Bidirectional, Longitudinal Associations Between Math Ability and Approximate Number System Precision in Childhood

Leanne Elliott, Lisa Feigenson, Justin Halberda & Melissa E. Libertus

To cite this article: Leanne Elliott, Lisa Feigenson, Justin Halberda & Melissa E. Libertus (2019) Bidirectional, Longitudinal Associations Between Math Ability and Approximate Number System Precision in Childhood, Journal of Cognition and Development, 20:1, 56-74, DOI: 10.1080/15248372.2018.1551218

To link to this article: https://doi.org/10.1080/15248372.2018.1551218

Published online: 30 Nov 2018.

Article views: 140
Bidirectional, Longitudinal Associations Between Math Ability and Approximate Number System Precision in Childhood

Leanne Elliott<sup>a</sup>, Lisa Feigenson<sup>b</sup>, Justin Halberda<sup>b</sup>, and Melissa E. Libertus<sup>a</sup>

<sup>a</sup>University of Pittsburgh; <sup>b</sup>Johns Hopkins University

ABSTRACT
Research suggests that individual differences in math abilities correlate with approximate representations of quantity that are supported by a primitive Approximate Number System (ANS). However, relatively little research has addressed the direction of this association in early childhood. Here we examined the development of the ANS and math ability longitudinally in 3- to 5-year-old children. Children were observed at three time points roughly six months apart; they completed a nonsymbolic numerical comparison task that measured ANS precision and a standardized math assessment. A series of cross-lagged panel models was then estimated to explore the associations between ANS precision and math ability over time. Bidirectional associations between ANS precision and math ability emerged: Early ANS precision was related to children’s later math skills, and early math ability also significantly predicted children’s later ANS precision. Evidence for mutual enhancement over time between the ANS and symbolic math ability adds to our growing understanding of the ANS and how the ANS and math knowledge interact.

Introduction
Early math achievement is a strong predictor of children’s later academic success (Duncan et al., 2007), prompting research into the factors promoting math skills. In addition to general cognitive abilities, including language, working memory, and executive functioning (Ginsburg, Lee, & Boyd, 2008), research implicates basic quantity representations as a predictor of school math skills. Specifically, the ability to estimate and mentally manipulate numerical quantities without formal symbols (e.g., numerals or number words) explains some of the variability in children’s symbolic math performance (e.g., Feigenson, Libertus, & Halberda, 2013). How might individual differences in the foundational Approximate Number System (ANS) that underlies our basic intuitions about quantity relate to mathematical thinking in the classroom or in standardized assessments? Several studies have attempted to uncover directional and even causal associations between the ANS and math with mixed success. In the present work we asked whether ANS precision and math performance related to one another over time in a longitudinal sample of children, using cross-lagged panel analyses.
**Correlations between the Approximate Number System and Math Abilities**

The ANS supports quantification without verbal counting and allows thinkers to perform number comparisons and approximate addition, subtraction, multiplication, and division using these intuitive quantity representations (Dehaene, 1997; McCrink & Wynn, 2004). In contrast to exact integer representations (e.g., Carey, 2009; Halberda, 2016; Halberda & Odic, 2014), ANS representations provide only approximate values, with representational uncertainty growing with the target quantity (Halberda, 2016). As a result, numerical discrimination (e.g., identifying the more numerous of two arrays) is harder with larger quantities than smaller ones, even when the absolute differences are equal—in other words, approximate number discrimination is ratio dependent (Dehaene, 1992). A second distinction between the ANS and integers concerns their ontogenetic and phylogenetic histories—whereas integer representations take years to construct (Wynn, 1990) and only emerge in human cultures that use number words (Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004), the ANS is present in newborns (Izard, Sann, Spelke, & Streri, 2009), nonhuman animals, and innumerate cultures (Gordon, 2004; Pica et al., 2004).

