# Number-related Brain Potentials Are Differentially Affected by Mapping Novel Symbols on Small versus Large Quantities in a Number Learning Task

Fabian C. G. van den Berg, Peter de Weerd, and Lisa M. Jonkman

## Abstract

■ The nature of the mapping process that imbues number symbols with their numerical meaning—known as the "symbol-grounding process"—remains poorly understood and the topic of much debate. The aim of this study was to enhance insight into how the nonsymbolic–symbolic number mapping process and its neurocognitive correlates might differ between small (1–4; subitizing range) and larger (6–9) numerical ranges. Hereto, 22 young adults performed a learning task in which novel symbols acquired numerical meaning by mapping them onto nonsymbolic magnitudes presented as dot arrays (range 1–9). Learning-dependent changes in accuracy and RT provided evidence for successful novel symbol quantity mapping in the subitizing (1–4) range only. Corroborating these behavioral results, the number processing related P2p component was only modulated by the learning/ mapping of symbols representing small numbers 1–4. The symbolic N1 amplitude increased with learning independent of symbolic numerical range but dependent on the set size of the preceding dot array; it only occurred when mapping on one to four item dot arrays that allow for quick retrieval of a numeric value, on the basis of which, with learning, one could predict the upcoming symbol causing perceptual expectancy violation when observing a different symbol. These combined results suggest that exact nonsymbolic–symbolic mapping is only successful for small quantities 1–4 from which one can readily extract cardinality. Furthermore, we suggest that the P2p reflects the processing stage of first access to or retrieval of numeric codes and might in future studies be used as a neural correlate of nonsymbolic–symbolic mapping/symbol learning. ■

## INTRODUCTION

Several species have been shown to distinguish between small sets of objects (Agrillo, Piffer, Bisazza, & Butterworth, 2012; Gross et al., 2009; Xia, Siemann, & Delius, 2000), but only humans develop a symbolic number system (SNS) in which each symbol (e.g., number word or Arabic digit) represents a specific quantity. Acquisition of the SNS is a long process that involves intensive training; it takes children about 2 years (between 2 and 4 years of age) to acquire the symbolic representations for magnitudes 1–4 (also called "the subitizing range"; Wynn, 1990, 1992). The importance of developing accurate symbolic representations in early childhood is underlined by its positive correlation with arithmetic skills later in life (Göbel, Watson, Lervåg, & Hulme, 2014; Holloway & Ansari, 2009; Siegler & Ramani, 2009; Duncan et al., 2007). Yet, the mechanism underlying the mapping process that imbues number symbols with their numerical meaning—known as the "symbol-grounding process"—remains poorly understood and the topic of much debate (Spelke, 2017; Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016; Siegler, 2016; Sullivan & Barner, 2013;

Piazza, 2010; Carey, 2009; Le Corre & Carey, 2007; Dehaene, 2001). Here, we aimed to gain more insight into the symbolic mapping process and its neurocognitive correlates by measuring adults' EEG while they performed a trial-and-error learning task in which they had to map meaningless novel symbols onto a specific magnitude (range 1–9) presented as dot arrays.

Number processing has predominantly been studied using number comparison paradigms. In such paradigms, participants decide which of two stimuli represents the largest quantity. The task can comprise symbolic and/or nonsymbolic stimuli (e.g., arrays of dots) and can entail comparisons between multiple simultaneously presented stimuli or between a target and a fixed internal reference number (e.g., the number 5). A well-documented behavioral phenomenon in such tasks is the numerical distance effect (NDE; Moyer & Landauer, 1967), referring to performance (RT, accuracy) becoming better when the numerical distance between two to-be-compared numerical stimuli gets larger. Dehaene (1992, 2001) and Dehaene, Dupoux, and Mehler (1990) have argued that the NDE arises as a result of numbers being represented spatially along a mental number line, along which representational precision or "tuning" decreases with increasing quantity.

Maastricht University

On such an ordered mental number line, two representations can interfere with each other if their tuning distributions overlap, resulting in the NDE.

Based on this literature, two different systems have been proposed to underlie the processing of nonsymbolic numerosities (numbers of objects). The first being an object tracking system (OTS) for exact representation of smaller numbers of objects (four or smaller; Feigenson, Dehaene, & Spelke, 2004). This OTS is a preattentive limited capacity system that is able to track ≈4 items in parallel. The second is the approximate number system (ANS) for making approximate estimations of the magnitude of larger numbers of objects (larger than four) without relying on language or symbols (Feigenson et al., 2004). A quantity of four or fewer items (OTS) can be estimated very quickly and without error (Piazza, Fumarola, Chinello, & Melcher, 2011; Revkin, Piazza, Izard, Cohen, & Dehaene, 2008; Lipton & Spelke, 2004; Trick & Pylyshyn, 1994). By contrast, estimations of quantities above four are less accurate, although the precision and robustness of representations of larger numbers does increase with development (Halberda & Feigenson, 2008). In addition to these two systems processing the magnitude of object sets, literate adults possess a third system, the SNS, in which each number symbol (e.g., the word "seven" or the digit "7") represents a specific quantity (e.g., a set of seven objects). Although the nonsymbolic OTS and ANS have been observed in various nonhuman species (Agrillo et al., 2012; Xia et al., 2000) and very young infants (Feigenson, Carey, & Hauser, 2002; Xu & Spelke, 2000), only humans develop the SNS (see for reviews on symbolic and nonsymbolic representations: Piazza, 2010; Nieder & Dehaene, 2009). As mentioned above, the process by which one acquires the SNS is still poorly understood, and different proposals on how number symbols (such as Arabic numerals) might gain their numerical meaning have been put forward.

For many years, the dominant view has been that number symbols acquire their numerical meaning by direct mapping onto already existing nonsymbolic representations of numbers in the ANS (Piazza, 2010; Dehaene, 2001). However, based on more recent findings and reviews of the number processing literature, several authors concluded that direct (exact) nonsymbolic (ANS)–symbolic mapping is not likely to apply for the full range of discrete quantities (Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016; Siegler, 2016) and might only occur for small numerosities 1–4 (Reynvoet & Sasanguie, 2016; Carey, 2009). Several recent developmental studies have indeed shown that 3- to 4-year-old children acquired nonsymbolic (dot arrays)–symbolic (number word/digit) mappings more easily for small (1–3) than for larger (4–9) numerosities, suggesting that children can only make use of a nonsymbolic system for symbol mapping in the case of small quantities for which one can easily derive the cardinality (Lira, Carver, Douglas, & LeFevre, 2017; Hurst, Anderson, & Cordes, 2017; Benoit, Lehalle, Molina, Tijus, & Jouen, 2013). Sullivan and Barner (2014) concluded that only the first six number words were directly (associatively) mapped onto their corresponding quantities in 5- to 7-year-old children. It has accordingly been suggested that once numerical representations for the first four symbols have been acquired, the learning of symbolic representations for larger quantities would rely on generalizing successor and cardinality principles without the need for direct nonsymbolic to symbolic mapping (Reynvoet & Sasanguie, 2016; Siegler, 2016; Carey, 2009).

The electrophysiological correlates of the learning process by which nonsymbolic to symbolic mapping occurs for numbers 1–4 has been studied in nonhuman primates (Diester & Nieder, 2007, 2010; Nieder, Diester, & Tudusciuc, 2006), but in humans equivalent, nonsymbolic to symbolic mapping studies are lacking. There are, however, numerous studies that have investigated the electrophysiological correlates of processing of already learned nonsymbolic and symbolic numbers in the adult brain. In such studies, two early ERP components gained most attention and have been linked to different number processing stages. The first is the N1, which peaks at approximately 150 msec poststimulus above occipital-temporal electrodes. The second is the P2p, which is a posterior positivity following the N1 which, in adults, peaks around 200–300 msec after stimulus onset and is measured at parietal-occipital-temporal electrodes (Rubinsten, Dana, Lavro, & Berger, 2013; Hyde & Spelke, 2009; Libertus, Woldorff, & Brannon, 2007; Temple & Posner, 1998; Dehaene, 1996). Most of these number processing ERP studies have made use of number comparison paradigms, taking the latency at which numerical ratio or distance effects first occur in the ERP as the time at which numerical magnitude representations are accessed. This approach, however, has yielded inconsistent conclusions about whether one already has access to numerical representations in the early processing window reflected by the N1 or whether such access is available only later in the P2p processing window. Although distance or ratio effects have been reported on the N1 in nonsymbolic number comparison tasks, these were confounded by changes in visual properties of nonsymbolic number stimuli, such as the size and density of dot arrays that vary with increasing quantity (Soltész & Szűcs, 2014; Hyde & Spelke, 2009; Libertus et al., 2007). This is in line with other studies showing that N1 topography can be modulated by stimulus format (e.g., number words vs. Arabic digits; Park, Chiang, Brannon, & Woldorff, 2014; Turconi, Jemel, Rossion, & Seron, 2004; Pinel, Dehaene, Rivière, & LeBihan, 2001; Dehaene, 1996).

