Acta Psychologica 230 (2022) 103765

Contents lists available at ScienceDirect

Acta Psychologica

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

No fingers, no SNARC? Neither the finger counting starting hand, nor its stability robustly affect the SNARC effect

Mateusz Hohol [a,*,1], Kinga Wołoszyn [b], Krzysztof Cipora [c,*,1]

[a] Mathematical Cognition and Learning Lab, Copernicus Center for Interdisciplinary Studies, Jagiellonian University, Krakow, Poland
[b] Institute of Psychology, Jagiellonian University, Krakow, Poland
[c] Centre for Mathematical Cognition, Loughborough University, Loughborough, UK

ARTICLE INFO

Keywords:
Finger counting
SNARC effect
Embodied cognition
Numerical cognition
Parity judgment task
Cognitive processing

ABSTRACT

The Spatial-Numerical Association of Response Codes (SNARC) effect (i.e., faster left/right sided responses to small/large magnitude numbers, respectively) is considered to be strong evidence for the link between numbers and space. Studies have shown considerable variation in this effect. Among the factors determining individual differences in the SNARC effect is the hand an individual uses to start the finger counting sequence. Left-starters show a stronger and less variable SNARC effect than right-starters. This observation has been used as an argument for the embodied nature of the SNARC effect. For this to be the case, one must assume that the finger counting sequence (especially the starting hand) is stable over time. Subsequent studies challenged the view that the SNARC differs depending on the finger counting starting hand. At the same time, it has been pointed out that the temporal stability of the finger counting starting hand should not be taken for granted. Thus, in this preregistered study, we aimed to replicate the difference in the SNARC between left- and right-starters and explore the relationship between the self-reported temporal stability of the finger counting starting hand and the SNARC effect. In line with the embodied cognition account, left-starters who declare more temporarily stable finger counting habits should reveal a stronger SNARC effect. Results of the preregistered analysis did not show the difference between left- and right-starters. However, further exploratory analysis provided weak evidence that this might be the case. Lastly, we found no evidence for the relationship between finger counting starting hand stability and the SNARC effect. Overall, these results challenge the view on the embodied nature of the SNARC effect.

1. Introduction

1.1. Where do Spatial-Numerical Associations come from?

Numerous experimental studies have shown that at the cognitive level, several aspects of numerical information are associated with some aspects of space. This broad range of phenomena is referred to as Spatial-Numerical Associations (see Cipora, Haman, Domahs, & Nuerk, 2020; Cipora, Schroeder, Soltanlou, & Nuerk, 2018; Toomarian & Hubbard, 2018 for reviews and taxonomies). The definition provided above seems to be somehow vague and unspecific, however, the large variety of phenomena of numbers being linked to space does not allow more specific SNA definitions, and it may be that Wittgenstein's notion of family resemblance (a number of features overlapping across SNAs but very few specific to all of them) might be useful in characterizing them (Cipora, He, & Nuerk, 2020). For instance, numbers may be associated with spatial extensions (larger numbers—larger extensions: extension SNAs) or with specific directions in space (directional SNAs).

Here we focus on the SNARC effect (Spatial-Numerical Association of Response Codes; Dehaene, Bossini, & Giraux, 1993), which is one of the most thoroughly studied directional SNA (see Wood, Willmes, Nuerk, & Fischer, 2008 for a meta-analysis, and Cipora, Soltanlou, Reips, & Nuerk, 2019 for a large-scale online replication). In speeded bimanual setups (e.g., parity judgment task), faster reactions to small/large numbers with left/right hand, respectively are observed. The SNARC effect is typically considered as evidence that numbers are cognitively represented along the left-to-right aligned mental number line (Dehaene, 2011; Restle, 1970; but see Fias & van Dijck, 2016; Gevers

* Corresponding authors.
E-mail addresses: mateusz.hohol@uj.edu.pl (M. Hohol), k.cipora@lboro.ac.uk (K. Cipora).
[1] Equal contribution.

https://doi.org/10.1016/j.actpsy.2022.103765
Received 31 December 2021; Received in revised form 12 July 2022; Accepted 6 October 2022
Available online 12 October 2022

et al., 2010; Proctor & Cho, 2006 for other accounts). Despite its robustness and replicability at the group level, there are considerable inter-individual differences in the SNARC effect (Cipora et al., 2019; Cipora, Soltanlou, et al., 2019). Several variables have been demonstrated to affect the SNARC effect, such as the reading/writing direction, the efficiency of cognitive inhibition, and the level of mathematical skills (see Cipora, He, & Nuerk, 2020; Wood et al., 2008 for reviews).

Among other factors, the SNARC effect is considered to be linked to the finger counting direction. Fischer (2008) has shown that the SNARC effect was not significant (associated *p*-value of .061) in participants starting finger counting with their right hand (right-starters). It differed significantly from the SNARC effect observed in left-starters. The latter group also revealed a significant SNARC effect. Moreover, the variance in the SNARC effect was greater among right-starters. This observation was only partly replicated in a large-scale online study (Cipora, Soltanlou, et al., 2019), which showed a difference between left- and right-starters in the same direction. Still, it was associated with a negligibly small effect size (Cohen's $d = 0.12$). However, Bayesian analysis has shown that the result was inconclusive and was leaning towards supporting the null hypothesis. At the same time, unlike in Fischer (2008), a robust SNARC effect was found in right-starters, and there was no significant difference in variances between left- and right-starters. Further studies have also demonstrated a robust SNARC in right-starters (Fabbri, 2013; Prete & Tommasi, 2020). Additionally, in several countries where the majority of people start finger counting with their right hand (e.g., Belgium and Italy), the SNARC effect has been observed in multiple studies (e.g., Cutini, Scarpa, Scatturin, Dell'Acqua, & Zorzi, 2014; Gevers, Ratinckx, de Baene, & Fias, 2006; Mapelli, Rusconi, & Umiltà, 2003). To sum up, there seems to be some evidence, however mixed, that finger counting is associated with the SNARC effect (see also Riello & Rusconi, 2011). Having seen these results, one might ask why the SNARC effect should be related to the finger counting direction. The research on the embodiment of numerical cognition can illuminate this issue.

1.2. Finger counting and numerical cognition

For a long time, psychologists considered finger counting mainly an immature behavioral strategy or an interim developmental step in dealing with numbers. They also believed that its role diminishes with formal mathematical education (see Gelman & Gallistel, 1986; Jordan, Kaplan, Ramineni, & Locuniak, 2008; Piaget, 1942). However, now it is well recognized that finger counting, apart from playing an important role in the development of numerical cognition (Moeller et al., 2012), is also practiced even by educated adults. Adults frequently employ their fingers to support calendar and arithmetic calculations, present arguments and plans, count syllables in a phrase, and communicate numbers to others (Bender & Beller, 2011; Hohol, Wołoszyn, Nuerk, & Cipora, 2018; Lucidi & Thevenot, 2014; Pika, Nicoladis, & Marentette, 2009). Various forms of finger counting are observed in the vast majority of cultures, and there is evidence for its use in the past, dating back to the prehistoric era (Göbel, Shaki, & Fischer, 2011; Overmann, 2014, 2021).