What is the relation between these two systems for thinking about quantity? Evidence for a link between the ANS and symbolic math abilities comes from correlations between the two systems. For example, some studies report that individuals who can make finer ANS discriminations tend to score higher on math tests (e.g., Fazio, Bailey, Thompson, & Siegler, 2014; Geary & vanMarle, 2016; Inglis, Attridge, Batchelor, & Gilmore, 2011; Lourenco, Bonny, Fernandez, & Rao, 2012; Mussolin, Nys, Leybaert, & Content, 2014). Although other studies fail to find such a link (Göbel, Watson, Lervåg, & Hulme, 2014; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Rodic et al., 2015; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013), meta-analyses reveal moderate correlations between ANS precision and math performance (e.g., r = .20 in cross-sectional analyses), even across studies that statistically control for factors such as IQ, language skills, and executive functions (Chen & Li, 2014; Schneider et al., 2016).

One possible explanation for the discrepancies between previous studies pertains to the debate over the numerical versus nonnumerical nature of performance in tasks designed to measure approximate number representations. Because numerosity is often confounded with visual cues in tasks measuring the ANS, there has been discussion about whether performance truly reflects numerical approximation, or simply domain-general visual processing (Allik, Tüülmets, & Vos, 1991; Dakin, Tibber, Greenwood, & Morgan, 2011; Gebuis, Kadosh, & Gevers, 2016; Gevers, Kadosh, & Gebuis, 2016; Henik, Gliksman, Kallai, & Leibovich, 2017; Leibovich, Katzin, Harel, & Henik, 2016; Tokita & Ishiguchi, 2010). Some amount of interaction between numerical and nonnumerical cues makes sense for efficient number processing, as nonnumerical cues—such as surface area (Tokita & Ishiguchi, 2010) or convex hull (Gebuis & Reynvoet, 2012)—may be helpful in natural environments. Importantly, these visual cues may influence children’s performance to a greater extent than adults’ (Szűcs, Nobes, Devine, Gabriel, & Gebuis, 2013). Although people tend to perform better on comparisons in which nonnumerical cues are congruent with number, the numerical ratio between arrays remains a significant predictor of performance over and above these visual cues (DeWind, Adams, Platt, & Brannon, 2015; DeWind, Park, Woldorff, & Brannon, 2018; Park, DeWind, Woldorff, & Brannon, 2015; Starr, DeWind, & Brannon, 2017). Furthermore, in congenitally blind individuals ANS acuity is unimpaired and correlates with math performance (Kanjlia, Feigenson, & Bedny, 2018).
Relatedly, it has been argued that nonnumerical aspects of discrimination tasks may contribute to the relation between the ANS and symbolic math performance (Clayton, Gilmore, & Inglis, 2015; Fuhs & McNeil, 2013; Gilmore et al., 2013). When nonnumerical and numerical cues are competing in visual arrays (e.g., the array with greater area has fewer items), additional inhibitory control may be required of the observer. Some have suggested that this executive processing may subserve the observed relation between the ANS and symbolic math ability (Bugden & Ansari, 2016; Gilmore et al., 2013; but see Keller & Libertus, 2015; Starr et al., 2017). Thus, studying the nature of these associations remains complex, given that numerical and nonnumerical magnitudes are linked by necessity, and therefore it is impossible to fully control for general cognitive skills such as inhibitory control.

**Directionality in Associations between the Approximate Number System and Math**

Critical to any link between the ANS and math ability is its causal direction. Does having a more precise ANS lead to better symbolic math abilities, does having better math ability sharpen ANS precision, or both? Several longitudinal studies suggest that ANS precision does predict later math performance (Geary & vanMarle, 2016; Libertus, Feigenson, & Halberda, 2013a; Mazzocco, Feigenson, & Halberda, 2011; Mussolin, Nys, Content, & Leybaert, 2014; Starr, Libertus, & Brannon, 2013), and a recent meta-analysis revealed small but significant associations between earlier approximation abilities and later math skills (Chen & Li, 2014). In addition, Chen and Li (2014) found evidence for the reverse: Earlier math skills predicted later numerical approximation across five independent samples (Desoete, Ceulemans, De Weerdt, & Pieters, 2012; Halberda, Ly, Wilmer, Naiman, & Germine, 2012; Libertus et al., 2013a).