Dehaene (1996) was the first to report a P2p amplitude decrease in adults in association with a decreased RT when increasing the numerical distance between an internal reference number (five) and a symbolic number stimulus (Arabic digit or number word representing numbers 1–9). This P2p NDE was lateralized to the right hemisphere and was not affected by stimulus format

(digits or words). These ERP findings led Dehaene (1996) to identify the processing stage reflected by the P2p as the stage at which one has first access to semantic numerical (magnitude) representations. The decrease in P2p amplitude with increasing numerical distance between two to-be-discriminated numbers is thought to reflect the need for less effortful processing when there is less representational overlap. These number distance effects on the P2p were replicated in several other number comparison studies with children or adults (Libertus et al., 2007; Pinel et al., 2001; Temple & Posner, 1998). Modulation of P2p by the distance or ratio between number stimuli was also reported in a passive viewing paradigm that did not require active number comparisons (Hyde & Spelke, 2009). A study by Rubinsten et al. (2013) showed that even when the primary task of the participants was to decide if three simultaneously presented dot arrays (range 1–20) were in the correct numerical order, the P2p was modulated by the absolute quantity (ratio) differences between the stimuli, whereas earlier components such as the P1 were only sensitive to ordinality. These consistent findings of number distance or ratio effects on the P2p amplitude across modalities and in passive viewing tasks as well as in tasks requiring active quantity comparisons corroborate the findings of Dehaene (1996) and make the P2p the best candidate for the processing stage where first numerical access to magnitude codes occurs.

The above reviewed studies have, however, investigated the effects of number processing on N1 and P2p components in adult participants who already have fully matured symbolic number representations that may no longer rely on nonsymbolic magnitude representations (Leibovich & Ansari, 2016; Lyons, Nuerk, & Ansari, 2015). It is thus not known whether the N1 and P2p components are (differentially) modulated by the process of nonsymbolic–symbolic mapping, through which novel symbols acquire their numerical meaning. To the best of the authors' knowledge, only two prior studies measured the effects of artificial number symbol learning on behavior as well as ERPs (Merkley, Shimi, & Scerif, 2016; Zhao et al., 2012). In both studies, adults acquired the numerical meaning of novel symbols by either mapping them onto dot arrays representing magnitudes (in both studies above 10) or by using ordinal information. To measure the effects of training, participants performed a number comparison task on the new symbols at different phases during learning (session effects) while monitoring the emergence of distance (ratio) effects on accuracy, RT, and ERPs. In both studies, P2p amplitudes to the novel symbols decreased as those symbols gained numerical meaning, regardless of whether symbols were learned on the basis of magnitude or ordinal information. Furthermore, both studies found significant NDEs on accuracy, RT, and P2p amplitude, the latter being larger for numerically close than far symbols. Merkley et al. (2016) also found NDEs on the N1 (close > far), whereas the

N1 was not studied in Zhao et al. (2012). Concluding, both Zhao et al. (2012) and Merkley et al. (2016) showed that newly learned symbols can elicit NDEs on the P2p in a number comparison task, suggesting their association with a spatial (nonsymbolic) representation of quantity. However, these studies did not report on ERPs during the mapping/learning task itself and only investigated symbols mapped onto quantities of 10 and higher that are processed by the ANS and are not likely to allow for exact mapping. These studies do thus not answer the question of whether nonsymbolic–symbolic mapping affects learning and N1 and P2p components differently for small (1–4) versus large (6–9) number ranges (i.e., inside or outside the subitizing range).

## This Study

The current study will investigate how behavioral performance as well as N1 and P2p amplitude change over time as novel symbols become increasingly more strongly associated with already acquired numerical representations in a trial and error learning task. The range of quantities used contained the subitizing range (1–4) and a nonsubitizing range (6–9), allowing us to assess potential differences in the dynamic mapping for quantities inside and outside the subitizing range. It is hypothesized that one-to-one nonsymbolic–symbolic mapping will only occur on trials in which the dot array automatically activates a specific magnitude with which the new symbol can subsequently be successfully associated on same trials and dissociated on different trials. Because the fast presentation rate will prohibit counting, exact symbol mapping is expected to only occur for low numbers (1–4) processed by the OTS. Dot arrays 6–9 are expected to only activate an approximate sense of number (ANS), which is expected to hinder the acquisition of exact numerical meaning of subsequently shown symbols. Based on this, it is hypothesized that symbol learning will be least successful on trials where dot arrays and symbols both represent a number from the 6–9 numerical range for which only approximate learning can take place. Because prior research points to the P2p as the most likely neural correlate of the processing stage at which retrieval of magnitude codes takes place, modulation of the P2p (and not N1) is expected to occur with the emergence of exact symbolic representations reflecting easier access/retrieval of magnitude codes. Thus, symbol learning effects are hypothesized to primarily occur on the P2p and to be strongest for symbols in the small 1–4 number range for which exact mapping is possible. Only a learning paradigm like that used in this study can disentangle effects of visual stimulus complexity and number magnitude processing on the N1 and P2p, because only the latter changes with learning while low-level stimulus differences are kept the same across learning blocks (for a similar approach in the language domain, see McCandliss, Posner, & Givón, 1997).

## METHODS

### Participants

Participants were 22 university students ($M_{age}$ = 22.3 years, $SD$ = 3.7 years, 20 women). All received a monetary compensation. One participant did not complete the training because of difficulties in tolerating the EEG cap and was removed from all analyses. The study was approved by the local ethical committee of the faculty of psychology and neuroscience at Maastricht University, and all participants signed an informed consent form.

### Procedure

All participants visited the lab on two consecutive days during which they performed several blocks of the nonsymbolic–symbolic mapping/learning task (see Figure 1). On the first day, two blocks of the symbol learning task were performed (each consisting of 144 trials), and a third block of 288 trials was performed the second testing day (for an elaborate description of the learning task, see Symbol Learning/Mapping Task section). The participants also performed a number comparison task with the novel symbols after each of the three learning blocks. Because this task, however, only involved deciding whether a presented symbol was smaller or larger than five (fixed internal reference number), the data from this task did not contribute to the main research question focused on the dynamic effects of the nonsymbolic–symbolic mapping/learning process itself on behavior and ERPs and were thus not reported on here. Finally, at the end of the second testing day, the participants performed a number line placement task in which they had to place randomly ordered cards of the novel symbols on a horizontal number line beginning with the smallest number on the left side. All solutions were photographed for later scoring.