The high prevalence of finger counting goes hand in hand with the diversity of finger counting patterns (Bender & Beller, 2012; Butterworth, 1999; Overmann, 2021). For the present study, the crucial finger counting characteristic is the starting hand. It seems to depend on multiple factors. Initially, it was postulated that reading and writing direction plays a prominent role (more left-starters in left-to-right reading cultures, more right-starters in right-to-left reading cultures; Lindemann, Alipour, & Fischer, 2011), but later studies have shown that this is not an ultimate explanation: There is a large variation within left-to-right reading Western and European cultures. While, in some countries, there is a majority of right-starters (e.g., Belgium, France, Italy; di Luca, Granà, Semenza, Seron, & Pesenti, 2006; Lindemann et al., 2011; Sato & Lalain, 2008), in some the proportions are relatively equal (e.g., 57 % right- and 43 % of left-starters in Poland; Hohol et al., 2018). Apart from reading direction, handedness seems to be relevant too, with left-handers being much more likely to be left- than right-starters, and a more equal share of left- and right-starters are among right-handers (Cipora, Gashaj, Gridley, Soltanlou, & Nuerk, 2021; Hohol et al., 2018).

Results of finger counting-related studies are frequently considered to be instances of embodied numerical cognition (e.g., Domahs, Moeller, Huber, Willmes, & Nuerk, 2010; Fischer, 2012; Sixtus, Fischer, & Lindemann, 2017; Wołoszyn & Hohol, 2017). According to this approach, in a nutshell, our bodily interactions play non-trivial causal roles in higher cognition (Barsalou, 2008, 2020; Ostarek & Bottini, 2021), including the processes involved in mathematical activities (Fischer, Felisatti, Kulkova, Mende, & Miklashevsky, 2021; Hohol, 2020; Lakoff & Núñez, 2000). Thus, the impact of finger counting reaches beyond a facilitative role, e.g., offloading working memory resources (Beller & Bender, 2011; Wiese, 2004). In line with the embodied approach, the specific sequence of finger counting (Fischer, 2008; Fischer & Brugger, 2011) is considered to shape SNAs. As Hohol et al. (2018) suggested, these "embodied" theories and studies on the functional role of the finger counting direction rest on an implicit assumption that finger counting routines are stable over time (i.e., are trait-like characteristics).

Nevertheless, finger counting is prone to situated influences, and could be flexibly adjusted according to task demands, e.g., when the preferred starting hand is busy (Hohol et al., 2018; Lucidi & Thevenot, 2014; Wasner, Moeller, Fischer, & Nuerk, 2014). Furthermore, recent studies have recognized that finger counting habits, specifically the starting hand, might not be stable over time (with an estimate of about 75 % of participants starting finger counting with the same hand when tested twice, which is not that high given that the chance level is 50 %; Hohol et al., 2018).

To the best of our knowledge, none of the previous studies have directly tested the relationship between the temporal stability of finger counting routines (with particular emphasis on the starting hand) and other aspects of elementary number processing, such as the SNARC effect. If finger counting routines are constitutive for the formation of directional SNAs (with the SNARC effect at the fore), as Fischer (2008) claims, these routines should be temporally stable, so that a given hand and finger occupying the given relative position (e.g., the first from the left) always corresponds to the same number (see Hohol et al., 2018). This is because the structure of abstract numerical concepts is assumed to be deeply grounded in relatively fixed motoric schemas (Lakoff & Núñez, 2000). Thus, one might expect that on top of the direction of finger counting routines, its stability should determine directional SNAs. So, the typical left-to-right SNARC should be most pronounced in individuals characterized by stable left-starting finger counting routines as the direction of the preferred finger counting routine (left-to-right) would be congruent with the direction of the SNARC effect. In both cases, relatively small numbers would be associated with the left hand/side and relatively large numbers – with the right hand/side. Moreover, it can also be that variations in temporal stability of finger counting routines might account for conflicting results on whether the SNARC effect and finger counting starting hand are related.

The temporal stability of finger counting routines was recently investigated by Hohol et al. (2018). Apart from the finding that finger counting routines were stable in about 75 % of participants when tested again after two months, what is important for the current study, participants were fairly accurate in judging the stability of their routines.

1.3. Objectives of the study and hypotheses

In this preregistered study, we aimed at investigating the relationship between the temporal stability of the finger counting starting hand and the SNARC effect. We also revisited the link between the finger counting starting hand and the SNARC effect (Fischer, 2008). To this end, we recruited 104 Polish native speakers (left-to-right readers). We tested them using the classic parity judgment task with single-digit Arabic numbers (Dehaene et al., 1993). Afterwards, we asked participants to count on their fingers from 1 to 10 (with free hands), memorize the sequence, and answer questions about their finger counting pattern and its temporal stability. Based on findings from a study by Hohol et al. (2018), we decided that it would be sufficient to obtain self-reports on the temporal stability of finger counting routines from our participants.

Firstly, we expected to replicate the SNARC effect in the Polish population (Cipora et al., 2016; Cipora & Nuerk, 2013). Secondly, we expected to replicate Fischer's (2008) finding that left-starters show a stronger SNARC effect than right-starters. Finally, we hypothesized that individuals who declare stability over time in terms of the finger counting starting hand should reveal a stronger SNARC effect than individuals declaring moderately stable or no stable preferences regarding this matter. In particular, left-starters who declare more temporarily stable finger counting habits should reveal a stronger SNARC effect.

2. Method

2.1. Participants

A total of 104 individuals participated in the study; all were native Polish speakers. We excluded 3 participants from the main analyses due to the error rate in the parity judgment task outside the 3 *SDs* from the sample mean. Thus, we finally analyzed the data from 101 individuals (56 females, 45 males) aged 18–35 years ($M = 24.9$, $SD = 3.1$). The group consisted mostly of university graduates ($n = 73$). Other participants were high school ($n = 25$), vocational school ($n = 2$), or junior high/middle school graduates ($n = 1$). Participants were recruited through internal online communication channels of Jagiellonian University as well as social media.

2.2. Materials and procedure

The design of the study was approved by the Ethics Committee for Research at the Institute of Philosophy, Jagiellonian University in Kraków. We preregistered the study on AsPridected.org under no. 54028; https://aspredicted.org/qw6zi.pdf. Data were collected between 27/04/2021 and 20/07/2021.

Participants performed all the tasks on desktops or laptops (we blocked the possibility of running the procedure on tablets and other portable devices). The experimental procedure was implemented and administered online in the Gorilla Experiment Builder (www.gorilla.sc; Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2020). The procedure is available at https://app.gorilla.sc/openmaterials/461468. Since Gorilla assures reliable timing accuracy and precision in behavioral experiments regardless of the operating system and/or Internet browser (Anwyl-Irvine, Dalmaijer, Hodges, & Evershed, 2021), we did not force any restrictions in this matter. During the recruitment and at the beginning of the procedure, participants were asked to perform tasks in a quiet room with a comfortable setting and without long breaks (although after each block of parity judgment tasks, they were encouraged to take a short break to rest their eyes). Participants were informed that they were free to withdraw from the procedure at any point by closing the Internet browser.

