Children’s spatial–numerical associations on horizontal, vertical, and sagittal axes

Sarah M. Cooney *, Corinne A. Holmes, Fiona N. Newell

School of Psychology and Institute of Neuroscience, Trinity College Dublin, College Green, Dublin 2, Ireland

A R T I C L E   I N F O

Article history:
Received 6 September 2019
Revised 18 March 2021
Available online 3 May 2021

Keywords:
Development
Number magnitude
Mental number line
Axes
SNARC
Numerical cognition

A B S T R A C T

There is substantial evidence linking numerical magnitude to the physical properties of space. The most influential support for this connection comes from the SNARC effect (spatial–numerical association of response codes), in which responses to small/large numbers are faster on the left/right side of space, respectively. The SNARC effect has been extensively replicated, and is understood as horizontal mapping of numerical magnitude. However, much less is known about how numbers are represented on the vertical and sagittal axes, and whether spatial–numerical associations on different axes emerge during childhood. To that end, we tested two groups of children, aged 5–7 years and 8 and 9 years, on a single-digit magnitude comparison task with response buttons positioned either upper/lower (vertical), left/right (horizontal) or near/far (sagittal). Our results provide evidence of spatial–numerical mapping on all three axes for both age groups that are similar in strength. This indicates that, even at an early stage of formal education, children can flexibly assign numerical magnitude to any spatial dimension. To examine the contribution of extracorporeal space and spatio-anatomical mapping to the SNARC effect across axes, these sources were pitted against each other by swapping the position of the response hands in Experiment 1b. Switching hand position did not reveal convincing evidence for SNARC effects on any axis. Results are discussed with respect to the utility of three-dimensional mental number lines, and potential avenues for future research are outlined.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

* Corresponding author at: School of Psychology, University College Dublin, Belfield, Dublin 4, Ireland.
E-mail address: s.cooney@ucd.ie (S.M. Cooney).

https://doi.org/10.1016/j.jecp.2021.105169
0022-0965/© 2021 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Introduction

There is substantial evidence linking numerical magnitude to the physical properties of extracorporeal space. In his investigations of visualized numerals, Francis Galton was the first to describe the spatial reference frame of the mental number line (MNL). Although the descriptions were idiosyncratic, they demonstrate that numbers are often tied to a specific spatial location that can be aligned vertically, horizontally, diagonally, or in depth (Galton, 1880). More than a century later, the most influential support for number–space mapping comes from the SNARC effect (spatial–numerical association of response codes), where left-sided key presses are faster when responding to relatively small numbers and right-sided key presses are faster when responding to relatively large numbers (Dehaene, Bossini, & Giraux, 1993; Dehaene, Dupoux, & Mehler, 1990). Extensively replicated (for a review, see Wood, Willmes, Nuerk, & Fischer, 2008), the SNARC effect demonstrates horizontal mapping of numerical magnitude and is typically interpreted as behavioral evidence of a horizontal mental number line oriented from left to right (Fias, 1996). Recent research with adults demonstrated that number-to-space mapping extends from horizontal to vertical and sagittal axes (Aleotti, Di Girolamo, Massaccesi, & Priftis, 2020; for a review, see Winter, Matlock, Shaki, & Fischer, 2015). But it is not clear whether or not children, like adults, map number to space on multiple axes.

Evidence that preverbal infants (Bulf, de Hevia, & Macchi Cassia, 2016; de Hevia, Girelli, Addabbo, & Macchi Cassia, 2014), newborns at birth (de Hevia, Izard, Coubart, Spelke, & Sterri, 2014; de Hevia, Veggiotti, Sterri, & Bonn, 2017; Di Giorgio et al., 2019), preschool children (Hoffmann, Hornung, Martin, & Schiltz, 2013; Opfer & Thompson, 2006; Opfer, Thompson, & Furlong, 2010), and some non-human animals such as primates and chicks (Drucker & Brannon, 2014; Gazes et al., 2017; Rugani, Vallortigara, Priftis, & Regolin, 2015) already associate small numerosities to the left side and large numerosities to the right side of the space supports the view that the spatial representation of number is an experience-independent feature of human cognition likely stemming from early brain development (see reviews by de Hevia, Girelli, & Macchi Cassia, 2012; McCrink & Opfer, 2014). This cognitive predisposition is considered to be encouraged further by aspects of cultural experience, such as reading and writing habits (Dehaene et al., 1993; Pitt & Casasanto, 2014; Shaki, Fischer, & Göbel, 2012; but see Pitt & Casasanto, 2020, for an alternative account), that have been shown to modulate the direction of SNARC effects (Shaki, Fischer & Petrusic, 2009; Zebian, 2005). That is, if attention is oriented from left to right during reading, this visuospatial scanning habit is then co-opted into the domain of numerical cognition. Although representational convention contributes significantly to creating and sustaining representations of linear order, more generally orthography might not be the only experience-dependent determinant for a horizontally oriented mental number line. Pitt and Casasanto (2020) argued that reading habits cannot create an automatic spatial representation of numbers because space and number are not correlated in the experience of reading, whereas space and temporal order are, and this experiential relationship facilitates a mental timeline rather than a mental number line. Importantly, number–space mapping on axes other than horizontal, as described later, are not explained by left–right reading direction (for a discussion, see Ito & Hatta, 2004). Spatial thinking is linked to increased mathematical proficiency (for a review, see Hawes & Ansari, 2020). Some research supports a positive relationship between the strength of the horizontal SNARC effect and math skill in children (Georges, Hoffmann, & Schiltz, 2017; but see He et al., 2020). Indeed, effort has been made to design and evaluate mental number line training for children, with positive training effects reported for pairing full body movement along the horizontal axis with spatially locating the position of numbers (e.g., Fischer, Moeller, Bientzle, Cress, & Nuerk, 2011). This builds on an embodied view of the development of numerical cognition; higher levels of cognitive concepts, such as the representation of number magnitude, are grounded in sensor–bodily functions (Barsalou, 2008). Sensorimotor experience in vertical space may also play a crucial role in the development of abstract number concepts; there is a strong correlation between vertical space and numerical magnitude in the world—referred to as geocentric mapping—such as the concept of “more is up” that is frequently depicted in cultural products such as graphs (e.g., Barsalou, 2008). In addition, if children map number to space in depth—where smaller numbers are represented close to the body and larger numbers are
represented far away from the body—as well as on horizontal and vertical axes, this would indicate
that, even at an early stage of formal education, children can flexibly assign numerical magnitude
to any spatial dimension (see Holmes & Lourenco, 2012).