### Symbol Learning/Mapping Task

The design of our nonsymbolic–symbolic learning/mapping task was similar to that used in nonhuman primates by Diester and Nieder (2007). Figure 1 shows the nine symbols (A) and an example trial of the learning task (B). In each trial, a dot array representing one of the nine quantities 1–9 was shown after a fixation square, followed by one of the symbols to be associated with these quantities. The quantity of five was excluded in all subsequent analyses, which is in line with previous research (Libertus et al., 2007; Temple & Posner, 1998; Dehaene, 1996). To control for nonnumerical factors associated with increasing nonsymbolic quantity (such as dot size or surface area), dot diameter was fixed in half of the trials whereas total surface area occupied by the dots was fixed in the other half (by adapting dot size) according to the procedure suggested by Dehaene, Izard, and Piazza (2005). Furthermore, dot array stimuli were randomized on surface area and dot size with 32 possible configurations (16 for fixed size and surface). The nine novel symbols used here were selected from a larger set of 70 symbols (that were created in Powerpoint [Microsoft]), on the basis of their discriminability (minimal overlapping features) and similar spatial occupation (so that they could be centered). Potential
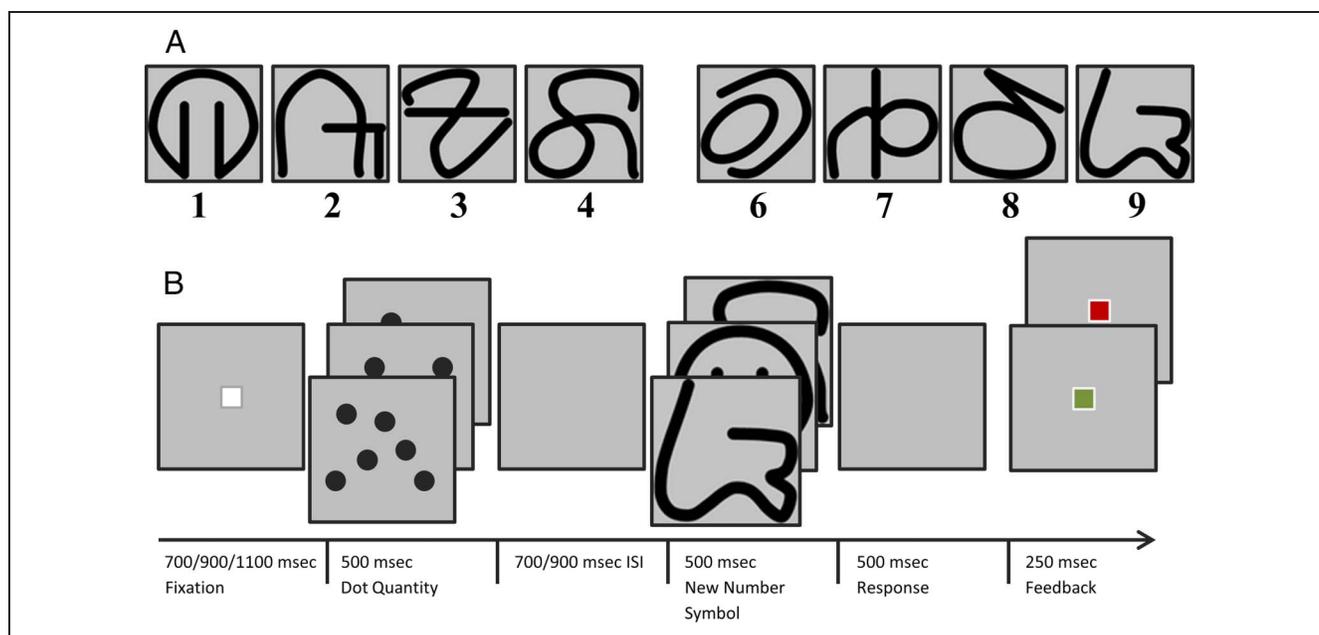


**Figure 1.** (A) The novel symbols representing cardinalities 1–9. (B) Example of a learning task trial: Participants were presented with a dot pattern between 1 and 9 followed by one of the novel number symbols, each symbol having a fixed association with a specific cardinality (see A). Participants had to indicate by button press if the presented symbol represented the same cardinality as the dot pattern (same trials) or a different cardinality (different trials) after which they received feedback (correct/wrong response).

(unwanted) differences in discriminability between the nine chosen symbols were piloted by administering a 1-back task, including the nine symbols to five participants (who did not participate in the later experiment). *n*-Back task results did not show any consistent differences in discrimination speed or accuracy between the nine symbols. In the learning task, the symbols were presented in bold, italic, or a narrow typeface to reduce the possible use of alternative strategies. All stimuli were presented in black on a gray background using Presentation (Version 18.0, Neurobehavioral Systems, Inc.). Stimuli were presented centrally at a viewing distance of 57 cm and subtended about 5° on the retina.

Participants had to indicate whether the magnitudes presented by the dot patterns and novel symbols matched/were the same (left button) or did not match/differed (right button). A block of 144 trials contained all possible pairings of dot arrays and symbols representing numbers in the 1–9 number range, with 72 trials in which paired dots and symbols represented a different quantity (each non-matching dot–symbol combination being presented only once) and 72 trials in which the dot array and the symbol represented the same quantity (same trials were repeated 8 times for each of the nine numbers within a 144-trial block). Trial presentation was fully randomized within 144-trial task blocks. Feedback was given after each correct (green square) and incorrect (red square) response (see Figure 1). Because behavioral learning effects were strongest between the first and the last learning block, only these two learning blocks were included in the analyses. Because the number of trials in the last block of the learning/mapping task was twice as large as that in the first block, we repeated all reported analyses with inclusion of only the first 144 trials of the last block (equalizing trial numbers). Because this did not change our results, we decided to keep the 288 trials in the last learning block to reduce signal/noise ratios and improve model fits. In the ERP analyses, only different trials (irrespective of response accuracy) were included to adequately capture ERP and RT variations due to improvements in dot–symbol mappings across learning sessions, as was the purpose of the study. This also equalized the number of trials and signal/noise ratios between learning task blocks (because this is a learning task with gradually increasing accuracy, including only correctly responded trials would have yielded large differences in trials for ERP averaging in the first versus the last learning block).

### EEG/ERP Acquisition and Analyses

EEG data were recorded using a 64-channel electrode cap (Ag/AgCl) and using the BrainAmp amplifier system (Brain products GmbH) and Brain Vision software. The data were filtered online at 0.01–225 Hz and continuously sampled at a rate of 500 Hz. The left mastoid (A1) served as online reference (A2 measured as active electrode), and POz served as ground. Horizontal and vertical EOG

was measured by electrodes placed on respectively the outer canthus of each eye and above/below the left orbit. All electrode impedances were kept below 5–10 kΩ.

EEG data were processed using EEGLAB/MATLAB 2014a. The data were resampled off-line to 250 Hz and re-referenced to the average signal. A band-pass filter of 0.1–70 Hz was applied to the data before execution of independent component analysis for removal of horizontal eye movements and blinks. The ocular artifact-free data were filtered using a 30-Hz low-pass filter, after which the data were epoched (based on the dot quantity and symbol quantity event codes) into 500 msec prestimulus and 1000 msec poststimulus windows. Baseline correction was performed using the 500-msec prestimulus interval. After this, remaining EEG artifacts were removed by applying an automatic artifact detection procedure rejecting trials with activity exceeding a ±75 μV threshold.

Visual inspection of the topography maps and ERPs identified a bilateral cluster of parietal-occipital electrodes of interest (see Figure 3). These included the left P3, P5, P7, PO3, and PO7 electrodes and the right P4, P6, P8, PO4, and PO8 electrode sites. Based on the number-processing ERP literature (see Introduction) and inspection of the grand-averaged ERP signals of different conditions and sessions, two analysis windows were identified: 130–185 msec (N1) and 210–250 msec (P2p; see Figure 3). Mean amplitude scores within these windows were entered in the statistical analyses.

### Statistical Analysis

To investigate whether performance and ERPs were differently affected by nonsymbolic–symbolic mapping of small numbers (1–4) versus larger numbers 6–9 (see Introduction), all "different" trials (where dot array and symbol represented a different quantity) were split up in 2 × 2 trial categories based on dot quantity (small, 1–4 or large, 6–9) and symbol quantity (small, 1–4 or large, 6–9). Behavioral dependent measures were mean accuracy and mean RT (in the RT analysis trials with RT of <150 or >1000 msec were excluded). Neural dependent variables were N1 and P2p amplitude. For dependent measures RT, N1, and P2p amplitude, we constructed linear mixed models. Accuracy data were analyzed using a generalized estimation equation model due to severe violations of normality. All analyses included within-subject factors hemisphere (two levels: left, right), session (two levels: Session 1 and Session 3), dot quantity (two levels: Small$_{1-4}$ and Large$_{6-9}$), and symbol quantity (two levels: Small$_{1-4}$ and Large$_{6-9}$). To be able to check whether the participants acquired more accurate nonsymbolic–symbolic mappings of small symbols 1–4 (subitizing range) than larger symbols (6–9) an extra 2 × 2 analysis with factors Session × Symbol Quantity was conducted for accuracy and RT on same trials where dot arrays and symbols represented the same magnitude.