After obtaining their informed consent, participants performed the parity judgment task. Next, they responded to the items on the finger counting routines questionnaire. Finally, they responded to the Edinburgh Handedness Inventory. The entire procedure lasted approximately 25 min.

2.2.1. Parity judgment task

To measure the SNARC effect, we used a computerized parity judgment task (Dehaene et al., 1993) modeled on the version previously applied in Cipora et al.'s (2016) study. Participants were asked to decide whether a number presented on the screen is even or odd through pressing the P or Q keys on a standard computer keyboard. The instruction emphasized both speed and accuracy. The task included two blocks with reversed response-to-key mapping. The order of the blocks was counterbalanced across participants. Numbers 1, 2, 3, 4, 6, 7, 8, and 9 were used. Black stimuli (font size 30) were presented against the white background. Each number was presented 30 times within each block (60 times within two blocks). Each block was preceded by a practice session (16 trials) to familiarize participants with the task. During the practice sessions, accuracy feedback was provided, and the required response mapping was indicated in the bottom line of the screen. The order of the trials was randomized with the restriction that each number could not appear more than twice in a row. Each trial started with an eye fixation cross presented for 300 ms. Subsequently, the number appeared. It was presented until the participants' response or for a maximum duration of 2 s. The next trial started after 500 ms.

2.2.2. Finger counting questionnaire

To investigate finger counting routines, we asked the questions from a survey previously used in a paper-and-pencil mode in Hohol et al.'s (2018) study. The questionnaire in Polish, as well as its English translation, can be accessed at https://osf.io/tg98s/. Here, the task began by asking the participants to count with their fingers from 1 to 10 and memorize how they did it. Then, they were asked whether they counted only with their left hand, right hand, or both hands. Subsequently, the participants were presented with a schematic drawing of one or two hands (consistently with their response) and asked to describe their finger counting sequence (ordering the names of the fingers by dragging them with the mouse/touchpad into a properly ordered list). This allowed us to clarify their starting hand. Next, the participants were asked to indicate whether they *always* followed the same sequence as they described: *usually* do so, or *do not have* any stable tendency in terms of the order of finger counting. Then, participants responded to the set of further questions regarding their finger counting routines. All these questions and collected data are available at https://osf.io/tg98s/. Here we report only the items which refer to the hypothesis regarding the relationship between the SNARC effect, starting hand and its temporal stability.

2.2.3. Handedness

Lastly, we used a computerized Edinburgh Handedness Inventory (Oldfield, 1971). The questionnaire consists of 10 items in which participants declare their preferred hand while performing daily activities.

2.3. Preregistration vs. reported study

In the study reported here, we made some deviations from the preregistered procedure. First, we tested more participants than initially declared ($n = 104$ vs. $n = 60$). We planned this sample size given the possibilities of participants' recruitment at the time of the preregistration; it was not guided by an a priori power analysis (see Lakens, 2022). While conducting the study, we had better recruitment opportunities than expected. We have also considered more analyses than initially preregistered. In the following parts, we clearly indicated which analyses were preregistered, and which were not.

2.4. Software

Data processing and analyses were conducted in the R language using RStudio (R Core Team, 2013). We used the following R packages: clinfun (Seshan & Whiting, 2022), cowplot (Wilke, 2020), dplyr (Wickham, François, Henry, & Müller, 2022), ggplot2 (Wickham, 2016), jmv (Selker, Love, Dropmann, & Moreno, 2022), psych (Revelle, 2022). All the data and the R analysis script are shared at the Open Science Framework (https://osf.io/tg98s/).

2.5. Parity judgment task – data preparation

Data from the practice series was not analyzed. Firstly, we excluded participants whose error rates in the parity judgment task exceeded the 3 *SD* relative to the sample mean (this step was not specified in the preregistration, which only stated that we would exclude participants with "excessive error rates"). Next, we excluded errors and anticipations, that is, trials in which RTs were shorter than 200 ms. Subsequently, we applied a sequential filtering method (the same as in Cipora et al., 2016; and Cipora & Nuerk, 2013) to exclude outlier RTs: we calculated RTs and *SD*s for each participant separately and then removed RTs outside ±3 *SD* from a participant's mean. This procedure was repeated until there were no further changes in means and *SD*s.

2.6. Parity judgment task – calculating the SNARC effect

2.6.1. Preregistered

To calculate the SNARC effect, we applied the method proposed by Fias, Brysbaert, Geypens, and D'Ydewalle (1996; see also Cipora & Nuerk, 2013; Cipora et al., 2016). Firstly, we computed dRT (RT right hand − RT left hand) for each number in the case of each participant. Positive dRT values indicate left-hand advantage, whereas negative dRT values indicate right-hand advantage. Next, we regressed dRT values on number magnitude. Unstandardized regression slopes served as a measure of the SNARC effect: A more negative slope corresponds to a stronger SNARC effect. To investigate whether there is a significant SNARC effect at the whole sample level, we tested slopes against zero with the one-sample *t*-test. Since there is a direct prediction regarding the directionality of the SNARC effect, a one-sided *t*-test for negative values was used.

2.6.2. Non-preregistered

Since measuring the SNARC effect size based on the non-standardized regression slope was criticized (Pinhas, Tzelgov, & Ganor-Stern, 2012), we also computed standardized SNARC slopes (standardized regression slopes were Fisher-z transformed to approximate normal distribution, see Cipora, van Dijck, et al., 2019). Finally, we estimated the reliability of both unstandardized and standardized SNARC slopes using the split-half method. We applied Spearman-Brown adjustments to attain the reliability estimate for the whole set of items (see Cipora, van Dijck, et al., 2019 for a detailed description of the method).

2.7. SNARC effect, finger counting starting hand, and stability

2.7.1. Preregistered

To compare the SNARC effect in left- and right-starters, we used the independent samples *t*-test. To investigate the relationship between the SNARC effect and temporal stability of finger counting, we performed the Jonckheere-Terpstra test for monotone trend with 2000 permutations. There are three response options in the stability question, and we expected a monotone decrease in the strength of the SNARC effect with decreasing declared stability of finger counting routines. The preregistered analysis was a suboptimal analytical choice as stability of left- and right-starting has a different meaning from the perspective of embodied numerical cognition.[2] Our original analysis plan did not consider the analysis of finger counting stability to be conducted separately for left- and right-starters. However, we report this analysis for transparency.

2.7.2. Non-preregistered

All analyses described above have also been conducted for the standardized SNARC. The Bayesian equivalent of the independent samples *t*-test was used to quantify evidence for null hypotheses.

The more suited analysis for investigating the effects of the stability of finger counting sequence is comparing left- and right-starters separately. In the case of left-starters, a one-sided test was used as there was a clear prediction regarding the direction of the effect. That was not the case for right-starters, so we used a two-sided test.

As there is no Bayesian equivalent of the Jonckheere-Terpstra test, to quantify the evidence for the null hypothesis, we compared the extreme groups (*I always follow the same sequence* and *I have no stable tendency*) with the Bayesian independent samples *t*-test.