The SNARC effect exemplifies one instance of spatial–numeric associations. However, investiga-
tions into the developmental trajectory of the SNARC effect are sparse and the findings are equivocal,
often depending on the nature of the task. For example, using a parity judgment task in which mag-
nitude processing is implicit, Berch, Foley, Hill, and Ryan (1999) required participants to judge the par-
ity (i.e., is this number odd or even?) of a single presented digit by pressing one of two horizontally
aligned response keys. Based on this task, a SNARC effect is demonstrated by faster response times
using the left-side button to small versus large numbers and by faster response times using the
right-side button to large versus small numbers. The authors reported a horizontal SNARC effect in
children aged 9 years and over using this method. However, it is not clear whether children under
9 years of age readily understand the concept of parity. To address this issue, White, Szucs, and
Soltész (2012) employed a parity judgment task that was more amenable to young children and
reported a SNARC effect in 7.5-year-olds, although no evidence was found for a SNARC effect in 6-
year-olds, whereas Yang et al. (2014) found parity-based SNARC effects in Chinese children as young
as 5.8 years. Developmental trajectories have been reported for tasks in which magnitude is made
more explicit such as a single-digit magnitude comparison task where children are asked whether
the presented digit (from 1 to 9) is smaller or larger than 5. When magnitude was relevant to the task,
the SNARC effect was reported for 7-, 8-, and 9-year-olds (van Galen & Reitsma, 2008) but not for 5-
year-olds (Hoffmann et al., 2013), although there was weak evidence for a SNARC effect for 5.8-year-
olds. Gibson and Maurer (2016) used a similar magnitude comparison task to examine the SNARC
effect in 6- to 8-year-olds and reported a SNARC effect in 7- and 8-year-olds but not in 6-year-olds,
consistent with the findings of van Galen and Reitsma (2008). Using a novel color judgment task,
Hoffmann et al. (2013) reported a SNARC effect in children as young as 5.5 years, thereby lending sup-
port to the proposal that number-to-space mapping emerges early in ontogenetic development.

Together, these studies support the idea that a directionally oriented left–right mental number line
is present from a young age and that proficiency in magnitude tasks may account for the inconsistency
in results across studies that attempt to chart the developmental trajectory of SNARC effects.

Whereas many studies have examined spatial–numerical associations along the horizontal axis, the
investigation of spatial–numerical associations on other spatial axes, particularly vertical and depth
(sagittal), has received surprisingly little attention in both the adult and developmental literatures.
Some adult studies of numerosity suggest indirect associations between spatial extent on different
axes and numeric judgments (Fischer & Campens, 2009; Hartmann, Grabherr, & Mast, 2012; for a
review, see Winter et al., 2015). For example, Fischer and Campens (2009) reported that blindfolded
participants tended to spontaneously gesture to positions on either the horizontal or vertical axis
when instructed to point to numbers in space. Moreover, Hartmann et al. (2012) found that passive
movement via an electronic motion platform affected participants’ ability to randomly generate
numbers; participants tended to generate larger numbers when moved upward.

Table 1
Mean response times and standard deviations for compatible and incompatible trials across axes by age group and experiment.

<table>
<thead>
<tr>
<th></th>
<th>Horizontal</th>
<th>Vertical</th>
<th>Sagittal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Compatible Mean (sd)</td>
<td>Incompatible Mean (sd)</td>
<td>Compatible Mean (sd)</td>
</tr>
<tr>
<td>Experiment 1a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YC (5–7 yrs)</td>
<td>910 (229)</td>
<td>992 (226)</td>
<td>937 (244)</td>
</tr>
<tr>
<td>OC (8–9 yrs)</td>
<td>694 (131)</td>
<td>716 (119)</td>
<td>703 (312)</td>
</tr>
<tr>
<td>Experiment 1b</td>
<td>813 (210)</td>
<td>891 (283)</td>
<td>796 (272)</td>
</tr>
</tbody>
</table>

Note. Means and standard deviations are reported in milliseconds. For Experiment 1b, response times are collapsed across age.
YC, young children; OC, older children.
Evidence from the SNARC paradigm suggests that vertical number–space mapping occurs in adults; small and large numbers are more readily associated with down and up, respectively (Gevers, Lammertyn, Notebaert, Verguts, & Fias, 2006; Ito & Hatta, 2004; Müller & Schwarz, 2007; Shaki & Fischer, 2012). However, there is a lack of clarity from these studies on the nature of the axes tested. For example, response buttons were typically positioned on a table along the participant’s mid-sagittal line (i.e., parallel with the tabletop) and therefore not positioned according to the true vertical axis. As such, the response keys representing up/down would be more accurately defined as lying along the depth plane and therefore at far/near distance from the participant. Using a vertical alignment of response keys (i.e., one on top of the other), Hartmann, Gashaj, Stahnke, and Mast (2014) reported the first true vertical SNARC effect in adults that was primarily driven by faster response times to small numbers with lower hand responses. Similarly, Holmes and Lourenco (2012) reported vertical number–space mapping when magnitude information was made explicit (i.e., when participants were instructed to consider each number as a different floor, ordered from the bottom–up of a building). More recently, Aleotti et al. (2020) employed a parity judgment task and an apparatus composed of three orthogonal bars that facilitated manual responses in horizontal, vertical, and sagittal space; they found equally strong SNARC effects on all three axes in adults. These findings suggest that adults have a three-dimensional mental representation of numbers.