The N1 model was constructed using within-subject factors hemisphere, session, dot quantity, and symbol quantity. Various models were tested, including random intercept and random slopes for session, dot, and symbol factors. None of these models had a better fit than the fixed effects model. Nonsignificant higher order interactions were removed in a step-wise manner starting with the four-way interaction Session × Hemisphere × Dot Quantity × Symbol Quantity ($p = .281$), followed by the three-way interaction between Hemisphere × Dot Quantity × Symbol Quantity ($p = .969$) and Session × Hemisphere × Dot Quantity ($p = .359$). The Session × Dot Quantity × Symbol Quantity interaction was kept based on theoretical expectations, and all two-way interactions also remained in the model. The final model used

for the N1 analyses (reported below) was a fixed effects model using an unstructured covariance matrix for repeated measures. Eight observations (single data points within participants) were removed as outliers based on their standardized residuals. The resulting model showed a significantly better fit.

The P2p model was constructed using the same within-subject factors and method as the N1 analysis. Nonsignificant higher order interactions were removed in a stepwise manner starting with the four-way interaction Session × Hemisphere × Dot Quantity × Symbol Quantity ($p = .304$), followed by the three-way interaction between Hemisphere × Dot Quantity × Symbol Quantity ($p = .853$), Session × Hemisphere × Symbol Quantity ($p = .311$), and Session × Hemisphere × Dot

**Table 1.** Means and Standard Deviations (in parentheses) for Behavioral and ERP Data in the Four Different Dot–Symbol Conditions on Different Response Trials (Rows with Dots 1–4 and Dots 6–9) and Behavioral Data on Same Response Trials in Learning Sessions 1 and 3

|  |  |  | Symbols 1–4 | Symbols 6–9 |
|---|---|---|---|---|
| Accuracy, % | Session 1 | Dots 1–4 | 54.96 (17.98) | 59.13 (15.57) |
|  |  | Dots 6–9 | 58.99 (16.45) | 55.06 (14.82) |
|  |  | Same trials | 63.58 (18.37) | 38.43 (14.35) |
|  | Session 3 | Dots 1–4 | 82.97 (11.89) | 93.58 (10.34) |
|  |  | Dots 6–9 | 83.14 (14.25) | 65.99 (14.83) |
|  |  | Same trials | 89.49 (13.45) | 61.77 (20.33) |
| RTs, msec | Session 1 | Dots 1–4 | 683.19 (100.31) | 698.26 (103.27) |
|  |  | Dots 6–9 | 706.51 (108.73) | 729.50 (91.20) |
|  |  | Same trials | 666.82 (115.98) | 706.18 (108.33) |
|  | Session 3 | Dots 1–4 | 618.53 (77.62) | 590.99 (87.47) |
|  |  | Dots 6–9 | 635.13 (92.58) | 627.55 (78.88) |
|  |  | Same trials | 553.31 (73.59) | 630.34 (90.83) |
| N1 (130–185 msec) | Session 1 | Dots 1–4 | −4.33 (3.44) | −2.90 (2.77) |
|  |  | Dots 6–9 | −3.66 (2.03) | −2.52 (2.62) |
|  | Session 3 | Dots 1–4 | −5.42 (2.82) | −4.09 (2.28) |
|  |  | Dots 6–9 | −3.31 (2.82) | −2.54 (2.88) |
| P2p (210–250 msec) | Session 1 | Dots 1–4 | 2.75 (3.16) | 3.37 (2.88) |
|  |  | Dots 6–9 | 3.67 (2.26) | 3.41 (2.25) |
|  | Session 3 | Dots 1–4 | 1.87 (2.52) | 2.73 (2.20) |
|  |  | Dots 6–9 | 2.66 (2.93) | 3.70 (2.67) |
| P2p (210–250 msec)[a] | Session 1 | Dots 1–4 | 3.45 (2.72) | 3.36 (2.68) |
|  |  | Dots 6–9 | 3.93 (2.22) | 3.13 (2.44) |
|  | Session 3 | Dots 1–4 | 3.11 (2.37) | 3.41 (2.55) |
|  |  | Dots 6–9 | 2.72 (2.72) | 3.28 (2.84) |

Values shown are estimated marginal means.

[a]P2p analysis results with N1 amplitude entered as covariate (estimated values for N1 = −3.5066 mV).
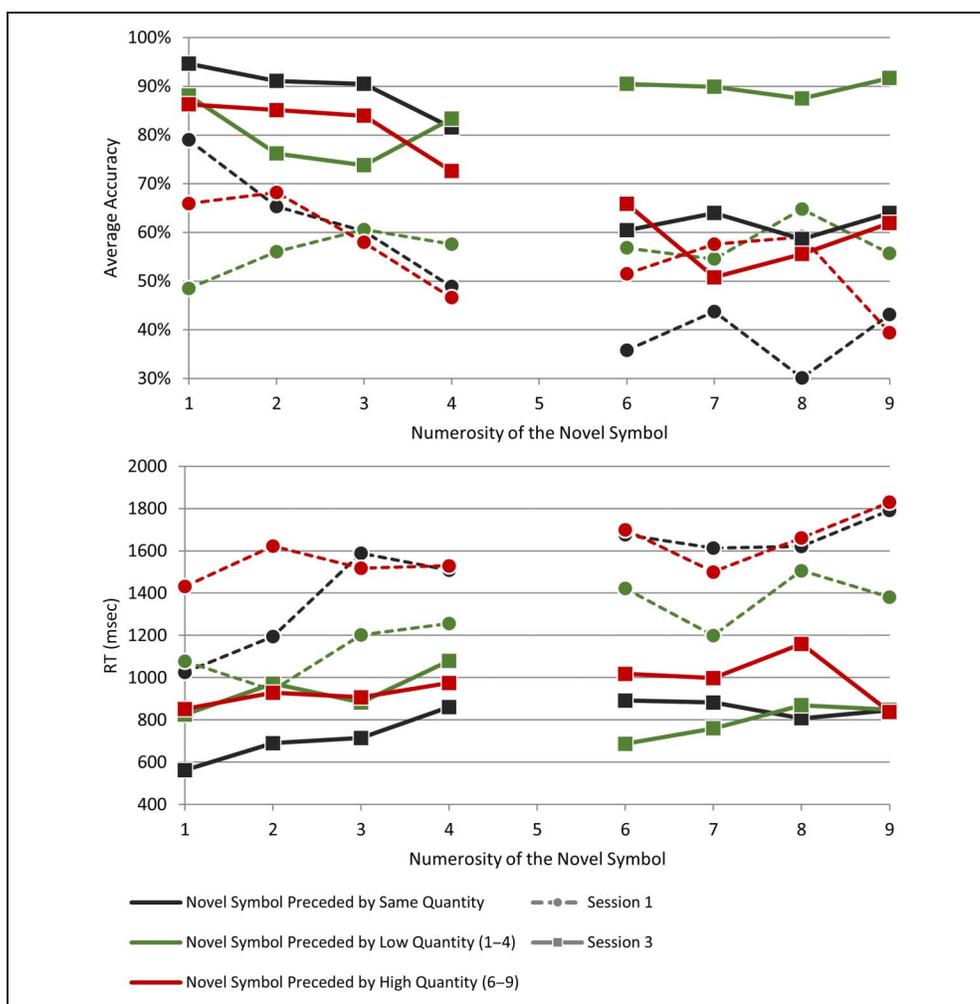
Quantity ($p = .493$). The Session × Dot Quantity × Symbol Quantity interaction was kept based on theoretical expectations, and all two-way interactions also remained in the model. The final model used for the P2p analyses (reported below) was a fixed effects model using an Unstructured Covariance Matrix for repeated measures. A single observation was removed as an outlier without significantly altering the fit of the model.

## RESULTS

### Novel Symbol Learning Task

Table 1 depicts all mean accuracy, RT, and N1, P2p amplitude data for the different dot–symbol combinations in the two learning task blocks (hereafter called sessions) on different response trials in the symbol learning task. The table also shows mean accuracy and RT data on same response trials. Learning curves for accuracy and RT data in the different learning task blocks across low and high number ranges are depicted in Figure 2. Average ERPs and topo-plots are shown in Figure 3.

### Behavioral Data

*Accuracy and RT effects on same response trials. Accuracy.* The accuracy analysis yielded a significant main effect of Session ($X_1^2 = 69.74, p < .001$), showing an overall improvement in accuracy of 24.63% from Sessions 1 to 3, regardless of whether symbols represented a number within the Small$_{1-4}$ or Large$_{6-9}$ number range ($M_{\text{diff-session}}$ Small was 25.92% and $M_{\text{diff-session}}$ Large was 23.34%). A main Symbol Quantity effect ($X_1^2 = 94.34, p < .001$) showed that accuracy was 26.43% higher for Symbols$_{1-4}$ than for Symbols$_{6-9}$ (Session 1: 25.14%; Session 3: 27.72%), indicating that Symbols$_{1-4}$ already acquired numerical meaning during the first 144 trials in Session 1 (see Figure 2). There was no significant Session × Symbol Quantity interaction ($X_1^2 = 0.17, p < .680$).