To investigate causal links from the ANS to math skills more rigorously, several researchers have attempted to manipulate the ANS and examine the resulting impact on math. Park and Brannon (2013) found that having adults practice adding and subtracting approximate quantities improved symbolic math performance, although they found no evidence that training improved ANS precision per se, and hence training may have improved mental calculation or other processing aspects (see Lindskog & Winman, 2016; Park & Brannon, 2016). Other studies with children found positive effects of ANS training on math performance (Hyde, Khanum, & Spelke, 2014; Khanum, Hanif, Spelke, Berteletti, & Hyde, 2016; Park, Bermudez, Roberts, & Brannon, 2016; Wilson, Dehaene, Dubois, & Fayol, 2009), although again there is mixed evidence regarding whether this training affected ANS precision, leaving the cause of the observed improvement unclear (Hyde, Berteletti, & Mou, 2016; Szücs & Myers, 2016). Finally, Wang, Odić, Halberda, and Feigenson (2016) found that temporary modulation of 5-year-olds’ ANS precision transferred to symbolic math performance but did not affect verbal performance. Despite some methodological concerns (e.g., small sample sizes, Szücs & Myers, 2016; lack of pretest measures, Merkley, Matejko, & Ansari, 2017; also see response by Wang, Odić, Halberda, & Feigenson, 2017), these training studies add to the growing body of work addressing possible causal links between the ANS and math.

Investigations of the reverse pathway, with math skills affecting ANS precision, have been fewer and even more mixed. One study asked whether math training improved adults’ ANS precision and found no such effect (Lindskog, Winman, & Poom, 2016). However, natural experiments suggest that math may affect the ANS. Pica et al. (2004)
showed that adult speakers of Mundurukú, a language lacking number words beyond five, represent approximate quantities less precisely than individuals from numerate Western cultures. Among the subset of Mundurukú who had experienced some formal education, education level correlated with ANS precision (Piazza, Pica, Izard, Spelke, & Dehaene, 2013). A similar pattern emerged among Western adults: Adults lacking formal math education were worse at comparing symbolic and nonsymbolic quantities than adults with some math education (Nys et al., 2013). Lindskog, Winman, and Juslin (2014) found that among undergraduates, year of study significantly related to ANS precision, such that third-year students (who had taken more math-related classes) had more precise approximate number representations than first-year students. Notably, these studies focused on adults and therefore do not speak to these pathways early in development.

A final approach to assessing directionality in the association between the ANS and math is to use advanced statistical methods. Cross-lagged models help approximate causal inference by examining the relative strengths of associations in constructs across time and modeling directional pathways in their relations (Newsom, 2015). Specifically, cross-lagged panel models examine longitudinal associations between constructs by addressing whether prior measures of one construct explain additional variability in a second construct even after controlling for the initial level of the dependent variable. In this way, these models estimate how a predictor relates to change in the outcome variable. This method, based on multiple regression, allows for the estimation of multiple pathways simultaneously and as such is particularly well suited to explore reciprocal or bidirectional correlations with longitudinal data. It is important to note that these models do not allow for full causal inference—although they do provide a window into causation across time (for more discussion see Berry & Willoughby, 2017; Rogosa, 1980).

One recent study with Chinese third to fifth graders using this type of statistical approach found that associations between the ANS and math abilities were primarily unidirectional; only pathways from early ANS precision to later arithmetic abilities were significant (He et al., 2016; see also Libertus et al., 2013a). However, Mussolin et al. (2014) examined cross-lagged correlations between preschoolers’ ANS precision and the ability to count and recognize digits and found the opposite: Symbolic number skills predicted improvements in ANS precision, whereas the reverse was not true. The reasons for these disparate findings across the two studies are unclear, although it is possible that developmental differences or variations in the aspects of math abilities tested may explain these inconsistencies.