However, little is known about whether spatial–numerical associations exist on multiple axes during childhood. There is evidence to support the early emergence of a left-to-right-oriented mental number line in children as young as 3 years (Opfer & Thompson, 2006; Opfer et al., 2010; Patro & Haman, 2012). In addition, 4- to 7-year-olds demonstrate comparably accurate placement of number to position on both horizontally and vertically oriented scales (Simms, Muldoon & Towse, 2013), and magnitude-relevant horizontal SNARC effects have been detected in children as young as 5–7 years (Gibson & Maurer, 2016; Hoffmann et al., 2013; van Galen & Reitsma, 2008). If children exhibit SNARC effects that extend beyond the horizontal axis, this implies that spatial–numerical associations are flexible and cannot be fully explained as deriving from experiences with cultural practices such as reading and writing. This would have important implications for how numeracy is instructed in primary school education and for the design and development of mathematical interventions that aim to exploit the embodied nature of spatial–numerical associations.

**Experiment 1a**

To address whether spatial–numerical mappings generalize to other axes, and whether these mappings exist during childhood, we employed a magnitude-relevant task to examine horizontal, vertical, and sagittal (depth) number–space mappings in two age groups of children.

**Method**

**Participants**

A total of 58 children (29 boys and 29 girls) with an average age of 7.25 years (SD = 1.43, range = 5–9) took part in the experiment, with all children being recruited from two primary schools within the city of Dublin, Ireland. All participants self-reported having normal or corrected-to-normal vision, and 7 participants self-reported as being left-handed for writing. The experiment was approved by the research ethics committee of the School of Psychology at Trinity College Dublin, and in accordance with the Declaration of Helsinki, all children’s parents or guardians provided written informed consent prior to the experiment. Both children and their parents/guardians were advised of the right to withdraw from the experiment at any time without prejudice. Each child received a small toy for his or her participation.

By 5-6 years of age, children attend school within the education system at a grade known as “senior infants”, which is akin to kindergarten (USA). At this age children have already been exposed to Arabic numerals, can count with ease from 1 to 10, and should competently perform a magnitude comparison.
task with digits 1 to 9 (see National Council for Curriculum and Assessment, 1999). Given the expected proficiency at magnitude comparison and familiarity with Arabic numerals, children were divided into two groups according to their age at the time of testing. The first group was the young children (YC) group comprising 31 children (17 boys and 14 girls) with an average age of 6.03 years (SD = 0.60, range = 5–7). This group included 23 children from senior infants (9 boys and 14 girls) and 8 children (all boys) from first class (akin to first grade USA). The second group was the older children (OC) comprising 27 children from third class (12 boys and 15 girls) with an average age of 8.67 years (SD = 0.48, range = 8–9). Counting abilities were confirmed in each child prior to testing by the experimenter; each child was required to count to 10. All children in the experiment could count to 10 with ease.

Materials

All testing occurred during the school year at Months 4 and 5 of the school calendar. Testing took place in a quiet classroom on the premises of each school. The computerized task was programmed in Presentation and administered using a Dell laptop with a 38-cm color monitor (1366 × 768 resolution). The laptop keyboard was not used and was covered with a piece of cloth. A wireless rectangular keyboard was used to record children’s responses and measured 2.5 cm in height, 12 cm in width, and 28 cm in length. Black cardboard was cut to size, laminated, and fixed to the keyboard, covering all but two response keys, located equidistant from the center of the keyboard, 21 cm apart. The wireless keyboard was mounted on a wooden board. To accommodate vertical positioning of the wireless keyboard, a wooden platform, 2.5 cm in height, was constructed. A clamp was fixed to a small board and attached to the wooden platform. A small rebate cut in the mounting board of the wireless keyboard accommodated the clamp without obstructing the face of the wireless keyboard.

The laptop was raised so that the screen was positioned at children’s eye level. Participants were seated 60 cm from the screen and were required to complete a single-digit comparison task in which numerical magnitude was made explicit; they were asked to decide whether the number on the screen was bigger or smaller than 5 (adapted from van Galen & Reitsma, 2008). Arabic numerals 1, 2, 3, 4, 6, 7, 8, and 9 served as stimuli, and digit number 5 served as the constant comparison value. Participants were told to respond as fast as possible and made their response by pressing one of two response keys aligned along the horizontal, vertical, or sagittal axis, depending on the experimental condition. Two stickers, one with a small circle and the other with a large circle, were placed beside the response keys as a visual prompt for “smaller” and “larger” than 5, respectively.

In the horizontal condition, the response box was positioned horizontally, with the two response keys located to the left and right of the body midline. In the vertical condition, the response box was fixed to the platform and positioned vertically, with the response buttons located above and below the participant’s midpoint on his or her chest. In the sagittal condition, the response box extended away from the body, aligned with the body midline, with response buttons located near or far relative to the participant’s body (see Fig. 1).

Design and procedure

For all conditions in this experiment, participants used their left hand to press the left response button (horizontal axis), lower button (vertical axis), or near button (sagittal axis) and used their right hand to press the right button (horizontal axis), upper button (vertical axis), or far button (sagittal axis). On spatially compatible trials—when the target number was less than 5—the corresponding response key was the left (horizontal axis), lower (vertical axis), or near (sagittal axis) response button. In contrast, on spatially incompatible trials, the location of the corresponding response buttons for small target numbers was reversed, such that the right (horizontal), upper (vertical), and far (depth) response buttons represented smaller magnitudes. The other response buttons (left, lower, and near for each axis, respectively) were used for target numbers greater than 5.

Each trial began with a circular placeholder (black bordered on a white background, subtending approximately 3 degrees of visual angle) that was presented in the middle of the screen. After 1000 ms, a number digit appeared at the center of the placeholder. Each target number was presented in black Arial 96-point font, subtending approximately 3 degrees of visual angle. The number stimulus disappeared once a response was made or after 5000 ms had elapsed. An intertrial interval followed a response, in which a blank screen was presented for a duration of 1000 ms. Number digits were
presented in a pseudorandom order, such that the same digit did not appear in consecutive trials. Response times were recorded and used as a measure of spatial–numerical associations.