2.8. Individual prevalence of the SNARC effect (non-preregistered)

Next, we investigated the presence of the SNARC effect at the individual level. Even though the regression method does not allow for making inferences about the presence of the effects of interest at the individual level, it is possible with a bootstrapping approach, specifically the H0 method proposed by Cipora, van Dijck, et al. (2019). This method aims at checking how (un)likely finding the SNARC effect as empirically observed in each participant would be if there was no association between numerical magnitude and response side. Therefore, separately for each participant and each number, we randomly sampled (with replacement) two sets of 30 responses. One of these sets we considered as "left-handed responses" and the other as "right-handed responses," and we subsequently used them to calculate the SNARC slope (both unstandardized and standardized). We repeated the bootstrapping procedure 5000 times. The slopes from these bootstrap-based regressions were considered as possible outcomes of the analysis if there was no SNARC effect. Finally, we checked whether empirically observed slopes were outside the mid 90 % of the distribution of the bootstrap slopes (i.e., the 90 % H0 confidence intervals). If that was the case, we classified the participant as revealing either a reliable SNARC effect (observed slope < lower bound of the H0 confidence interval), reliable reversed SNARC (observed slope > upper bound of the H0 confidence interval), or non-reliable SNARC (observed slope within the H0 confidence interval; see also van Dijck, Fias, & Cipora, 2022; Hohol et al., 2020).

2.9. Individual prevalence of SNARC effect and finger counting routines (non-preregistered)

To compare the proportions of left-starters and right-starters among the participants revealing reliable, reliable reverse, or no reliable SNARC, we used Fisher's exact test. We applied the same procedure to investigate the relationship between finger counting stability over time and individual prevalence. In both cases, we used unstandardized and standardized SNARC slopes.

2.10. Handedness (non-preregistered)

To investigate handedness, we calculated the laterality quotient (LQ) which could range from −100 to +100. We applied Oldfield's (1971) recommendations, categorizing participants as right-handers if LQ > 40, ambidextrous if 40 ≤ LQ > 0, and left-handers if LQ ≤ 0. The vast majority of the group were right-handers (n = 93); only one person was ambidextrous, and seven were left-handers. Given a small number of non-right-handers, relationships between neither handedness and the SNARC effect nor handedness and finger counting routines were not analyzed. Additionally, previous studies (Cipora, Soltanlou, et al., 2019 Supplementary material 2; Dehaene et al., 1993 Experiment 5; Huber et al., 2014) did not show relationships between left- and right-handers in the SNARC effect.

[2] We wish to thank anonymous reviewers for bringing it to our attention.

3. Results

3.1. Preregistered analyses

3.1.1. Data preprocessing

Firstly, we excluded participants with accuracy lower than 3 *SD* from the sample mean. This led to the exclusion of 3 participants. The overall accuracy of the remaining 101 participants was 94 %. We also excluded the anticipations (correct responses < 200 ms; <0.1 % of the trials). Subsequently, we performed sequential trimming for ±3 *SD* from the individual mean (Cipora et al., 2016; Cipora & Nuerk, 2013; Cipora, Soltanlou, et al., 2019; Hohol et al., 2020). This excluded 5.8 % of the data. Thus, we included 88.5 % of all trials in the further analysis. Mean reaction time was 602,8 ms (*SD* = 203.3).

3.1.2. Unstandardized SNARC slopes

As expected, we found a significant SNARC effect at the whole sample level. Mean slope was −4.71 (*SD* = 5.70), and it differed significantly from 0 ($t_{100} = -8.29$, $p < .001$; one-sided).

3.1.3. Unstandardized SNARC and starting hand

More participants declared starting finger counting with their left (67 persons) than with their right hand (34 persons). Mean slope of left-starters was −5.17 (*SD* = 5.80), and mean slope of right-starters was −3.78 (*SD* = 5.48). The *t*-test showed no significant difference between left-starters and right-starters ($t_{99} = -1.16$, $p = .249$; $d = 0.24$). The (non-preregistered) Bayesian *t*-test showed anecdotal support for the null hypothesis; $BF_{01} = 2.51$. See Fig. 1, Panel A. Moreover, the variance in SNARC slopes is almost the same in both groups (unlike in Fischer, 2008).

3.1.4. Unstandardized SNARC and temporal stability of finger counting

Most participants declared moderate stability over time of the order of finger counting (44; "I usually follow the same sequence"). A similar proportion of participants reported no preference (28; "I have no stable tendency") and strong stability (29; "I always follow the same sequence"). The Jonckheere-Terpstra test (2000 permutations) did not show evidence for a monotone trend for directional alternative hypothesis that more robust stability is linked to stronger SNARC ($T_{JT} = 1798$, $p = .809$ for directional alternative hypothesis). See Fig. 2 Panel A. Comparing the extreme groups with the (non-preregistered) Bayesian *t*-test supported the null hypothesis $BF_{01} = 3.37$.

3.2. Non-preregistered analyses

3.2.1. Reliability of the SNARC effect

Reliability (split-half; Spearman-Brown corrected) of the SNARC slopes was 0.40 and 0.33 for unstandardized and standardized SNARC respectively.

3.2.2. Unstandardized SNARC and temporal stability of finger counting – separately for left- and right-starters

The declared stability split by starting hand was "always" for 20 left- and 8 right-starters; "sometimes" for 28 left- and 16 right-starters; "not stable" for 19 left- and 10 right-starters. In both left- and right-starters there was no relationship between the stability of the finger counting and the strength of the SNARC effect. In the case of left-starters, it was tested with one-sided Jonckheere-Terpstra test, $T_{JT} = 760$, $p = .605$. Bayesian *t*-test supported the null hypothesis, $BF_{01} = 3.18$. In the case of right-starters, we two-sided Jonckheere-Terpstra test was used, $T_{JT} = 212$, $p = .404$. Bayesian *t*-test was largely inconclusive, $BF_{01} = 2.11$.

3.2.3. Standardized SNARC slopes

As expected, we found a significant SNARC effect at the whole sample level. Mean slope was −0.34 (*SD* = 0.39), and it differed significantly from 0 ($t_{100} = -8.84$, $p < .001$; one-sided).