The Current Study

In the present work we sought to characterize longitudinal associations between the ANS and math in children using cross-lagged panel models. Specifically, we asked whether ANS precision predicts changes in preschool- to early elementary-school-aged children’s math skills, and whether math ability predicts changes in ANS precision, across three time points in early childhood. Our approach combines the strengths of He et al. (2016) and Mussolin et al. (2014) in examining younger children during a developmental period in which math skills are growing rapidly.
Methods

Participants

We assessed 193 children (94 girls) drawn from a larger longitudinal study addressing the development of math and language skills (Keller & Libertus, 2015; Libertus et. al., 2011, 2013a, 2013b); children were between 3 and 5 years old at the first observation. Data were originally collected from 204 children, but seven children had to be excluded due to refusal to participate (n = 3), external interference (n = 1), or developmental disorders (n = 3). Further, three 2-year-olds and one 6-year-old were excluded because they fell outside of our specified age range (i.e., 3–5 years). Among participants’ parents, 80% of mothers and 87.2% of fathers reported having a college degree or higher. The study was approved by the local Institutional Review Board, and parents provided informed written consent prior to children’s participation. Descriptive information is given in Table 1.

Measures

Math ability

Children completed the Test of Early Mathematics Abilities (TEMA-3, Ginsburg & Baroody, 2003) at each of the three time points (see Procedure). This standardized mathematics assessment measures numbering skills (e.g., verbally counting objects), number-comparison (e.g., determining which of two numbers is larger), numeral literacy (e.g., reading numerals), and calculation skills (e.g., solving simple addition and subtraction problems). Children’s performance was measured using age-normed standard scores. Form A was used at the first and second time points (T1 and T2 respectively), and Form B was used at the third (T3).

ANS precision

At each time point, children completed a nonsymbolic number-comparison task similar to that of Halberda, Mazzocco, and Feigelson (2008). Children saw arrays of blue and yellow dots, spatially separated on a 13-inch laptop screen. They were told that these were blue balls belonging to the character Grover and yellow balls belonging to Big Bird and were asked to verbally indicate who had more. The experimenter pressed a corresponding key on a keyboard as soon as children responded. Children first completed six practice trials with verbal feedback from the experimenter. Then, at T1 and T2, children completed 60 test trials. In these test trials children saw blue and yellow dot arrays of numerosities

Table 1. Descriptive Statistics on All Study Variables at Each Time Point.

<table>
<thead>
<tr>
<th></th>
<th>Time Point 1</th>
<th>Time Point 2</th>
<th>Time Point 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANS Precision (percent correct)</td>
<td>65.09 (15.08)</td>
<td>76.42 (13.57)</td>
<td>81.78 (9.30)</td>
</tr>
<tr>
<td>Size-Controlled Trials</td>
<td>62.66 (16.18)</td>
<td>73.12 (15.68)</td>
<td>82.64 (10.37)</td>
</tr>
<tr>
<td>Area-Controlled Trials</td>
<td>66.83 (16.94)</td>
<td>79.50 (14.15)</td>
<td>79.30 (10.15)</td>
</tr>
<tr>
<td>Math Ability (standard = 100; SD = 15)</td>
<td>107.58 (14.94)</td>
<td>109.76 (14.89)</td>
<td>110.65 (15.79)</td>
</tr>
<tr>
<td>Age (in years [y] and months [m])</td>
<td>4y3m (8m)</td>
<td>4y9m (8m)</td>
<td>5y3m (9m)</td>
</tr>
<tr>
<td>Inhibition (percent of commission errors)</td>
<td>57.01 (24.95)</td>
<td>57.01 (24.95)</td>
<td>57.01 (24.95)</td>
</tr>
<tr>
<td>Nonverbal IQ (standard = 100; SD = 15)</td>
<td>121.55 (18.69)</td>
<td>121.55 (18.69)</td>
<td>121.55 (18.69)</td>
</tr>
<tr>
<td>N</td>
<td>172</td>
<td>165</td>
<td>134</td>
</tr>
<tr>
<td>Gender (percent female)</td>
<td>49%</td>
<td>48%</td>
<td>49%</td>
</tr>
</tbody>
</table>

Note. ANS = approximate number system.
between 4 and 15 each, drawn randomly from one of four numerical ratio bins (larger numerosity divided by smaller numerosity): 1.17, 1.33, 1.5, and 2.0 (e.g., 10 yellow vs. 5 blue dots). At T3, children completed 64 test trials, with numerosities ranging from 5 to 22 and test numerosities drawn from one of four ratio bins: 1.14, 1.17, 1.5, and 2.5. These included ratios that were both harder (1.14) and easier (2.5) than those presented at T1 and T2. Because ANS precision increases over development (Halberda & Feigenson, 2008), we wanted to include a more challenging ratio to preclude ceiling effects. Yet there were also some children who, at T1 and T2, struggled with the numerical discriminations; we included the easier ratio to keep them motivated.