The experiment comprised 240 trials, with 80 trials tested per axis, and it was completed over 2 days. For each axis, there were two testing blocks (compatible and incompatible conditions) of 40 trials each, and participants could take a self-timed break after 20 trials were completed. Each digit (1, 2, 3, 4, 6, 7, 8, or 9) was presented five times per block. To avoid any possible confusion across the tasks, compatible and incompatible blocks for the same axis were completed on separate days. Thus, children completed 120 trials per day, with 40 trials tested in each of the three axes (horizontal, vertical, and sagittal). The presentation order of blocked trials for each of the axes and response mode (compatible or incompatible) were counterbalanced across participants. Children completed 12 practice trials at the beginning of each block.

Results

Data were analyzed in R (R Development Core Team, 2010) using analyses of variance (ANOVAs). Greenhouse–Geisser corrections were used when Mauchly's test for sphericity was significant, and effect sizes are presented as partial eta squared ($\eta^2_p$). We used two commonly employed methods to examine the SNARC effect: (a) the ANOVA approach, where the SNARC effect is supported by faster response times on spatially compatible trials compared with spatially incompatible trials (Gevers

![Fig. 1. Illustration of the experimental apparatus and setup. Top panel: Hand (spatio-anatomical) and response button location (extracorporeal) mapping in Experiment 1a for spatially compatible trials across each of the horizontal (left of figure), vertical (center), and depth (right of figure) axes. Bottom panel: On incompatible trials, the response code positions (<5 and > 5) on the apparatus were switched; hands remained in same positions. Gray oval depicts the back view of the vertical apparatus.](image)
et al., 2006; Imbo, De Brauwer, Fias, & Gevers, 2012), and (b) the t test of regression coefficients of the difference in response times (dRTs) to the different spatial positions across number magnitude (Fias, 1996; van Galen & Reitsma, 2008). In addition, Bayesian hypothesis testing was performed in JASP (JASP Team, 2019). Bayes factors range from 0 to \( \infty \), where a Bayes factor of 1 indicates that both hypotheses predict the data equally well. A BF\(_{10} > 100\) yields extreme evidence for the alternative hypothesis, and a BF\(_{10} < 1/100\) yields extreme support for the null hypothesis (see Wagenmakers et al., 2018, for classification of values between 1/100 and 100).

Response times of less than 200 ms (which indicated anticipation errors; 0.38% of responses) and outliers (1.2%), identified through visual analysis of box plots with outliers labeled for each condition, were not included in the final data set. Consistent with previous studies, overall error rates were low (8.5% of responses). The YC group made more errors (5.86% of responses) than the OC group (2.64% of responses). The error rates were comparable across all three axes (horizontal: 2.87%; vertical: 3.02%; sagittal: 2.61%) and relatively low. Error data were not further analyzed. For the correct responses, the median response times were calculated for each participant and subjected to statistical analyses (Table 1).

As a first step, we ran a 3 \( \times \) 2 \( \times \) 2 mixed-model ANOVA with axis (horizontal, vertical, or sagittal) and spatial compatibility (compatible or incompatible) as within-participant factors and age group (younger or older) as the between-participants factor. A main effect of spatial compatibility, in which response times were faster to spatially compatible versus incompatible trials, is indicative of a SNARC effect.

The main ANOVA yielded significant main effects for all three factors. For the effect of age group, \( F (1, 56) = 26.95, p < .001, \eta^2_p = .33, \text{BF}_{10} = 2195\), the older children (\( M = 716 \text{ ms}, SD = 137 \)) responded significantly faster than the younger children (\( M = 953 \text{ ms}, SD = 202 \)). For the effect of spatial compatibility, \( F (1, 56) = 12.29, p < .001, \eta^2_p = .18, \text{BF}_{10} = 10828\), children responded faster to the spatially compatible trials (\( M = 815 \text{ ms}, SD = 206 \)) than to the spatially incompatible trials (\( M = 886 \text{ ms}, SD = 241 \)). Although the main effect for age was significant, \( F (1,78, 99.72) = 4.00, p = .025, \eta^2_p = .07, \text{BF}_{10} = 0.174\), the post hoc pairwise comparisons did not survive Bonferroni correction (\( \alpha = .05/3 = .017 \)) and, as indicated above, the Bayesian ANOVA yielded a BF of less than 1, which is indicative of support for the null hypothesis. Therefore, the response times across all three axes were equivalent (all \( ps > .03 \)). None of the interactions was significant (all \( ps > .32 \), all BF\(_{10} < 1\)). In particular, there was no evidence for a three-way interaction, \( F (2, 112) = 0.29, p = .75, \eta^2_p = .005, \text{BF}_{10} = .001\), suggesting that the SNARC effect was present on all axes and did not differ by age group.\(^2\) The relative decrease in response times to the spatially compatible trials over the incompatible trials is plotted in Fig. 2; negative values indicate faster response times to spatially compatible trials. All three axes show a SNARC effect when collapsed across age group (see Fig. 2). An additional repeated-measures ANOVA examining the effect of age on the “compatibility effect,” \( F (2, 114) = 1.06, p = .35, \eta^2_p = .02, \text{BF}_{10} = 0.146\), further confirmed that the strength of the SNARC effect was comparable across all three axes.

The primary analyses revealed that the spatial compatibility effects found on each axis were not modulated by the age group of participants; therefore, for the purpose of the following analyses, the data from all participants were pooled. First, on each axis, we calculated each participant’s median response time for each number separately provided that the response was correct. The dRTs were then calculated by subtracting each participant’s median response time to the left, lower, and near response buttons from the participant’s median response time to the right, upper, and far response buttons for each number digit along each of the axes. For example, for the responses along the horizontal axes, we subtracted the response times to the left response button from those to the right response button (and similarly for the vertical axis [upper – lower] and sagittal axis [far – near]). If there was an association between response position and number magnitude, we would expect an increase in numerical magnitude to be associated with an increasingly negative dRT because responses to the right, upper, and far positions should be faster to relatively large number magnitudes, whereas responses to the left, lower, and near positions should be faster to small number magnitudes.