*RT.* Significant main effects of Session, $F(1, 22.12) = 21.52, p < .001$, and Symbol Quantity, $F(1, 22.22) = 60.32, p = .001$, showed that responses were significantly faster in Session 3, $t(22.12)_{\text{Session1–Session3}} = 4.64, p < .001$, and for Small$_{1-4}$ symbols, $t(22.22)_{\text{Small-Large}} = -7.77, p = .001$. The Session × Symbol Quantity interaction was marginally significant, $F(1, 20.91) = 3.22, p = .087$.
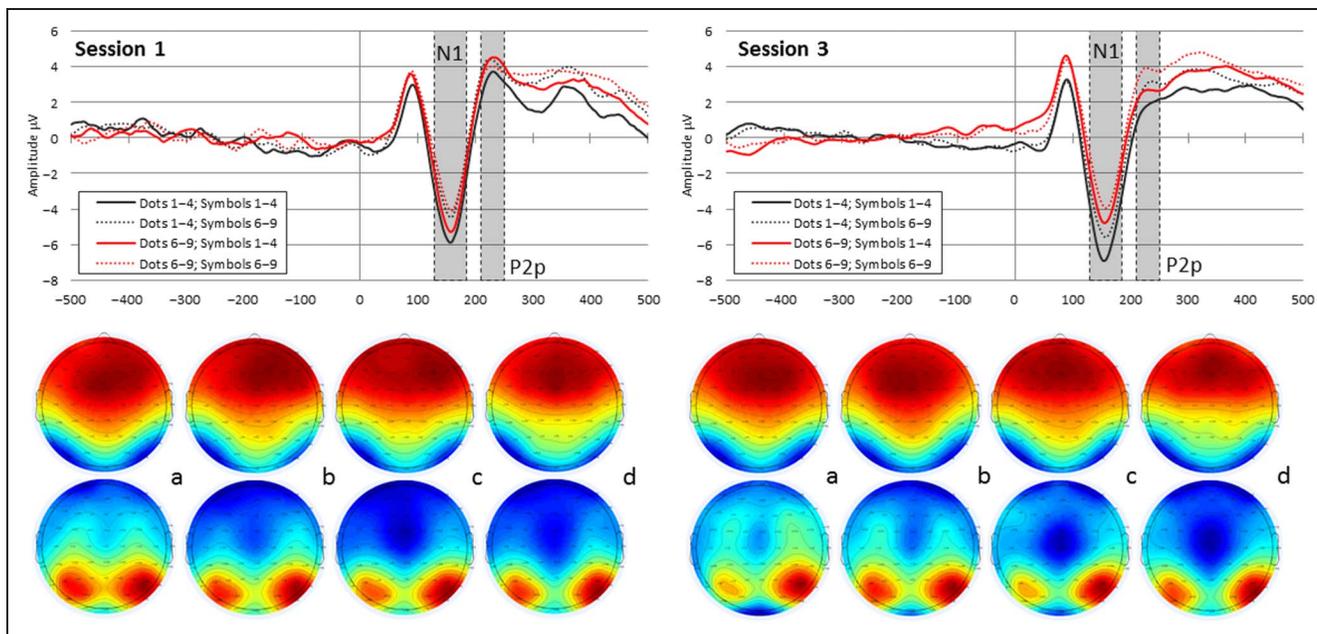
**Figure 2.** Behavioral data from the learning task showing accuracy and RT to the novel symbols in the first (Session 1: dotted lines) and the last (Session 3, solid lines) learning task block in small (1–4) and large (6–9) number ranges. Symbols preceded by a dot quantity equal to the symbol (same trials) are shown in black. Symbol processing on different trials, in which symbols were preceded by dot arrays that did not match in quantity, were split up in (1) novel symbols preceded by a dot array from the small (1–4) number range (green lines) or (2) novel symbols preceded by a dot array from the large (6–9) number range (red lines). For novel symbols 1–4, same trials (shown in black) show stronger training gain in accuracy and speed of responding than different trials (shown in green/red). For symbols 6–9, this effect is reversed with higher training gain in accuracy and RT on different than same trials, but only when symbols 6–9 had to be mapped onto a preceding dot pattern depicting a numerosity in the low 1–4 number range. Both effects match with exact quantity–symbol mapping for only the first four symbols.

**Figure 3.** Grand-averaged ERPs (averaged over left and right parietal-occipital hemisphere clusters) and topography plots (top: N1; bottom: P2p) in the four dot–symbol categories and the first session (left) and the last session (Session 3; right) of the symbol learning task. (a) Small Dots$_{1-4}$/Small Symbol$_{1-4}$, (b) Small Dots$_{1-4}$/Large Symbol$_{6-9}$, (c) Large Dots$_{6-9}$/Small Symbol$_{1-4}$, (d) Large Dots$_{6-9}$/Small Symbol$_{1-4}$.

*Accuracy effects on different response trials.* This analysis yielded significant main effects of Session ($X_1^2 =$ 104.56, $p <$ .001) and Dot Quantity ($X_1^2 =$ 23.63, $p <$ .001), with an accuracy increase from Sessions 1 to 3 and higher accuracy in trials where symbols were preceded by Small$_{1-4}$ compared with Large$_{6-9}$ dot arrays. These main effects were qualified by two-way interactions of Session × Dot Quantity ($X_1^2 =$ 15.16, $p <$ .001) and Dot Quantity × Symbol Quantity ($X_1^2 =$ 23.55, $p <$ .001) and a three-way Session × Dot Quantity × Symbol Quantity interaction ($X_1^2 =$ 15.15, $p <$ .001).

Further testing of the three-way Session × Dot Quantity × Symbol Quantity interaction (Sidak-corrected comparisons) showed significant Session effects (accuracy increases with learning) for all symbol and dot levels (all $p$s ≤ .001). A subsequent analysis was executed on the difference scores (Session 3 − Session 1), showing a significant Dot Quantity × Symbol Quantity interaction ($X_1^2 =$ 8.61, $p =$ .003). Further (Sidak-corrected) comparisons revealed this to be due to a smaller learning effect in trials in which both dot arrays and symbols represented a number in the high 6–9 number range than in any of the other dot–symbol trial combinations: Dots$_{6-9}$–Symbols$_{6-9}$ versus Dots$_{6-9}$–Symbols$_{1-4}$, $t(1) =$ 2.86, $p =$ .025; Dots$_{6-9}$–Symbols$_{6-9}$ versus Dots$_{1-4}$–Symbols$_{6-9}$, $t(1) =$ 5.05, $p <$ .001; and Dots$_{6-9}$–Symbols$_{6-9}$ versus Dots$_{1-4}$–Symbols$_{1-4}$, $t(1) =$ 2.60, $p =$ .054.

*RT effects on different response trials.* Main effects of Session, $F(1, 17.60) =$ 32.46, $p <$ .001, and Dot Quantity, $F(1, 15.64) =$ 15.20, $p =$ .001, were found; RT was significantly faster in Session 3 than in Session 1, $t(17.60)_{Session1-Session3} =$

$-5.70$, $p <$ .001, and in trials in which symbols were preceded by Small$_{1-4}$ compared with Large$_{6-9}$ dot arrays, $t(15.64)_{Small-Large} = -3.90, p =$ .001.

A two-way Session × Symbol Quantity, $F(1, 19.34) =$ 8.80, $p =$ .008, interaction was further explored by performing pairwise comparisons (Sidak-corrected). Testing of the symbol quantity effect per level of session showed no significant differences in RT between Small$_{1-4}$ and Large$_{6-9}$ symbols in Session 1, $t(19.63)_{Small-Large} =$ $-1.64$, $p =$ .118, whereas RT was significantly slower for Small$_{1-4}$ than Large$_{6-9}$ symbols in Session 3, $t(20)_{Small-Large} =$ 2.59, $p =$ .018. Session effects were significant for both Small$_{1-4}$, $t(18.37)_{Session1-Session3} =$ 4.24, $p <$ .001, and Large$_{6-9}$ symbols, $t(16.85)_{Session1-Session3} =$ 6.28, $p <$ .001, with the largest difference for the latter.