Children heard computerized feedback throughout: high-pitched tones for correct responses and low-pitched tones for incorrect ones, as previous work suggests that feedback in these tasks can increase participant motivation (Lindskog, Winman, & Juslin, 2013), and feedback has been used in previous studies (e.g., Halberda & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011; Libertus et al., 2013a; Libertus, Feigenson, & Halberda, 2013b; Mazzocco et al., 2011; Odic, Libertus, Feigenson, & Halberda, 2013; Park & Brannon, 2013; Wang et al., 2016). Which array was more numerous was counterbalanced, and dot size varied within each array.

To ask whether the directional associations between the ANS and math differ when ANS precision is measured with arrays in which numerosity is congruent with a salient visual cue (i.e., cumulative surface area) versus arrays in which numerosity and surface area are incongruent, we systematically varied the surface area of the stimulus arrays. On a random half of the trials the cumulative surface area of the arrays was equivalent (i.e., area controlled), and on the other half average dot size within the arrays was equivalent (i.e., size controlled). Default dot size was 60 pixels (approximately 2° of visual angle from a viewing distance of 18 inches), and dots could vary up to ± 35% of default dot size. Convex hull was estimated post hoc by running the testing program 20 times and saving the generated images. We then calculated the convex hulls of all arrays and found that the convex hull of the arrays correlated negatively with number on 15% of trials overall (i.e., the collection with fewer items had a larger convex hull). As numerical discrimination difficulty increased (i.e., as numerical ratios approached 1), the proportion of trials in which convex hull negatively correlated with number also increased. For example, at the hardest numerical ratio at T3, convex hull was negatively correlated with number on 42% of the trials.

Spearman-Brown adjusted split-half reliabilities for numerical discrimination performance at T1, T2, and T3 were .84, .84, and .76 respectively, similar to previous reports in adults (Libertus, Odic, & Halberda, 2012; but see Clayton et al., 2015; DeWind & Brannon, 2016 for concerns regarding test-retest reliability).

**Control variables**

To assess inhibitory control, children completed the Conners Kiddie Continuous Performance Test (K-CPT; Conners, 2006), which is similar to conventional go/no-go tasks and other measures of behavioral inhibition (Ballard, 2001; Nigg, 2000). Continuous performance tasks are widely used to measure attention and inhibition in clinical settings (e.g., Edwards et al., 2007; Hall et al., 2016; Vogt & Williams, 2011). Here, children saw a stream of images and pushed a button after every image except for images of a ball. Children completed five blocks in subblocks of 20 trials. Images were presented for 500 ms with an interstimulus interval of either 1,500 or 3,000 ms; total testing time was
7.5 minutes. Commission errors are considered to measure lack of inhibition. Thus, performance was quantified as the percentage of trials on which a child failed to inhibit a response. Nonverbal IQ was assessed with the Primary Test of Nonverbal Intelligence (PTONI; Ehrler & McGhee, 2008), which asks children to look at pictures and point to the picture that does not belong with the others. Items increase in difficulty, with early items measuring lower-order reasoning (e.g., visual and spatial perception) and later items measuring higher-order reasoning (e.g., analogical thinking; sequential reasoning). Performance was measured using age-normed standard scores.