3.2.4. Standardized SNARC slopes and starting hand

Mean slope of left-starters was −0.36 (*SD* = 0.36), and mean slope of right-starters was −0.30 (*SD* = 0.43). The *t*-test showed no significant

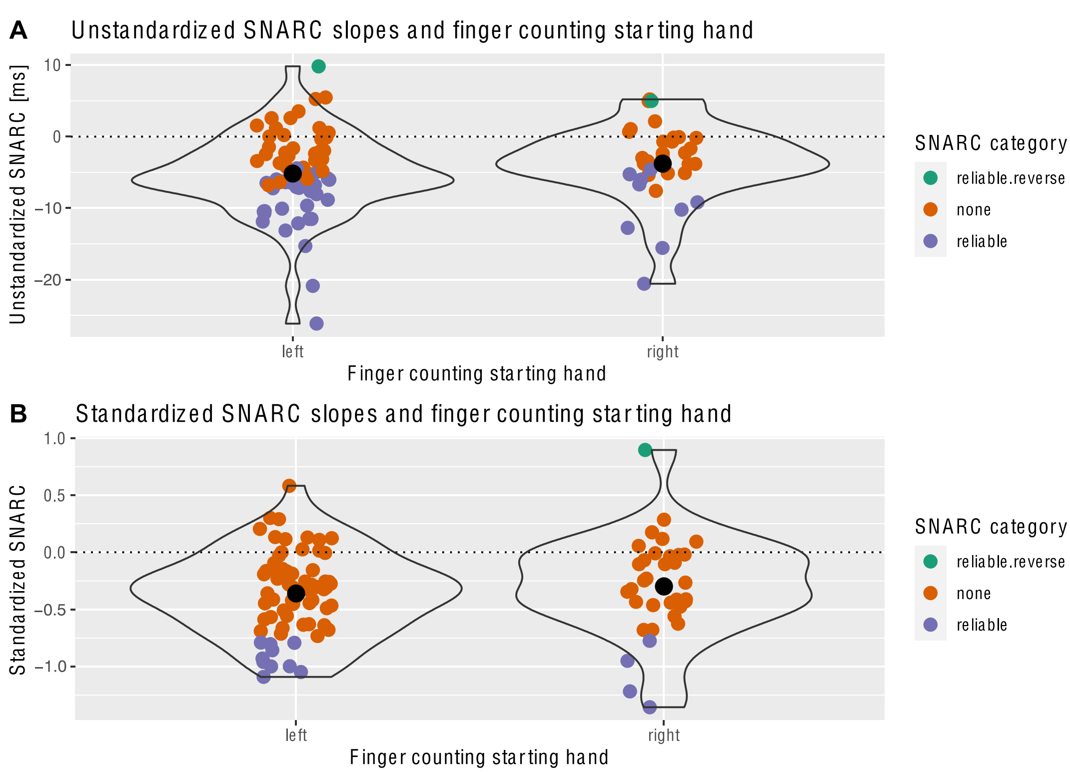

Fig. 1. Unstandardized (panel A) and standardized (panel B) SNARC effect and finger counting starting hand. Colors represent the results of the H0 bootstrapping. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

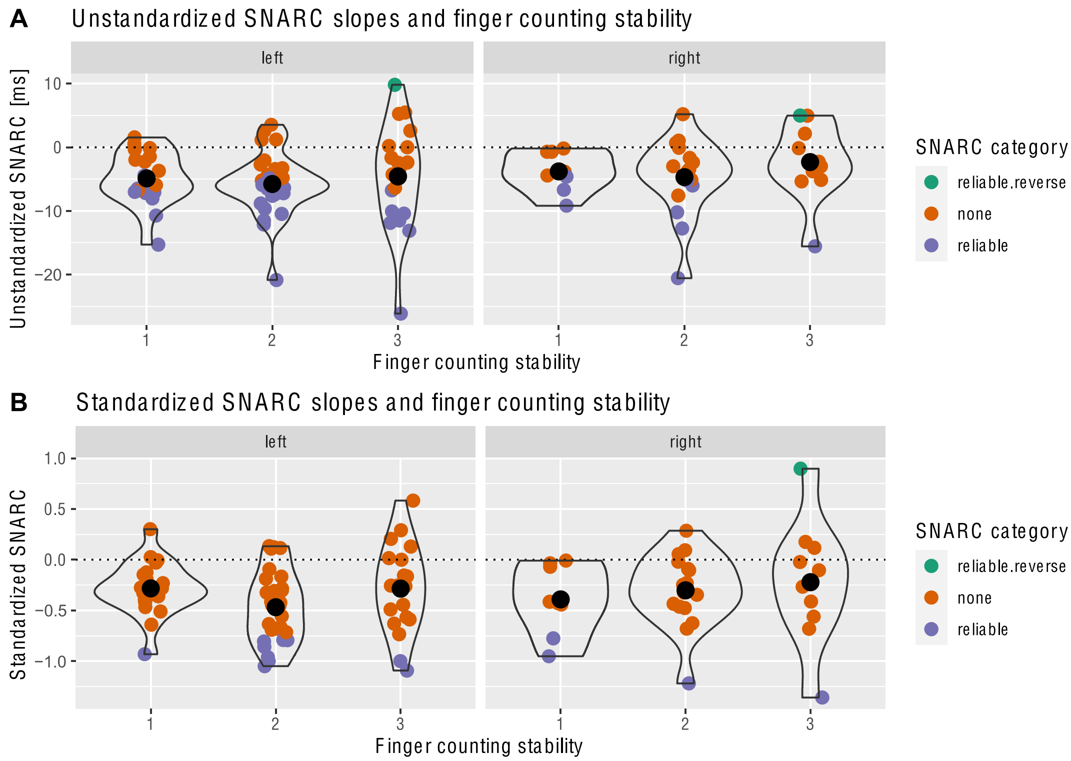

Fig. 2. Unstandardized (panel A) and standardized (panel B) SNARC effect and finger counting starting hand stability (1 = always, 2 = sometimes, 3 = not stable). Left and right subpanels correspond to left- and right-starters accordingly. Colors represent the results of the H0 bootstrapping. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

difference between left-starters and right-starters ($t_{99} = -0.76$, $p = .448$; $d = 0.16$). The Bayesian analysis supported the null hypothesis ($BF_{01} = 3.51$). See Fig. 1, Panel B.

3.2.5. Standardized SNARC and temporal stability of finger counting

The Jonckheere-Terpstra test (2000 permutations) did not show evidence for a monotone trend for directional alternative hypothesis that stronger stability is linked to stronger SNARC ($T_{JT} = 1719$, $p = .653$ for directional alternative hypothesis). See Fig. 2 Panel B. The comparison of the extreme groups with the Bayesian *t*-test supported the null hypothesis $BF_{01} = 3.40$.

3.2.6. Standardized SNARC and temporal stability of finger counting – separately for left- and right-starters

Neither in left- nor in right-starters, there was a relationship between the stability of the finger counting and the strength of the standardized SNARC effect. In the case of the left-starters, it was tested with a one-sided Jonckheere-Terpstra test, $T_{JT} = 719$, $p = .425$. Bayesian *t*-test supported the null hypothesis, $BF_{01} = 3.21$. In case of right-starters, a two-sided Jonckheere-Terpstra test was used, $T_{JT} = 208$, $p = .444$. Bayesian *t*-test was largely inconclusive $BF_{01} = 2.04$.

3.2.7. Individual prevalence of SNARC effect

For unstandardized SNARC, bootstrapping revealed that 42.5 % of participants had a reliable effect. 2 % of participants displayed a reliable reverse effect. Most importantly, 55.4 % of participants did not have a reliable effect. On the other hand, for standardized SNARC, only 15 % of participants had a reliable effect. 1 % of participants displayed a reliable reverse effect. Crucially, 86 % of participants did not have a reliable effect.