## Event-related Potentials

*N1 amplitude (130–185 msec).* The N1 analysis yielded significant main effects of Dot Quantity, $F(1, 19.98) =$ 29.23, $p <$ .001, and Symbol Quantity, $F(1, 15.07) =$ 17.41, $p =$ .001; the N1 amplitude in response to symbols was larger when they were preceded by a small (1–4 item) dot array than when preceded by a large (6–9 item) dot array. Furthermore, small Symbols$_{1-4}$ elicited higher N1 amplitude than large Symbols$_{6-9}$. These main effects were qualified by a two-way interaction of Session × Dot Quantity, $F(1, 18.42) =$ 13.74, $p =$ .002, and a three-way interaction of Session × Hemisphere × Symbol Quantity, $F(1, 20.30) =$ 7.44, $p =$ .013.

The two-way Session × Dot Quantity interaction showed a significant Session effect (larger N1 amplitude

in Session 3 than in Session 1) for trials where symbols were preceded by small dot arrays (1–4 items), $t(19.87)_{Ses1–Ses3} = 2.52$, $p = .020$, but not for trials on which symbols were preceded by large (6–9 item) dot arrays, $t(19.10)_{Ses1–Ses3} = 0.54$, $p = .599$ (see Table 1 and Figure 3). Follow-up tests of the three-way Session × Hemisphere × Symbol Quantity interaction revealed no significant Session × Symbol Quantity interactions in the left nor right hemisphere, left: $F(1, 20) = 0.05$, $p = .942$; right: $F(1, 17.12) = 1.11$, $p = .306$.

There were no other significant main effects (Session, $F(1, 19.81) = 1.32$, $p = .265$, and Hemisphere, $F(1, 17.53) = 0.49$, $p = .493$) nor interaction effects (Session × Hemisphere, $F(1, 17.87) = 0.18$, $p = .681$; Session × Symbol Quantity, $F(1, 19.55) = 0.68$, $p = .419$; Hemisphere × Dot Quantity, $F(1, 17.94) = 0.43$, $p = .518$; Hemisphere × Symbol Quantity, $F(1, 20.33) = 0.19$, $p = .667$; Dot Quantity × Symbol Quantity, $F(1, 20.02) = 0.89$, $p = .356$; and Session × Dot Quantity × Symbol Quantity, $F(1, 20.19) = 0.56$, $p = .463$).

*P2p amplitude (210–250 msec).* The P2p analysis yielded significant main effects for Hemisphere, $F(1, 19.89) = 9.58$, $p = .006$; Dot Quantity, $F(1, 19.95) = 10.06$, $p = .005$; and Symbol Quantity, $F(1, 18.37) = 11.94$, $p = .003$. The P2p amplitude in response to the symbols was larger in the right than the left hemisphere, was highest when symbols were preceded by large (vs. small) dot arrays, and was higher to large (6–9) than small (1–4) symbols. These main effects were qualified by a significant two-way interaction between Session × Symbol Quantity, $F(1, 19.45) = 12.82$, $p = .002$, further testing of which showed a significant session effect (smaller P2p in Session 3 than Session 1) for small Symbols$_{1–4}$, $t(19.90)_{Ses1–Ses3} = 2.23$, $p = .037$, but not for large Symbols$_{6–9}$, $t(20)_{Ses1–Ses3} = -0.47$, $p = .642$ (see Table 1 and Figure 3).

Marginally significant ($p < .1$) effects were found for the Hemisphere × Symbol interaction, $F(1, 18.61) = 3.52$, $p = .076$, and the three-way Session × Dot Quantity × Symbol Quantity interaction, $F(1, 18.64) = 3.08$, $p = .096$. No other effects were significant.

To verify that the above P2p effects were not driven by effects on the preceding N1, a second P2p model was run including the N1 amplitude as a covariate (keeping it constant at the overall N1 mean of $-3.51$ μV) and showed that the N1 amplitude covariate was significant, $F(1, 83.24) = 833.78$, $p < .001$. After inclusion of the N1, all effects involving the dot quantity factor were no longer significant (all $p$s > .11) and were thus driven by earlier N1 effects. All effects involving hemisphere and symbol quantity remained significant and/or only increased in strength after covarying for N1 amplitude: Hemisphere, $F(1, 18.32) = 29.95$, $p < .001$; Hemisphere × Symbol Quantity, $F(1, 20.61) = 6.56$, $p = .018$; and Session × Symbol Quantity, $F(1, 19.30) = 21.202$, $p < .001$. No other effects were significant. Further testing of the Hemisphere ×

Symbol Quantity interaction showed significantly higher P2p amplitude at the right than the left hemisphere in response to both small Symbols$_{1–4}$, $t(18.40)_{Left-Right} = -6.99$, $p < .001$, and large Symbols$_{6–9}$, $t(18.64)_{Left-Right} = -3.93$, $p = .001$, but this effect was slightly more pronounced for small Symbols$_{1–4}$. Further testing of the two-way Session × Symbol Quantity interaction showed a significant session effect (smaller P2p in Session 3 than Session 1) for small Symbols$_{1–4}$, $t(20.16)_{Ses1–Ses3} = 2.70$, $p = .014$, but not for large Symbols$_{6–9}$, $t(19.68)_{Ses1–Ses3} = -0.32$, $p = .749$.

### Number Line Placement (Ordering) Task with the Novel Symbols

An exact position score was computed for all participants by assigning 1 point for each correctly placed symbol on the number line. A Wilcoxon signed-ranks test, comparing low (1–4) and high (6–9) number ranges, showed significant differences in the exact position score, $Z = -2.51$, $p = .012$, $r(21) = -.55$, with mean position scores being 3.57 ($SD = 1.12$, $Mdn = 4$, Min = 0, Max = 4) for the 1–4 number range and 2.76 ($SD = 1.30$, $Mdn = 2$, Min = 0, Max = 4) for the 6–9 number range. These data confirm lower acquisition of exact numerical meaning for higher symbols as was also shown by the behavioral data in the learning task. Participants were also questioned about (a) specific associations with symbols, (b) whether some symbols were easier to learn than others, and (c) whether specific strategies were used. There were no obvious patterns in symbol associations (Question 1), whereas eight participants (34.78%) in response to Questions 2 and 3 indicated in some form that the smaller range (1–4) was easier to learn. This was based on responses such as "trial and error and then I tried to remember the first 4 and the rest by chance" and "I tried to remember the symbols belonging to digits 1–4, since those I could see quickly enough for an exact digit." So, mostly, participants were not aware that, in their learning, low-range numbers had a special status.

### DISCUSSION

The aim of this study was to obtain more insight into how the nonsymbolic–symbolic mapping process and its neurocognitive correlates might differ between low (1–4) and high (6–9) numerical ranges. Previous studies suggest that nonsymbolic–symbolic mapping will be most successful between small (<4) arrays of objects and small digits because one can in this case make use of nonsymbolic systems (OTS) for representing individuals of small arrays, allowing for exact one-to-one mapping. Without counting, such quick one-to-one mapping is suggested to be impossible for arrays of objects above the subitizing range for which enumeration is approximate, based on estimation (Huang, Spelke, & Snedeker, 2010; Le Corre & Carey, 2007; Le Corre, Van de Walle, Brannon, & Carey, 2006; Benoit, Lehalle, & Jouen, 2004). To investigate this,

adults' EEG was measured during performance of a trial-and-error symbol learning task in which they had to acquire the numerical meaning of novel symbols representing a small (1–4) versus a larger (6–9) number by mapping them onto their nonsymbolic counterparts (presented as arrays of dots).