Procedure

Children were tested individually in a quiet room in a preschool or laboratory. Testing sessions occurred roughly six to seven months apart ($M = 205.65$ days, $SD = 45.23$ days between T1 and T2, and $M = 192.24$, $SD = 43.34$ days between T2 and T3). At T1, children completed the TEMA-3 and then the ANS precision task. T2 was split into two sessions approximately two weeks apart ($M = 12.84$ days, $SD = 12.71$ days), with the TEMA-3 and K-CPT administered in the first session and the ANS precision task in the second. At T3, children completed the ANS precision task, followed by the TEMA-3 and PTONI. Additional measures collected in this study but not analyzed here include parental report of children’s expressive vocabulary, and working memory, which were relevant for empirical questions not tested here.

Analysis plan

Missing data

Some children were unavailable for testing at certain time points, resulting in attrition across waves of data collection. Of the 193 children included here, 21 did not contribute data at T1, 28 at T2, and 59 at T3. Full information maximum likelihood (FIML) was used to handle missing data given evidence that this method performs well even with relatively high levels of missing data (Enders & Bandalos, 2001); however, the alternative approach of deleting cases with any missing data resulted in no differences in the observed pattern of results. Similarly, removing children who performed below chance on the ANS task (less than 50% accuracy; $n = 45$) yielded similar results. Given minimal differences across specifications, the full sample of 193 children was used in all further analyses to maximize power and increase the precision of our estimates.

Cross-lagged panel analysis

To assess longitudinal associations between math skills and ANS precision, cross-lagged panel models were estimated using data from the three time points, with the goal of determining the most statistically supported and appropriate model for our data set (see Newsom, 2015). These models include autoregressive pathways, or estimates of how strongly prior measures of a variable predict later measures of that same variable; these reflect the stability of constructs over time. Models also include cross-lagged paths, where prior levels of one variable are used to predict change in another. In the present study, we also included residual correlations between constructs at T1 and T3. These residual correlations capture carry-over effects—or potential shared development in both math
and ANS precision, such that a positive residual correlation would indicate that children with higher residuals on one construct would also have higher residuals on the other. Importantly, the addition of these correlations does not influence the estimates of either the autoregressive or cross-lagged pathways but instead only affects overall model fit (i.e., extent to which the parameter estimates produced accurately represent the observed data) and thus increases the precision of all parameter estimates. We evaluated model fit using conventional fit indices (i.e., nonsignificant model chi-square, RMSEA < .06, CFI > .95, and SRMR < .08; Hu & Bentler, 1999) and compared nested models with likelihood ratio tests, where significant improvements in fit indicated that additional parameters in the model should be included (whereas nonsignificant improvements in fit indicated that model parsimony should be prioritized over improving model fit).

**Cross-lagged panel model specifications**

Models were estimated in MPlus 7 (Muthén & Muthén, 2012). We first ran a series of models to detect associations between math performance and ANS precision. Math and ANS precision at T1 were included as predictors of both math and ANS precision at T2, and math and ANS precision at T2 were used to predict both variables at T3. This model estimates the autoregressive associations in math abilities and in ANS precision over time (e.g., whether individuals with high scores at T1 also had high scores at T2) as well as the cross-lagged paths, which reflect how one construct relates to changes in the other. Additionally, the models included estimates of the residual correlations between math ability and ANS precision at T1 and between the residuals of math ability and ANS precision, after controlling for prior math and ANS precision, at T2 and T3. We were primarily interested in how earlier ANS precision related to later math abilities and how earlier math abilities related to later ANS precision; we therefore focused on cross-lagged paths. Children’s current age, IQ, and inhibitory control were included as predictors of math and ANS precision at T2 and T3. Specifically, math and ANS precision were also regressed on children’s age, IQ, and inhibitory control. The cross-lagged paths of interest represent how one construct related to changes in the other, controlling for each of these factors.

