Association between Number Magnitude and Space is Amodal. *Experimental Psychology, 52*(3), 187–194. https://doi.org/10.1027/1618-3169.52.3.187

OpenAI. December). ChatGPT (*GPT-4o*) [Large language model]. https://chat.openai.com.

Opfer, J. E., Thompson, C. A., & Furlong, E. E. (2010). Early development of spatial-numeric associations: Evidence from spatial and quantitative performance of preschoolers. *Developmental Science, 13*(5), 761–771. https://doi.org/10.1111/j.1467-7687.2009.00934.x

Patro, K., Fischer, U., Nuerk, H., & Cress, U. (2016). How to rapidly construct a spatial–numerical representation in preliterate children (at least temporarily). *Developmental Science, 19*(1), 126–144. https://doi.org/10.1111/desc.12296

Patro, K., & Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *Journal of Experimental Child Psychology, 111*(3), 534–542. https://doi.org/10.1016/j.jecp.2011.09.006

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

Robinson, W. S. (1950). Ecological Correlations and the Behavior of individuals. *American Sociological Review, 15*(3), 351–357.

Rouder, J. N., & Haaf, J. M. (2021). Are there Reliable Qualitative Individual Difference in Cognition? *Journal of Cognition, 4*(1), 46. https://doi.org/10.5334/joc.131

Sauce, B., & Matzel, L. D. (2013). The causes of variation in learning and behavior: Why individual differences matter. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00395

Schwarz, W., & Keus, I. M. (2004). Moving the eyes along the mental number line: Comparing SNARC effects with saccadic and manual responses. *Perception & Psychophysics, 66*(4), 651–664. https://doi.org/10.3758/BF03194909

Sella, F., Lucangeli, D., & Zorzi, M. (2019). Spatial order relates to the exact numerical magnitude of digits in young children. *Journal of Experimental Child Psychology, 178*, 385–404. https://doi.org/10.1016/j.jecp.2018.09.001

Shaki, S., & Fischer, M. H. (2008). Reading space into numbers – a cross-linguistic comparison of the SNARC effect. *Cognition, 108*(2), 590–599. https://doi.org/10.1016/j.cognition.2008.04.001

Shaki, S., Fischer, M. H., & Göbel, S. M. (2012). Direction counts: A comparative study of spatially directional counting biases in cultures with different reading directions. *Journal of Experimental Child Psychology, 112*(2), 275–281. https://doi.org/10.1016/j.jecp.2011.12.005

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review, 16*(2), 328–331. https://doi.org/10.3758/PBR.16.2.328

Siegler, R. S. (1987). The Perils of Averaging Data over strategies: An example from Children`s Addition. *Journal of Experimental Psychology: General, 116*(3), 250–264.

Treccani, B., & Umiltà, C. (2011). How to cook a SNARC? Space may be the critical ingredient, after all: A comment on Fischer, Mills, and Shaki (2010). *Brain and Cognition, 75*(3), 310–315. https://doi.org/10.1016/j.bandc.2010.11.006

Van Dijck, J., Abrahamse, E., & Fias, W. (2020). Do preliterate children spontaneously employ spatial coding for serial order in working memory? *Annals of the New York Academy of Sciences, 1477*(1), 91–99. https://doi.org/10.1111/nyas.14430

Van Dijck, J.-P., & Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition, 119*(1), 114–119. https://doi.org/10.1016/j.cognition.2010.12.013

van Dijck, J.-P., Gevers, W., & Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition, 113*(2), 248–253. https://doi.org/10.1016/j.cognition.2009.08.005

van Galen, M. S., & Reitsma, P. (2008). Developing access to number magnitude: A study of the SNARC effect in 7- to 9-year-olds. *Journal of Experimental Child Psychology, 101*(2), 99–113. https://doi.org/10.1016/j.jecp.2008.05.001

Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The Cognitive Mechanisms of the SNARC effect: An Individual differences Approach. *PLoS One1, 9*(4), Article e95756. https://doi.org/10.1371/journal.pone.0095756

Vogel, E. K., & Awh, E. (2008). How to Exploit Diversity for Scientific Gain: Using Individual differences to Constrain Cognitive Theory. *Current Directions in Psychological Science, 17*(2), 171–176. https://doi.org/10.1111/j.1467-8721.2008.00569.x

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: A meta-analysis of the SNARC effect. *Psychology Science Quarterly, 50*(4), 489–525.

Yang, T., Chen, C., Zhou, X., Xu, J., Dong, Q., & Chen, C. (2014). Development of spatial representation of numbers: A study of the SNARC effect in chinese children. *Journal of Experimental Child Psychology, 117*, 1–11. https://doi.org/10.1016/j.jecp.2013.08.011

Zhang, P., Cao, B., & Li, F. (2022). The role of cognitive control in the SNARC effect: A review. *PsyCh Journal, 11*(6), 792–803. https://doi.org/10.1002/pchj.586

Zhang, Y., Yu, X., Qi, Y., Zhang, H., Xu, J., & Chen, Y. (2024). Cognitive factors contribute to the symbolic and the non-symbolic SNARC effects in children and adults. *Cognitive Development, 70*, Article 101422. https://doi.org/10.1016/j.cogdev.2024.101422