The behavioral (accuracy and RT) data indeed provided evidence for easier mapping/better learning of symbols representing magnitudes 1–4 in the subitizing range. A three-way interaction effect showed that, on trials where dot array and subsequent symbol represented a different magnitude (and required a "different" response), accuracy increased across learning sessions for all four small/large dot–symbol combinations, but this accuracy increase was significantly lowest (only from 55% in Session 1 to 66% in Session 3) when both dots and symbols represented a numerical value in the large (6–9) number range. In the other three categories, accuracy increased much more strongly from 55–59% in Session 1 to 82–93% in Session 3. This can be explained by the fact that numerosities for dot arrays in the low range 1–4 can be readily extracted by subitizing, leading to fast retrieval of an exact numeric representation/code from memory on which novel symbols (1–4) can be easily and exactly mapped in a one-to-one fashion. Support for successful exact mapping of only novel symbols 1–4 was further confirmed by analyses of behavioral data on trials where dots and symbol represented the same magnitude (same trials) and also by data from the number line placement task, showing almost perfect placement accuracy for symbols representing numerosities 1–4, whereas this was much lower for symbols 6–9. An accurate "same" judgment in the learning task requires knowing the exact numerosity belonging to the novel symbol (one-to-one mapping), which was hypothesized to be only possible for dot arrays 1–4 for which a numerical code could be readily extracted. Indeed, symbols 1–4 showed a significantly larger increase in mean accuracy across learning sessions (from 64% in the first session to 89% in the last session) than symbols 6–9, the latter showing a mean accuracy increase from 38 to only 62%. The development of exact numeric representations for symbols 1–4 in turn made it easier to decide that small dot arrays (1–4) and large symbols (6–9) represent a different numerosity, explaining the large accuracy increase from Sessions 1 to 3 in this trial category.

The time to decide whether dots and novel symbols represented a different numerosity decreased across learning sessions for all symbols, but a significant two-way Session × Small/Large Symbol interaction effect showed no significant difference in RTs between small/large symbols in Session 1, whereas in Session 3, RTs were significantly slower (about 40 msec) for symbols in the small (1–4) than large (6–9) number range, independent of whether they were preceded by small or large dot arrays. One likely explanation for this difference is that, with the increase in accurate mapping of the symbols representing numerosities 1–4, it will take more time to retrieve

their unique numerical codes from memory to decide whether the symbol and dot array represent a different (or the same) numerosity as was required in the present learning task. In the case a symbol cannot be mapped onto one specific numerosity, as in the case of a numerosity outside the subitizing range for which no exact numerical code can be extracted without counting (for which the presentation time was too short in the current task), one has to guess/approximate its numerical value, which is a faster process than retrieving an exact numerical code from memory.

The current learning task paradigm also allowed us to investigate whether two early ERP components (N1 and P2p), that in earlier studies have been linked to the processing stage at which one has first numerical access to analog numerical information (Rubinsten et al., 2013; Hyde & Spelke, 2009; Libertus et al., 2007; Temple & Posner, 1998; Dehaene, 1996), were differentially modulated by number symbol learning in our nonsymbolic–symbolic mapping task. The N1 amplitude increased across learning blocks in response to all symbols independent of their numerical range, but only in trials in which the symbols were preceded by dot arrays representing a magnitude in the 1–4 subitizing range. Previous number processing ERP studies often attributed the presence of number distance effects on the N1 in nonsymbolic tasks to perceptual differences in stimulus displays, such as size, contrast, or surface area (Soltész & Szűcs, 2014; Gebuis & Reynvoet, 2013; Libertus et al., 2007). It is, however, unlikely that the current N1 training effects are due to such variations in visual stimulus characteristics because ERP responses were measured in response to the novel symbols that only show small perceptual differences. More importantly, the N1 amplitude change occurred across learning blocks that comprised the same symbolic stimuli and can thus not be due to perceptual stimulus variations between task blocks. If perceptual differences between the novel symbols would have been responsible for the N1 amplitude effect, it should have been present already in the first learning session, which was not the case; the N1 effect only emerged with learning. Instead, we interpret the increased N1 response to the novel symbols across learning blocks in only small dot array trials as a sign of perceptual surprise or expectancy violation, caused by a mismatch between the predicted and actual visual input. More specifically, we propose that only the small (1–4) dot arrays preattentively activated a specific magnitude code that, on different trials, created a perceptual mismatch with the magnitude code evoked by the novel symbol once it was learned in the last learning block. For such perceptual expectancy violations to occur, one should have acquired one-to-one numerosity–symbol mappings, and in the current paradigm with short stimulus viewing times, the development of such mappings was only possible for the dot arrays 1–4 in the subitizing range for which numerosity is preattentively activated. Such an explanation is supported by recent studies by Johnston et al. (2017) and Robinson,

Breakspear, Young, and Johnston (2018) in which an enhanced parieto-occipital N1/N170 response was found when viewing faces or objects that rotated in an unpredicted direction, which was interpreted as a sign of perceptual surprise or expectancy violation.

The second component of interest was the P2p. The literature has established in multiple paradigms that the P2p component is responsive to experimental manipulations requiring access to numerical/magnitude codes. One such task that has been used most frequently is a number comparison task in which one has to decide whether two number stimuli represent the same or a different magnitude, decisions that can be only made when there is magnitude knowledge that has to be accessed to make such numerical decisions (Rubinsten et al., 2013; Hyde & Spelke, 2009; Libertus et al., 2007; Temple & Posner, 1998; Dehaene, 1996). On the basis of the assertion that the P2p reflects the processing stage at which one has first access to numerical representations (Hyde & Spelke, 2009; Libertus et al., 2007; Pinel et al., 2001; Temple & Posner, 1998; Dehaene, 1996), we hypothesized a modulation of the P2p with learning only for novel symbols that can be successfully mapped onto their corresponding numerosity, which we expected to be limited to the subitizing range (Huang et al., 2010; Le Corre & Carey, 2007; Le Corre et al., 2006; Benoit et al., 2004). In the learning stage in which the novel symbols have not yet gained any numerical meaning or in stimulus conditions where exact symbol–quantity mapping is not possible (i.e., in the case of larger arrays of objects that can only be approximated when counting is not allowed), no such number information will/can be retrieved, and there should thus be no P2p modulation. Indeed, a two-way interaction effect between training session and small/large symbol range showed a significant learning effect, with P2p amplitude reducing from Session 1 to Session 3, only for symbols in the 1–4 numerical range and not for symbols in the 6–9 range. The current learning-related reduction of P2p amplitude only for symbols 1–4 is congruent with the by now firmly established finding of smaller P2p amplitude in tasks with easy discriminable (e.g., spatially further apart) compared with harder discriminable number stimuli (e.g., the distance effect). Accordingly, the current P2p amplitude reduction with learning can be interpreted as a sign of lower attentional effort needed to compare/discriminate the numerosities represented by small (<5) dot arrays and symbols once symbol–quantity associations increase in accuracy and strength due to learning. Finally, the finding that this symbol learning effect on the P2p for magnitudes in the subitizing range occurred independent of the numerical size of the preceding dot arrays (there were no dot set size effects on the P2p) and occurred after correction for N1 effects excludes the possibility that the P2p effects were influenced by changes in the visual properties (or priming effects) of preceding nonsymbolic stimuli.

An interesting question is whether the novel symbols representing numbers 1–4 were directly mapped onto their respective nonsymbolic (quantity) representations or indirectly via number words. Because the current study population existed of college students with a fully developed SNS and considerable exposure to complex mathematics, it is very well possible that the P2p effect reflects the mapping of the novel symbols onto number words that were automatically activated when seeing dot arrays in the subitizing range for which quick and exact magnitude estimates could be extracted. Related to this, it is assumable that there is a change in the nature of numerical representational systems across the life span. When first learning number symbols at preschool age, they might be mapped onto early existing, nonsymbolic quantity representations, at least for the first four symbols (Carey, 2009). Later on, with more experience with and exposure to number symbols, this may gradually shift to symbol–symbol (e.g., number word–digit) mapping and result in the development of a discrete semantic representational system as has recently been suggested in adults with an advanced SNS (Marinova, Sasanguie, & Reynvoet, 2018; Krajcsi, Lengyel, & Kojouharova, 2016; Reynvoet & Sasanguie, 2016).

There are as of yet few studies that have investigated symbol learning, especially the mapping of Arabic numerals on magnitude in early childhood. Although there is no consensus yet on the exact order of mapping, all studies agree that magnitude–spoken number word mapping takes place before magnitude–Arabic numeral mapping at respectively 3 and 4–5 years of age (Hurst et al., 2017; Lira et al., 2017; Odic, Le Corre, & Halberda, 2015; Benoit et al., 2013). More importantly for this study, all these developmental studies report behavioral evidence for earlier and more accurate nonsymbolic–symbolic mapping for small (1–3) than medium/large (4–9) numbers for which it is more difficult to determine cardinality, which is in line with the current mapping/learning task results in adults. A study by Sullivan and Barner (2014), for example, concludes that children map words to magnitudes via associative mapping for the lower range (up until approximately six) but rely on an inferential process to acquire quantities above that. This is further expanded upon in a review by Carey and Barner (2019), in which it is proposed that learning of larger number words outside the subitizing range requires the knowledge and usage of counting algorithms and linguistic properties such as ordinality. Future studies might further investigate the electrophysiological correlates underlying the different stages of symbol number learning/mapping in early childhood.