\(^2\) Separate two-way analyses of variance on response times revealed main effects of spatial compatibility and age to each of the individual axes, with no evidence of an interaction between these factors (see online supplementary material).
Regression lines were fit separately for each axis, with digits (1–9 without 5) entered as the predictor and dRT as the dependent variable. Then, to contrast the SNARC effects between axes, we extracted participant-wise regression weights for each axis (thereby identifying the SNARC effect at the individual level), and these results were included as the dependent variable in one-sample t tests. Linear regression for dRTs are graphed for each axis in Fig. 3.

The expected negative slope between number magnitude and dRT deviated significantly from zero on each of the horizontal, vertical, and sagittal axes (see Table 2). On the horizontal axis, the linear regression analysis yielded a regression weight of \(-23.80\). A one-sample t test on the extracted individual coefficients showed that the regression weight of the horizontal axis dRT slope significantly differed from zero, \(t(57) = -2.96, p = .002, d = 0.49, BF_{10} = 14.36\), thereby supporting a SNARC effect. For the vertical axis, the linear regression analysis yielded a regression weight of \(-37.57\). A one-tailed t test also confirmed the presence of a SNARC effect; the regression weight of the vertical axis slope significantly differed from zero, \(t(57) = -3.41, p = .001, d = 0.45, BF_{10} = 45.57\). Finally, for the sagittal axis, the linear regression analysis yielded a regression weight of \(-19.34\). As with the other axes, a one-sample t test showed that the extracted slope significantly differed from zero by conventional standards, but with a BF close to 1 implying only anecdotal evidence, \(t(57) = -1.87, p = .03, d = 0.25, BF_{10} = 1.404\). To compare the strength of the SNARC effect across the three axes, we used one-tailed pairwise comparisons between the individual participant coefficients (\(\infty = .05/3 = .016\)). Participants’ coefficients did not significantly differ between the horizontal and vertical axes, \(t(57) = 1.417, p = .162, d = 0.19, BF_{10} = 0.369\), between the horizontal and sagittal axes, \(t(57) = -.392, p = .697, d = 0.05, BF_{10} = 0.154\), or between the vertical and sagittal axes, \(t(57) = -1.57, p = .121, d = 0.21, BF_{10} = 0.459\). In sum, the results of the linear regression analyses are consistent with those from the mixed ANOVA and provide additional support of spatial–numerical associations on all three axes in children aged 5 to 9 years.

**Experiment 1b**

In Experiment 1a, we found evidence of spatial–numerical associations on the horizontal, vertical, and sagittal axes with no difference in the magnitudes of the SNARC effect across the different age groups. However, both the spatio-anatomical information (i.e., position of the hand relative to the body) and the extracorporeal information (i.e., side of space) were constantly aligned. For example, the left hand was associated with lower numbers when positioned in left, lower, or near space (see Fig. 1); therefore, it remains unclear whether the spatio-anatomical or extracorporeal mapping of magnitude influenced the SNARC effect.
Both spatio-anatomical (hand-based) and extracorporeal (location) hypotheses have previously been proposed to account for the horizontal SNARC effect. Initially, because a SNARC effect was observed even when participants crossed their hands, such that their left hand was now positioned in the right side of space and their right hand was positioned in the left side of space (Dehaene et al., 1993; Müller & Schwarz, 2007; for a review, see Viarouge, Hubbard, & Dehaene, 2014), the effect was explained as numbers mapped to a spatial reference frame external to the body. This is because crossing the hands induces a conflict between body-centered and external coordinate spatial reference frames. If the effect persists, even with this conflict, it is the position of the response codes in space that is driving the effect.

However, evidence that extracorporeal spatial information has a greater influence over anatomical information on the SNARC effect has not been consistent. Several studies have provided evidence supporting the hand-based hypothesis, indicating an association between the side of the body and numerical magnitude. For example, Wood, Nuerk, and Willmes (2006) found that the SNARC effect was significantly reduced when the hands were crossed, suggesting that anatomical maps also likely contribute to the SNARC effect. Other studies have also highlighted the importance of body-centered reference frames in the SNARC effect, with a variety of effectors such as eye movements (Fischer, Castel, Dodd, & Pratt, 2003) and hand movements (Schwarz & Keus, 2004) being known to influence the SNARC effect, suggesting that the lateralization of participants’ effectors is a necessary condition for the horizontal SNARC effect.

In Experiment 1b, we switched the position of the response hands along each axis but maintained the same numerical magnitude mappings as in Experiment 1a. In this way, we produced a condition of conflict between extracorporeal space and spatio-anatomical mapping for horizontal and vertical axes, and different predictions can be made on the outcome. First, if numerical magnitude is mainly mapped

<table>
<thead>
<tr>
<th></th>
<th>Horizontal</th>
<th>Vertical</th>
<th>Sagittal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta$</td>
<td>$R^2$</td>
<td>$t$</td>
</tr>
<tr>
<td>Experiment 1a</td>
<td>-23.81</td>
<td>.21</td>
<td>-4.62</td>
</tr>
<tr>
<td>Experiment 1b</td>
<td>-19.26</td>
<td>.02</td>
<td>-2.10</td>
</tr>
</tbody>
</table>
to location in extracorporeal space, such that the position of the hands would not influence the SNARC effect, then we should see evidence for SNARC effects—faster response times to spatially compatible trials than to spatially incompatible trials. Alternatively, if the SNARC effect is driven by spatio-anatomical mapping, then evidence for a SNARC effect would be found only when the numerical magnitude is compatible with the responding hand—faster response times on anatomically compatible trials compared with anatomically incompatible trials. Finally, if numerical magnitude is mapped to a combination of extracorporeal location and the effector body site, then we would expect that the mismatch between these influences (on either spatially compatible or spatially incompatible trials) should result in the elimination of, or considerable reduction in, the SNARC effect. This cannot be applied to the sagittal axis because the hands are either near or far from the body, thereby precluding a dissociation of extracorporeal and spatio-anatomical locations. However, for consistency with Experiment 1a, all axes were tested, with hand position swapped on the sagittal axis.