3.2.8. Individual prevalence of SNARC effect and starting hand

As the cell sizes were smaller than 5 in the case of reliable positive unstandardized SNARC, we used Fisher's exact test to compare the proportions of participants revealing reliable, reliable reverse, or no reliable SNARC. Fisher exact test ($n = 101$) $p = .037$. [Note that the χ^2 was not significant; $\chi^2_2 = 5.48$, $p = .065$.] This indicates that significantly more left-starters revealed reliable SNARC. In a subsequent analysis, we excluded three participants who revealed reliable positive (reversed) SNARC. The observed effect got even more pronounced, $\chi^2_1 = 5.26$, $p = .022$. See color coding in Fig. 1, Panel A. The same analysis for standardized SNARC did not reveal a significant effect, Fisher exact test ($n = 101$) $p = .486$. See color coding in Fig. 1, Panel B.

3.2.9. Individual prevalence of SNARC effect and declared temporal stability of finger counting routines

There was no significant relationship between declared finger counting pattern stability over time and individual prevalence of unstandardized SNARC effect. Fisher exact test ($n = 101$) $p = .111$. For standardized SNARC $p = .516$.

3.2.10. Consistency of the finger counting stability measures

As our questionnaire contained more items concerning finger counting stability, we looked at the congruence between these methods. First, we looked at the relation between declared stability and actions that participants reported taking when they count with their fingers while their preferred hand was occupied holding something (options: *using the other hand, trying to count with the occupied hand anyway*, *putting away the held object*). The majority of participants declared that they would use the other hand ($N = 62$). The analysis with the Fisher's exact test did not reveal a significant effect $p = .186$, suggesting that there was no relation between responses to these two items. So, despite declared stability, most participants declared that if their preferred hand was busy, they would most likely simply count with the non-preferred hand.

Secondly, we examined the relation between declared stability and the degree of naturalness and comfort while counting with their non-preferred hand (both reported on 5-point Likert-like scales). This was tested by Jonckheere-Terpstra tests. In both cases, the effect was significant, $T_{JT} = 2035.5$, $p = .007$, and $T_{JT} = 1961.6$, $p = .024$ for directional alternative hypotheses for *unnatural* and *discomfort* items respectively. This indicates that participants who declared having stable routines declared higher discomfort and felt it less natural if they had to count on fingers with a non-preferred hand as compared to participants who declared that they do not have strong preferences.

4. Discussion

4.1. Overview

In this study, we first aimed to replicate the SNARC effect and then replicate the results reported by Fischer (2008) that individuals who start finger counting with their right hand show weaker (or no) SNARC, while the stronger effect is present in left-starters. Finally, we intended to check whether the larger temporal stability of starting hand while finger counting is related to the stronger SNARC effect. In the line of embodied cognition account, left-starters who declare more temporarily stable finger counting habits should reveal a stronger SNARC effect.

During the preregistered analyses, we found a robust SNARC effect at the whole sample level. Still, we neither found significant differences between right- and left-starters in the SNARC effect nor differences related to finger counting stability over time.

Results of the subsequent (non-preregistered) analyses were mainly in line with these observations. However, the bootstrapping suggested that more left-starters than right-starters show a reliable SNARC effect at an individual level. Further analyses provided some evidence for consistency between various measures of finger counting stability over time.

4.2. Replicating the SNARC effect

We observed a robust SNARC effect at a group level comparable in size to the effects reported in the previous studies with Polish participants (Cipora et al., 2016; Cipora & Nuerk, 2013). At the same time, the reliability of the SNARC effect was relatively low. However, this estimate was similar to those observed in the online setup (Cipora, Soltanlou, et al., 2019). The proportion of participants displaying the reliable SNARC effect at an individual level was analogous to the one reported in previous studies (Cipora, van Dijck, et al., 2019).

Expectedly, the conclusions from both unstandardized and standardized SNARC slopes are very similar when it comes both to the presence of the SNARC and the (lack of) relationships between SNARC and finger counting. While we observed a difference in proportions of participants revealing a reliable SNARC effect, this observation is in line with conclusions from other studies using these methods (see Cipora, van Dijck, et al., 2019; Hohol et al., 2020).

4.3. Finger counting routines

The proportion of participants declaring stable or relatively stable finger counting routines was approximately 73 %. This ratio is slightly lower than the one observed in the previous study by Hohol et al. (2018), which was 86 %. At the same time, the proportion obtained in the current study closely resembles the actual performance when participants counted with their fingers twice, a few months apart (i.e., 75 %; Hohol et al., 2018). As reported in this study, there is a fair consistency between finger counting starting hand stability measures.

Moreover, the proportion of left- (66 %) and right-starters (34 %) differed from the results reported by Hohol et al. (2018), who observed that the majority (57 %) of participants were right-starters. On the other hand, another online study showed more equal proportions with a predominance of left-starters among Polish participants (Cipora, Soltanlou, et al., 2019, Supplementary material).

4.4. The SNARC effect and finger counting starting hand

We did not observe a discrepancy in the SNARC effect between left- and right-starters. We also found no evidence for a difference in variances between these two groups, thus, we did not replicate Fischer's (2008) results. On the contrary, there was a robust SNARC effect in both groups, which is in line with the results of other studies (Cipora, Soltanlou, et al., 2019, Fabbri, 2013; Prete & Tommasi, 2020). The Bayesian analysis also leaned towards no between-group difference, however, the results remained inconclusive.

At the same time, we need to acknowledge that there might be some evidence for a more robust SNARC effect among left-starters since at the level of descriptive statistics, the SNARC effect was stronger in this group. Additionally, the non-preregistered, exploratory analysis using the bootstrapping method revealed that the proportion of participants showing a reliable SNARC effect at the individual level was larger among left-starters. This was only observed in the case of the unstandardized SNARC. This might be due to general problems with H0 bootstrapping approach for standardized slopes.[3] However, based on current results and the results of previous studies (Cipora, Soltanlou, et al., 2019; Fabbri, 2013; Prete & Tommasi, 2020), we can conclude that even if there is a difference in the SNARC effect between left- and right-starters, it is not as robust as initially proposed by Fischer (2008).

4.5. The SNARC effect and finger counting stability over time

None of the analyses revealed between-group differences in the SNARC effect depending on the declared stability of starting hand while finger counting. Whether the participants declared strong stability of their finger counting routines was not linked to the strength of their SNARC effect, no matter whether they declared to be right- or left-starters. The results of the Bayesian analysis also supported the lack of between-group differences. The null result held no matter which method of quantifying the SNARC effect was used (unstandardized and standardized slopes, and to a degree, the reliability of the effect at the individual level). This observation directly contradicts our preregistered hypotheses that both finger counting direction and stability of finger counting routines are linked to the strength of directional Spatial-Numerical Associations.

4.6. Limitations of the study

While bringing several new insights, our investigation has some limitations that should be addressed in subsequent studies. Firstly, the group sizes are unequal due to the quasi-experimental design, which affects experimental power. We acknowledge that, despite the overall consistency of the observed results, the evidential strength of reported Bayes factors is limited, and this might require further validation in the following studies. However, our experiment shows clearly that the effects of finger counting on the SNARC (if there are any) are small, and relatively large samples would be required to provide more conclusive evidence.

The reliability of the SNARC effect was relatively low. Considering the previous results (Cipora, Soltanlou, et al., 2019), that was to be expected in the online study. However, we acknowledge that this might also have affected the quality of our results.