Concluding, the combined behavioral data from both our trial-and-error nonsymbolic–symbolic mapping/learning task and the number line placement task suggest that adults only developed exact numeric representations for novel symbols representing a magnitude in the 1–4 (subitizing) range. This would be in agreement with suggestions that the foundation for symbolic number knowledge lies in

the subitizing range, where exact numerical information can be quickly extracted from nonsymbolic scenes and associated with words and symbols (Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016; Siegler, 2016; Carey, 2009). Furthermore, only the P2p component showed nonsymbolic–symbolic mapping-related changes with learning that are indicative of gradually increased numerical access when the novel symbols acquired numerical meaning. The N1 amplitude did increase with learning for all symbols, but this effect was modulated by dot set size in the sense that it only occurred on trials where symbols were preceded by dots in the 1–4 subitizing range. This symbolic N1 modulation on only small dot trials was explained by expectancy violation that only occurred when one could quickly subitize an exact magnitude from the dot array that led to a prediction about the corresponding learned symbol, which subsequently caused a perceptual mismatch with the observed symbol. In contrast, the reduction of the P2p across learning blocks, interpreted as a decrease in effort to retrieve the numeric symbol code with learning, was not modulated by dot set size and was only present for novel symbols in the low subitizing range, for which exact representations were acquired as shown by the behavioral data. These results support the conclusion that the P2p reflects the processing stage of first access to or retrieval of numeric codes and might in future studies thus be used as a neural correlate of symbol learning in the sense that it is only modulated when symbols acquire numerical meaning.

## Acknowledgments

## REFERENCES

Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One*, 7, e31923.

Benoit, L., Lehalle, H., & Jouen, F. (2004). Do young children acquire number words through subitizing or counting? *Cognitive Development*, 19, 291–307.

Benoit, L., Lehalle, H., Molina, M., Tijus, C., & Jouen, F. (2013). Young children's mapping between arrays, number words, and digits. *Cognition*, 129, 95–101.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Carey, S., & Barner, D. (2019). Ontogenetic origins of human integer representations. *Trends in Cognitive Sciences*, 23, 823–835.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, 44, 1–42.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience*, 8, 47–68.

Dehaene, S. (2001). Précis of the number sense. *Mind & Language*, 16, 16–36.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 626–641.

Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in numerosity experiments (unpublished manuscript). Available at www.unicog.org/pm/pmwiki.php/Main/Arithmetics.

Diester, I., & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLoS Biology*, 5, e50294.

Diester, I., & Nieder, A. (2010). Numerical values leave a semantic imprint on associated signs in monkeys. *Journal of Cognitive Neuroscience*, 22, 174–183.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428–1446.

Feigenson, L., Carey, S., & Hauser, M. (2002). The representations underlying infants' choice of more: Object files versus analog magnitudes. *Psychological Science*, 13, 150–156.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307–314.

Gebuis, T., & Reynvoet, B. (2013). The neural mechanisms underlying passive and active processing of numerosity. *Neuroimage*, 70, 301–307.

Göbel, S. M., Watson, S. E., Lervåg, A., & Hulme, C. (2014). Children's arithmetic development: It is number knowledge, not the approximate number sense, that counts. *Psychological Science*, 25, 789–798.

Gross, H. J., Pahl, M., Si, A., Zhu, H., Tautz, J., & Zhang, S. (2009). Number-based visual generalisation in the honeybee. *PLoS One*, 4, e4263.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457–1465.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, 103, 17–29.

Huang, Y. T., Spelke, E. S., & Snedeker, J. (2010). When is *four* far more than *three*? Children's generalization of newly acquired number words. *Psychological Science*, 21, 600–606.

Hurst, M., Anderson, U., & Cordes, S. (2017). Mapping among number words, numerals, and nonsymbolic quantities in preschoolers. *Journal of Cognition and Development*, 18, 41–62.

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: An electrophysiological investigation of small and large number representations. *Journal of Cognitive Neuroscience*, 21, 1039–1053.

Johnston, P. J., Robinson, J., Kokkinakis, A., Ridgeway, S., Simpson, M., Johnson, S., et al. (2017). Temporal and spatial

localization of prediction-error signals in the visual brain. *Biological Psychology*, *125*, 45–57.

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The source of the symbolic numerical distance and size effects. *Frontiers in Psychology*, 7, 1795.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395–438.

Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, *52*, 130–169.

Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology*, *70*, 12–23.

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions*, *3*, 1.

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, *5*, 271–290.

Lira, C. J., Carver, M., Douglas, H., & LeFevre, J.-A. (2017). The integration of symbolic and nonsymbolic representations of exact quantity in preschool children. *Cognition*, *166*, 382–397.

Lyons, I. M., Nuerk, H.-C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General*, *144*, 1021–1035.

Marinova, M., Sasanguie, D., & Reynvoet, B. (2018). Symbolic estrangement or symbolic integration of numerals with quantities: Methodological pitfalls and a possible solution. *PLoS One*, *13*, e0200808.

McCandliss, B. D., Posner, M. I., & Givón, P. T. (1997). Brain plasticity in learning visual words. *Cognitive Psychology*, *33*, 88–110.

Merkley, R., Shimi, A., & Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations: The role of magnitude and ordinal information. *ZDM Mathematics Education*, *48*, 279–289.

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

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience*, *32*, 185–208.

Nieder, A., Diester, I., & Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science*, *313*, 1431–1435.

Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition*, *138*, 102–121.

Park, J., Chiang, C., Brannon, E. M., & Woldorff, M. G. (2014). Experience-dependent hemispheric specialization of letters and numbers is revealed in early visual processing. *Journal of Cognitive Neuroscience*, *26*, 2239–2249.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*, 542–551.

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*, 147–153.

Pinel, P., Dehaene, S., Rivière, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage*, *14*, 1013–1026.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*, 607–614.

Reynvoet, B., & Sasanguie, D. (2016). The symbol grounding problem revisited: A thorough evaluation of the ANS mapping account and the proposal of an alternative account based on symbol–symbol associations. *Frontiers in Psychology*, 7, 1581.

Robinson, J. E., Breakspear, M., Young, A. W., & Johnston, P. J. (2018). Dose-dependent modulation of the visually evoked N1/N170 by perceptual surprise: A clear demonstration of prediction-error signalling. *European Journal of Neuroscience*. https://doi.org/10.1111/ejn.13920.

Rubinsten, O., Dana, S., Lavro, D., & Berger, A. (2013). Processing ordinality and quantity: ERP evidence of separate mechanisms. *Brain and Cognition*, *82*, 201–212.

Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science*, *19*, 341–361.

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*, 545–560.

Soltész, F., & Szűcs, D. (2014). Neural adaptation to non-symbolic number and visual shape: An electrophysiological study. *Biological Psychology*, *103*, 203–211.

Spelke, E. S. (2017). Core knowledge, language, and number. *Language Learning and Development*, *13*, 147–170.

Sullivan, J., & Barner, D. (2013). How are number words mapped to approximate magnitudes? *Quarterly Journal of Experimental Psychology*, *66*, 389–402.

Sullivan, J., & Barner, D. (2014). Inference and association in children's early numerical estimation. *Child Development*, *85*, 1740–1755.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences, U.S.A.*, *95*, 7836–7841.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.

Turconi, E., Jemel, B., Rossion, B., & Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cognitive Brain Research*, *21*, 22–38.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220–251.

Xia, L., Siemann, M., & Delius, J. D. (2000). Matching of numerical symbols with number of responses by pigeons. *Animal Cognition*, *3*, 35–43.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

Zhao, H., Chen, C., Zhang, H., Zhou, X., Mei, L., Chen, C., et al. (2012). Is order the defining feature of magnitude representation? An ERP study on learning numerical magnitude and spatial order of artificial symbols. *PLoS One*, *7*, e49565.