Method

Participants

Because the spatial compatibility effects detected in Experiment 1a were not modulated by age, Experiment 1b was not designed to examine developmental trends. A total of 26 children participated in Experiment 1b (9 boys and 17 girls; \(M_{\text{age}} = 7.59\) years, \(SD = 0.97\), range = 6–9). Data from 2 participants could not be used due to an error in the experimental setup. Participants were recruited and randomly selected from First and Third Class of two primary schools within the city of Dublin. All children self-reported normal or corrected-to-normal vision, and 4 participants self-reported as being left-handed for writing. None of the children took part in the previous experiment, and all children were naïve to the purpose of the task. The study was approved by the research ethics committee of the School of Psychology at Trinity College Dublin, and in accordance with the Declaration of Helsinki, all participants’ parents or guardians provided written informed consent in advance of the study. All children were advised of their right to withdraw from the study at any time without prejudice. Each child received a small toy for his or her participation.

Design and procedure

The children were tested at the end of the school year at Month 9 of the academic year. Testing took place in a quiet classroom space on the premises of each school. The design and procedure were identical to those of Experiment 1a with the exception that here we swapped the location of the hands on each axis (see Fig. 4).

Results

Errors were made on 10.03% of trials (horizontal: 3.21%; vertical: 3.31%; sagittal: 3.51%) and were not further analyzed. Response times less than 200 ms (i.e., representing anticipation errors; fewer than 0.63% of trials) were removed prior to analysis. Because there was a wider age range of children in Experiment 1b (6- to 9-year-olds) than in Experiment 1a (5- to 7-year-olds and 8- and 9-year-olds), and because we were not looking at developmental trends, a more stringent filtering process was applied to detect outliers and to account for intra-individual variability in response times (Eckert & Eichorn, 1977; Whelan, 2008). Response times 2.5 standard deviations from the participant’s mean in each condition were removed from the data set prior to analysis, accounting for less than 3.5% of the data. For trials to which each participant correctly responded, the median response times were calculated and subjected to statistical analyses. Given the smaller sample size, to investigate potential insensitivity to effects, we also report Bayesian alongside frequentist statistics. In cases where null findings occurred, we were able to determine cases where comparisons may have been underpowered versus convincing evidence for the null. In addition, we estimated the internal consistency measure of reliability of the SNARC and employed a permutation-based split-half reliability approach as recommended by Parsons, Kruijt, and Fox (2019). Subsequently, we applied the Spearman–Brown adjustment to obtain the reliability estimate for the whole set of items (see Table 3). All Spearman–Brown reliability estimates are > .70, indicating moderate to high internal reliability.
A 3 × 2 repeated-measures ANOVA with axis (horizontal, vertical, or sagittal) and spatial compatibility (compatible or incompatible) as within-participant factors revealed no main effect of axis, $F(2, 46) = 2.80, p = .07, \eta^2_p = .11, 90\% \text{ CI} [.00, .24], BF_{10} = 0.429$, or of spatial compatibility, $F(1, 23) = 2.02, p = .17, \eta^2_p = .08, 90\% \text{ CI} [.00, .24], BF_{10} = 0.497$, and found no interaction between axis and spatial compatibility, $F(2, 46) = 1.51, p = .24, \eta^2_p = .06, 90\% \text{ CI} [.00, .11], BF_{10} = 0.177$. Note that the interaction term BF is < 1/3, indicating moderate evidence for the null hypothesis, with BFs for the main effects > 1/3 indicating only anecdotal evidence for the null hypothesis.

Although the main effect of spatial compatibility was not statistically significant, it was important to check whether there was any convincing evidence of a spatial compatibility effect separately for each axis. Separate analyses were conducted on the response times to each axis with spatial compatibility (compatible or incompatible) as the only within-participant factor. On the horizontal axis, there was a main effect of spatial compatibility, $F(1, 23) = 4.73, p = .04, \eta^2_p = .17, 90\% \text{ CI} [.00, .38], BF_{10} = 0.605$, with faster response times to the spatially compatible condition ($M = 813 \text{ ms}, SD = 210$) than to the spatially incompatible condition ($M = 891 \text{ ms}, SD = 283$). However, the BF was close to 1, showing anecdotal evidence for the null hypothesis. No effect of spatial compatibility was detected on either the vertical axis, $F(1, 23) = 1.11, p = .30, \eta^2_p = .05, 90\% \text{ CI} [.00, .23], BF_{10} = 0.438$, or the sagittal axis, $F(1, 23) = 0.07, p = .80, \eta^2_p = .003, 90\% \text{ CI} [.00, .10], BF_{10} = 0.344$.

Fig. 5 plots the mean dRTs between spatially compatible and incompatible trials on each axis separately. From Fig. 5 (horizontal axis), whereas the 90% confidence intervals for the effect size does not include zero, the confidence intervals for the mean difference touch the zero line, indicating that the SNARC effect is small. In addition, the response time difference between conditions is comparable across axes, $F(2, 46) = 1.51, p = .23, \eta^2_p = .06, BF_{10} = 0.403$.

In line with Experiment 1a, we also tested the SNARC effect by means of linear regression analyses. To be consistent with the analyses of Experiment 1a, the dRT scores were calculated by subtracting the median response time to the left, lower, and near response buttons from the median response time to the right, upper, and far response buttons, respectively.