In this study, we did not use a test-retest design for establishing the temporal stability of finger counting. Even though previous results have shown the validity of self-declared stability (Hohol et al., 2018), other studies have shown general problems with the adequacy of self-

[3] Only recently Cipora, van Dijck, et al., 2019 (preprint updated in 2022), argued that H0 bootstrapping should not be used in case of standardized slopes, as it overestimates the proportion of participants revealing non-reliable SNARC, so the results for unstandardized slopes should be taken into consideration.

reporting in establishing the finger counting direction (Lucidi & Thevenot, 2014; Morrissey & Hallett, 2018). Thus, future research on the relationship between finger counting routines and the SNARC effect should include a more controlled, preferably lab-based protocol.

5. Conclusion

A temporally stable correspondence between the fingers and numbers (indexed by finger counting starting hand) seems crucial for building explanations of the SNAs as shaped by finger counting habits, thus formulating the "embodied" interpretations of the SNARC effect. Despite replicating the well-established SNARC effect, our preregistered study did not provide evidence for its dependence on the finger counting direction. Simultaneously, it did not support the claim that the SNARC effect depends on the temporal stability of finger counting routines, particularly in terms of the starting hand. While the embodied approach undoubtedly has heuristic potential for the study of mathematical cognition, our results challenge its power in explaining the genesis of directional SNAs.

Funding

MH research and open access publication have been supported by a grant from the Priority Research Area FutureSoc under the Strategic Programme Excellence Initiative at Jagiellonian University. KW was supported by the grant 2017/25/N/HS6/01052 funded by the National Science Centre, Poland.

Compliance with ethical standards

This study was performed in line with the principles of the Declaration of Helsinki. The design of the study was approved by the Ethics Committee for Research at the Institute of Philosophy, Jagiellonian University in Krakow. Informed consent was obtained from all individual participants included in the study.

CRediT authorship contribution statement

Conceptualization: MH, KW, KC; Methodology: MH, KW, KC; Data collection: MH; Analysis: MH, KW, KC; Interpretation: MH, KW, KC; Writing – original draft preparation: MH, KC; Writing – review, editing, revision: MH, KW, KC.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

The data as well as analysis scripts are available at https://osf.io/tg98s/. The preregistration is available at https://aspredicted.org/qw6zi.pdf. The procedure is available at https://app.gorilla.sc/openmaterials/461468.

Acknowledgements

We would like to thank Aleksandra Kolny and Piotr Szymanek for contributing to data collection, and anonymous reviewers for their constructive comments.

References

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods, 53*, 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2020). Gorilla in our MIDST: An online behavioral experiment builder. *Behavior Research Methods, 52*, 388–407. https://doi.org/10.3758/s13428-019-01237-x

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Barsalou, L. W. (2020). Challenges and opportunities for grounding cognition. *Journal of Cognition, 3*(1), 1–24. https://doi.org/10.5334/joc.116

Beller, S., & Bender, A. (2011). Explicating numerical information: When and how fingers support (or hinder) number comprehension and handling. *Frontiers in Psychology, 2*(214). https://doi.org/10.3389/fpsyg.2011.00214

Bender, A., & Beller, S. (2011). Fingers as a tool for counting? Naturally fixed or culturally flexible? *Frontiers in Psychology, 2*(256). https://doi.org/10.3389/fpsyg.2011.00256

Bender, A., & Beller, S. (2012). Nature and culture of finger counting: Diversity and representational effects of an embodied cognitive tool. *Cognition, 124*(2), 156–182. https://doi.org/10.1016/j.cognition.2012.05.005

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

Cipora, K., Gashaj, V., Gridley, A. S., Soltanlou, M., & Nuerk, H.-C. (2021). Universalities and cultural specificities of finger counting and montring: Evidence from Amazon tsimane' people. *PsyArxiv Preprints*. https://doi.org/10.31234/osf.io/suhej

Cipora, K., Haman, M., Domahs, F., & Nuerk, H.-C. (2020). *On the development of space-number relations: Linguistic and cognitive determinants, influences, and associations*. In (p. 11(182).). Frontiers in Psychology. https://doi.org/10.3389/fpsyg.2020.00182.

Cipora, K., He, Y., & Nuerk, H.-C. (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., Hohol, M., Nuerk, H.-C., Willmes, K., Brożek, B., Kucharzyk, B., & Nęcka, E. (2016). Professional mathematicians differ from controls in their spatial-numerical associations. *Psychological Research, 80*(4), 710–726. https://doi.org/10.1007/s00426-015-0677-6

Cipora, K., & Nuerk, H.-C. (2013). 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. *The Quarterly Journal of Experimental Psychology, 66*(10), 1974–1991. https://doi.org/10.1080/17470218.2013.772215

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: The role of spatial reasoning in mathematical thought* (pp. 77–116). Cham: Springer. https://doi.org/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*, 1–17. https://doi.org/10.3758/s13428-019-01213-5

Cipora, K., van Dijck, J.-P., Georges, C., Masson, N., Göbel, S. M., Willmes, K.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. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/bwyr3

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

Dehaene, S. (2011). *The number sense* (Revised ed.). Oxford: 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. https://doi.org/10.1037/0096-3445.122.3.371

di Luca, S., Granà, A., Semenza, C., Seron, X., & Pesenti, M. (2006). Finger-digit compatibility in Arabic numeral processing. *The Quarterly Journal of Experimental Psychology, 59*(9), 1648–1663. https://doi.org/10.1080/17470210500256839

Domahs, F., Moeller, K., Huber, S., Willmes, K., & Nuerk, H.-C. (2010). Embodied numerosity: Implicit hand-based representations influence symbolic number processing across cultures. *Cognition, 116*(2), 251–266. https://doi.org/10.1016/j.cognition.2010.05.007

Fabbri, M. (2013). Finger counting habits and spatial-numerical association in horizontal and vertical orientations. *Journal of Cognition and Culture, 13*(1–2), 95–110. https://doi.org/10.1163/15685373-12342086

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

Fias, W., & van Dijck, J. P. (2016). The temporary nature of number-space interactions. *Canadian Journal of Experimental Psychology, 70*(1), 33–40. https://doi.org/10.1037/CEP0000071

Fischer, M. H. (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. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cognitive Processing, 13*(S1), 161–164. https://doi.org/10.1007/s10339-012-0477-5

Fischer, M. H., & Brugger, P. (2011). When digits help digits: Spatial, numerical associations point to finger counting as prime example of embodied cognition. *Frontiers in Psychology, 2*. https://doi.org/10.3389/fpsyg.2011.00260

Fischer, M. H., Felisatti, A., Kulkova, E., Mende, M. A., & Miklashevsky, A. (2021). Measuring the mathematical mind: Embodied evidence from motor resonance, negative numbers, calculation biases, and emotional priming. In M. D. Robinson, & L. E. Thomas (Eds.), *Handbook of embodied psychology* (pp. 149–170). Cham: Springer.

Gelman, R., & Gallistel, C. (1986). *The child's understanding of number*. Cambridge, MA: Harvard University Press.

Gevers, W., Ratinckx, E., de Baene, W., & Fias, W. (2006). Further evidence that the SNARC effect is processed along a dual-route architectures: 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

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. *Journal of Cross-Cultural Psychology, 42*(4), 543–565. https://doi.org/10.1177/0022022111406251

Hohol, M. (2020). *Foundations of geometric cognition*. London-New York: Routledge. https://doi.org/10.4324/9780429056291

Hohol, M., Willmes, K., Nęcka, E., Brożek, B., Nuerk, H.-C., & Cipora, K. (2020). Professional mathematicians do not differ from others in the symbolic numerical distance and size effects. *Scientific Reports, 10*(11531). https://doi.org/10.1038/s41598-020-68202-z

Hohol, M., Wołoszyn, K., Nuerk, H.-C., & Cipora, K. (2018). A large-scale survey on finger counting routines, their temporal stability and flexibility in educated adults. *PeerJ, 6* (e5878). https://doi.org/10.7717/peerj.5878

Huber, S., Klein, E., Graf, M., Nuerk, H.-C., Moeller, K., & Willmes, K. (2014). Embodied markedness of parity? Examining handedness effects on parity judgments. *Psychological Research, 79*(6), 963–977. https://doi.org/10.1007/s00426-014-0626-9

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2008). Development of number combination skill in the early school years: When do fingers help? *Developmental Science, 11*(5), 662–668. https://doi.org/10.1111/j.1467-7687.2008.00715.x

Lakens, D. (2022). Sample size justification. *Collabra:Psychology, 8*(1), 33267. https://doi.org/10.1525/collabra.33267

Lakoff, G., & Núñez, R. (2000). *Where mathematics comes from*. New York: Basic Books.

Lindemann, O., Alipour, A., & Fischer, M. H. (2011). Finger counting habits in middle eastern and Western individuals: An online survey. *Journal of Cross-Cultural Psychology, 42*(4), 566–578. https://doi.org/10.1177/0022022111406254

Lucidi, A., & Thevenot, C. (2014). Do not count on me to imagine how I act: Behavior contradicts questionnaire responses in the assessment of finger counting habits. *Behavior Research Methods, 46*(4), 1079–1087. https://doi.org/10.3758/s13428-014-0447-1

Mapelli, D., Rusconi, E., & Umiltà, C. (2003). The SNARC effect: An instance of the Simon effect? *Cognition, 88*(3), B1–B10. https://doi.org/10.1016/S0010-0277(03)00042-8

Moeller, K., Fischer, U., Link, T., Wasner, M., Huber, S., Cress, U., & Nuerk, H.-C. (2012). Learning and development of embodied numerosity. *Cognitive Processing, 13*(S1), 271–274. https://doi.org/10.1007/s10339-012-0457-9

Morrissey, K., & Hallett, D. (2018). Cardinal and ordinal aspects of finger-counting habits predict different individual differences in embodied numerosity. *Journal of Numerical Cognition, 4*(3), 613–635. https://doi.org/10.5964/jnc.v4i3.138

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Ostarek, M., & Bottini, R. (2021). Towards strong inference in research on embodiment – Possibilities and limitations of causal paradigms. *Journal of Cognition, 4*(1), 1–21. https://doi.org/10.5334/joc.139

Overmann, K. A. (2014). Finger-counting in the upper Palaeolithic. *Rock Art Research, 31* (1), 63. https://doi.org/10.31235/osf.io/wgbe5

Overmann, K. A. (2021). Finger-counting and numerical structure. *Frontiers in Psychology, 12*(723492). https://doi.org/10.3389/fpsyg.2021.723492

Piaget, J. (1942). *Child's conception of number*. London: Routledge.

Pika, S., Nicoladis, E., & Marentette, P. (2009). How to order a beer: Cultural differences in the use of conventional gestures for numbers. *Journal of Cross-Cultural Psychology, 40*(1), 70–80. https://doi.org/10.1177/0022022108326197

Pinhas, M., Tzelgov, J., & Ganor-Stern, D. (2012). Estimating linear effects in ANOVA designs: The easy way. *Behavior Research Methods, 44*(3), 788–794. https://doi.org/10.3758/S13428-011-0172-Y/TABLES/4

Prete, G., & Tommasi, L. (2020). Exploring the interactions among SNARC effect, finger counting direction and embodied cognition. *PeerJ, 7*(9155). https://doi.org/10.7717/peerj.9155

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

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

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*(2p1), 274–278. https://doi.org/10.1037/h0028573

Revelle, W. (2022). *psych: Procedures for psychological, psychometric, and personality research*. Retrieved from. Illinois: Evanston https://cran.r-project.org/package=psych.

Riello, M., & Rusconi, E. (2011). Unimanual SNARC effect: Hand matters. *Frontiers in Psychology, 2*(372), 1–11. https://doi.org/10.3389/fpsyg.2011.00372

Sato, M., & Lalain, M. (2008). On the relationship between handedness and hand-digit mapping in finger counting. *Cortex, 44*(4), 393–399. https://doi.org/10.1016/j.cortex.2007.08.005

Selker, R., Love, J., Dropmann, D., & Moreno, V. (2022). Jmv: The "jamovi" analyses. *Retrieved from*. https://cran.r-project.org/package=jmv.

Seshan, V. E., & Whiting, K. (2022). Clinfun: Clinical trial design and data analysis functions. *Retrieved from*. https://cran.r-project.org/package=clinfun.

Sixtus, E., Fischer, M. H., & Lindemann, O. (2017). Finger posing primes number comprehension. *Cognitive Processing, 15*, 1–12. https://doi.org/10.1007/s10339-017-0804-y

Toomarian, E. Y., & Hubbard, E. M. (2018). On the genesis of spatial-numerical associations: Evolutionary and cultural factors co-construct the mental number line. *Neuroscience and Biobehavioral Reviews, 90*, 184–199. https://doi.org/10.1016/j.neubiorev.2018.04.010

van Dijck, J.-P., Fias, W., & Cipora, K. (2022). Spatialization in working memory and its relation to math anxiety. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/nyas.14765

Wasner, M., Moeller, K., Fischer, M. H., & Nuerk, H.-C. (2014). Aspects of situated cognition in embodied numerosity: The case of finger counting. *Cognitive Processing, 15*(3), 317–328. https://doi.org/10.1007/s10339-014-0599-z

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Retrieved from. New York: Springer-Verlag https://ggplot2.tidyverse.org.

Wickham, H., François, R., Henry, L., & Müller, K. (2022). Dplyr: A grammar of data manipulation. *Retrieved from*. https://cran.r-project.org/package=dplyr.

Wiese, H. (2004). *Numbers, language, and the human mind*. Cambridge: Cambridge University Press.

Wilke, C. O. (2020). Cowplot: Streamlined plot theme and plot annotations for "ggplot2". *Retrieved from*. https://cran.r-project.org/package=cowplot.

Wołoszyn, K., & Hohol, M. (2017). Commentary: The poverty of embodied cognition. *Frontiers in Psychology, 8*(845). https://doi.org/10.3389/fpsyg.2017.00845

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.