![Fig. 4. Illustration of the experimental apparatus. Both the spatio-anatomical hand position and position of response codes (extracorporeal spatial information) are shown for the magnitude comparison task in Experiment 1b.](image-url)
The individual participant-wise extracted regression weights show no statistically significant difference from zero on the vertical axis, $t(23) = -1.55$, $p = .07$, $d = 0.32$, $BF_{10} = 1.13$, horizontal axis, $t(23) = -1.60$, $p = .06$, $d = 0.33$, $BF_{10} = 1.22$, or sagittal axis, $t(23) = 0.68$, $p = .75$, $d = 0.14$, $BF_{10} = 0.139$. BFs for effects on the vertical and horizontal axes are anecdotal, whereas the BF for the sagittal axis shows moderate support for the null hypothesis. From this, there is no convincing evidence of SNARC effects on any axis.

**Discussion**

No main effect of spatial compatibility was found in the omnibus axis by spatial compatibility ANOVA, indicating that there was no overall SNARC effect. This is further supported by no interaction between axis and spatial compatibility. Furthermore, when spatial compatibility was examined separately for each axis, the significant effect on the horizontal axis should be interpreted with caution because the Bayesian ANOVA indicates that the evidence for this is weak, with a $BF < 1$. We do not make any conclusions where only anecdotal support for the alternative or null hypothesis is present.

On the vertical axis, there was no evidence for a benefit of spatial compatibility over spatial incompatibility on response times; however, the results of the linear model provide anecdotal evidence that, overall, participants responded faster to numbers less than 5 when the response button was in the upper position and to numbers greater than 5 when the response button was in the lower position (irrespective of hand position) (see Fig. 6). Still, this finding should be interpreted carefully because the individual coefficients show no convincing statistical evidence for the horizontal or vertical axes. The results of the responses to the sagittal axis were consistent; no SNARC effect was evident from either analysis.
General discussion

The main goal of this study was to examine whether children spatially represent number in three dimensions. In Experiment 1a, we reported, for the first time, SNARC effects on three axes in children. Whereas older participants (8 and 9 years) were faster than younger participants (5–7 years) at the task, the SNARC effect, as defined by relatively faster response times on spatially compatible trials versus spatially incompatible trials, was not modulated by age group. As a follow-up, Experiment 1b explored the influence of body-based (spatio-anatomical) spatial reference frames and external to the body (extracorporeal) reference frames on these multidimensional number–space mappings.

Horizontal axis

Our initial finding of a magnitude-relevant SNARC effect on the horizontal axis in children is consistent with previous reports from the developmental literature (Chan & Wong, 2016; Crollen, Dormal, Seron, Lepore, & Collignon, 2013; Gibson & Maurer, 2016; Nava, Rinaldi, Bulf, & Macchi Cassia, 2017; van Galen & Reitsma, 2008). Thus, our study provides further support for a left-to-right-oriented mental number line in children. When the same magnitude comparison task was given, but children were required to cross their hands, we did not find convincing evidence of a SNARC effect on the horizontal axis.

In adults, the horizontal mental number line, as indicated by SNARC effects, has been shown to rely on both body-based (e.g., Wood et al., 2006) and external (object-based) spatial coordinate systems (Dehaene et al., 1993; Müller & Schwarz, 2007). Previous work has explored the relative importance of body-centered versus external spatial reference frames in different populations, including children with high and low visuospatial abilities (Crollen et al., 2013; Crollen & Noël, 2015; Nava et al., 2017). There is some evidence of a developmental shift in the spatial reference frames employed by children when mapping number to space. By examining the spatial representation of number in 5- and 6-year-old children under different conditions (vision or no vision) and hand posture (crossed or uncrossed), a recent study showed that preschool-aged children use both spatio-anatomical and extracorporeal frames of reference for representing number in space (Nava et al., 2017). On the horizontal axis, 5- and 6-year-olds show SNARC effects when hands are uncrossed and when visual input is available (Nava et al., 2017). The results from Experiment 1a are in line with those reported by Nava et al. (2017). This is particularly important when we consider that, combined with the results of Nava et al. (2017), there is evidence of a developmental shift toward a purely extracorporeal referencing of number to space, with 10-year-olds demonstrating horizontal SNARC effects irrespective of...
whether their hands are crossed or uncrossed (Crollen & Noël, 2015). Because the group in Experiment 1b included children aged 6 to 9 years, this shift toward an extracorporeal representation of numbers in space may begin to emerge before 10 years of age, and the wider age range may explain why we found only partial evidence for a SNARC effect with crossed hands on the horizontal axis.

Vertical axis

In line with research with adults (Aleotti et al., 2020; Gevers et al., 2006; Ito & Hatta, 2004; Müller & Schwarz, 2007; Shaki & Fischer, 2012), we found that spatial–numerical associations extend beyond the horizontal axis to the vertical axis in children. What is the source of magnitude space mapping along the vertical plane? One speculation is that these associations are learned from the environment; for example, many products that depict quantity conform to the idea that “more is up,” including thermometers, graphical depictions of battery power, price fluctuations, and height-measuring tools. Moreover, our linguistic descriptions of quantity are often rooted in vertical space (e.g., “high” number, “low” number). From this perspective, the vertical spatial association of magnitude may be based on sensorimotor contingencies established at the time of concept acquisition (Barsalou, 2008; Lakoff & Johnson, 1980). Lakoff and Johnson (1980) first highlighted the example of adding water to a glass as a learned contingency between an increase in level and “more” as it relates to water level rising and less associated with a reduction in volume. Similarly, gravity may establish the ground as origin, with everything extending upward as increments from zero, which functions to cognitively link magnitude with vertical linear extent (Holmes & Lourenco, 2012). Likewise, children perceive themselves as growing “bigger” by measuring their height. Our finding of a vertical SNARC effect in 6- to 9-year-olds in Experiment 1a supports the vertical spatial representation of numbers—evident even from the beginning of formal education. This makes sense, when one considers that, in addition to the cultural products and sensorimotor contingencies mentioned above that emphasize vertical space and numerical magnitude, children’s learning environments from the beginning of preschool contain horizontal and vertical presentations of number lines on classroom walls and in textbooks.

In contrast to this study, Hartmann et al. (2014, Experiment 1) found a vertical association of small numbers with the lower response position in adults that was not modulated by anatomic response side (whether participants’ right or left hand was placed at the lower response position); we found no spatial compatibility effect on the vertical axis when hand position was swapped in Experiment 1b. Moreover, there is other research with children suggesting that spatial–numerical associations on the vertical axis may be less common. For example, Aulet and Lourenco (2018) showed no vertical effect in children with their “number memory” task. The source of vertical number–space magnitude mapping is likely a complex interplay of cultural and linguistic conventions that children are exposed to from an early age combined with how verticality and quantity are represented in the natural world (see Winter et al., 2015).

Some authors have suggested that the vertical spatial association of numerical magnitude should be more salient than horizontal associations because it is universal and instantiated a priori to the cultural habits of reading and writing (Fischer, 2012; Shaki & Fischer, 2018; Sixtus, Lonnemann, Fischer, & Werner, 2019). However, the data from the current study are not consistent with this interpretation given that no difference in SNARC effect between axes was found in Experiment 1a. This finding extends the recent finding in adults of SNARC effects of comparable strength on all three axes (Aleotti et al., 2020) and supports the existence of three independent mental number lines. Furthermore, the results point toward a three-dimensional (3-D) representation of number in space that has emerged by early school age. Future research could examine whether or not these 3-D mental number lines are facilitated by cultural depictions of number, quantity, and space—such as those commonly found in classrooms during early school years—and whether these representations exist a priori to such exposure.

Sagittal axis

The results of Experiment 1a suggest that children map numerical magnitude along the sagittal dimension from near (small) to far (large). Sagittal mapping can be understood in terms of
distance-based magnitude mapping from one’s body; nearby space is a small magnitude away from
the observer and more readily maps onto small numbers, whereas far space is a large magnitude away
from the observer and more readily maps onto large numbers (for a discussion, see Winter et al.,
2015). When presented with near and far response locations, observers flexibly assign magnitude to
these spatial dimensions based on the spatial heuristic that large numerical magnitudes represent
“more” and that “more” corresponds to greater distance from one’s body. Sagittal mapping of number
to space may be constrained to experimental conditions that are preferable for egocentric spatial ref-
erence frames—as in Experiment 1a. It is important to recognize that the sagittal axis does not strictly
permit an extracorporeal spatial reference frame given that the position of the response codes is
always relative to the body of the observer rather than to the left or right (horizontal) or at the top
or bottom (vertical) of the response apparatus.

The different results from Experiments 1a and 1b tell us that hand position influences the sagittal
SNARC effect for children. This is in contrast to the finding of (Chen, Zhou, & Yeh, 2015), where adults
demonstrated a SNARC effect in a parity judgment task on the near–far plane—event when hand posi-
tion had been counterbalanced. Task differences may explain the contradictory findings between the
two studies; Chen et al. (2015) used a magnitude–irrelevant task, whereas the task in the current study
made magnitude explicit. An alternative explanation is that during childhood the left–right repres-
ation of number and space is stronger than near/small–far/large and switching hand position reduces
or eliminates the effect. However, if this were the case, we would see an interaction between spatial
compatibility and axis, which we did not; this is in line with the recent findings in adults (Aleotti et al.,
2020).

Cognitive mechanisms other than the visuospatial account of the mental number line have been
proposed to account for SNARC effects; the verbal–spatial account (Gevers et al., 2010; Proctor &
Cho, 2006) posits that associations are derived from verbal coding of space such as associations
between small/left and large/right. In the context of the current study, the verbal association can also
apply to small/down and large/up. However, near and far are not verbally coded by small and large,
respectively, and as such cannot explain the findings on the sagittal axis.

Related to the verbal–spatial account is the principle of polarity correspondence (Proctor & Cho,
2006). Here, instead of SNARC effects reflecting a mental number line, they can be understood in terms
of stimulus–response compatibility. The response dimensions on each axis are binary, and according
to polarity correspondence one element of each pair is preferred. “Right” and “up” occur more fre-
quently and therefore have a positive polarity—because linguistically they are assumed to occur more
frequently—with “left” and “down” allocated a negative polarity. However, linguistic frequency is not
clear with respect to near and far on the sagittal axis, and as such polarity correspondence cannot
explain SNARC effects on this axis, but it is an explanatory candidate for SNARC effects on the horizon-
tal and vertical axes. When response–stimulus compatibility is controlled, research has shown that
adult participants make magnitude-driven shifts of attention on the horizontal axis but not on the ver-
tical axis (Holmes, Ayzenberg, & Louengo, 2016). Future research that investigates whether or not chil-
dren make such spontaneous magnitude-driven eye movements, on all three axes, would help to
elucidate the cognitive mechanisms that account for SNARC effects.

A limitation of the current study is the broad age range and smaller sample size in Experiment 1b,
which restricts the examination of the developmental trajectory of SNARC effects. Examining develop-
mental trends was not prioritized in Experiment 1b due to the finding from Experiment 1a that SNARC
effects were not modulated by age. Further research is needed to identify when, in children under
10 years of age, a predominantly extracorporeal reference frame emerges on the horizontal axis and
whether this is the same for the vertical axis.

Conclusion

This is the first study to report magnitude-relevant SNARC effects in children on three spatial axes.
The finding that children associate numerical magnitude in terms of height (vertical) and distance
(sagittal) supports the idea that the spatial representation of number is not limited to a dominant glo-
bal left–right orientation. Although future research will need to tease apart the relative contribution of
spatial reference frames to the emergence of spatial–numerical associations across multiple axes in
young children and infants, our findings provide novel insights into the existence of 3-D mental number lines during development.

Acknowledgment

This project received funding from the European Union’s Horizon 2020 research and innovation program (732391) weDRAW - Exploiting the best sensory modality for learning arithmetic and geometrical concepts based on multisensory interactive Information and Communication Technologies and serious games.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2021.105169.

References


JASP Team. (2019). JASP (Version 10.2) [Computer software].