**Results**

**Computing Individual Differences in Numerical Approximation**

First, we computed estimates of children’s ANS precision by examining their accuracy on the discrimination task as a function of the ratio between the two numerosities on each trial and as a function of trial type (dot size controlled, cumulative area controlled). Consistent with previous results (e.g., Halberda & Feigenson, 2008; Halberda et al., 2008), children were more accurate on trials in which the relative difference between numerosity was larger, as shown by linear contrast effects, T1: $F(1, 171) = 40.18, p < .001$; T2: $F(1, 155) = 249.69, p < .001$, and T3: $F(1, 133) = 587.22, p < .001$. One-sample $t$-tests revealed that at all time points, children performed above chance on both trial types (all $p$s < .001). It is also noteworthy that children performed better on area-controlled than size-controlled trials at T1 and T2, $t(171) = 3.78, p < .001$, and $t(154) = 6.672, p < .001$, whereas at T3, the opposite pattern was seen, $t(132) = 4.71, p < .001$. Children’s accuracy at each time point was $z$-scored so that performance relative to the group across time could be examined.
Model Comparison and Selection

Correlations between all study variables are shown in Table 2. A cross-lagged panel model was estimated with ANS precision and math ability at each time point, with autoregressive paths between time points within each construct and cross-lagged paths from ANS precision to math ability and math ability to ANS. Covariates were included in this model, as well as residual correlations between ANS precision at T1 and T3 and between math ability at T1 and T3. Because unexplained variance in math and ANS precision at T3 (after accounting for math and ANS precision at T2) correlated with individual differences in these measures at T1 in preliminary models (data not shown but available upon request), we added these residuals to the model. This model had good fit to the data based on the fit criteria described in our analytic plan, \( \chi^2(6) = 9.18, p = .16; \) RMSEA = .05 [95% CI = .00, .12]; CFI = .997; SRMR = .02. Additionally, we constrained pathways to be equivalent over time (i.e., stationarity constraints). Given the relatively wide age range of children in our sample compared to the spacing of time points, similar coefficients were expected between T1 and T2 as between T2 and T3. As such, autoregressive and cross-lagged pathways were set to equivalence with the exception of the autoregressive pathways for ANS precision, which did not result in a significant change in model fit from the model with no such constraints, \( \Delta \chi^2(3) = 6.35, p = .10. \) Preliminary analyses suggested that the ANS autoregressive pathways differed significantly, and so these were not constrained to equivalence.

In the final model, auto-regressive pathways in math ability and ANS precision were positive and significant, indicating stability in both ANS precision and math ability across time, as shown by the horizontal lines in Figure 1. ANS precision became increasingly stable over time, as indicated by the nonstationary autoregressive pathways (i.e., the horizontal lines at the bottom of Figure 1), but autoregressive pathways in math ability did not change significantly from T1 to T2 compared to T2 to T3 (i.e., the horizontal lines at the top of Figure 1). All cross-lagged pathways were also positive and significant, suggesting bidirectional relations in ANS precision and math ability over time (as shown in the diagonal arrows between math ability and ANS precision in Figure 1). Specifically, arrows from ANS precision to later math ability suggest that prior ANS precision relates to changes in math over time. Similarly, arrows from math ability to later ANS precision reflect how math ability relates to changes in ANS precision.

Table 2. Correlations for All Study Variables (\( N = 193 \)).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ANS Precision</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. ANS Precision</td>
<td>.54***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. ANS Precision</td>
<td>.53***</td>
<td>.66***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Math Ability</td>
<td>.48***</td>
<td>.51***</td>
<td>.43***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Math Ability</td>
<td>.51***</td>
<td>.51***</td>
<td>.44***</td>
<td>.76***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Math Ability</td>
<td>.34***</td>
<td>.33***</td>
<td>.29***</td>
<td>.64***</td>
<td>.66***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Age</td>
<td>.50***</td>
<td>.54***</td>
<td>.45***</td>
<td>.24***</td>
<td>.19**</td>
<td>−.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Age</td>
<td>.50***</td>
<td>.51***</td>
<td>.44***</td>
<td>.24***</td>
<td>.20**</td>
<td>−.07</td>
<td>.99***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. IQ</td>
<td>−.16*</td>
<td>−.13†</td>
<td>−.17*</td>
<td>.30***</td>
<td>.26***</td>
<td>.28***</td>
<td>−.14†</td>
<td>−.12†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Inhibition</td>
<td>−.22**</td>
<td>−.27***</td>
<td>−.17***</td>
<td>−.31***</td>
<td>−.27***</td>
<td>−.31***</td>
<td>−.10</td>
<td>−.13†</td>
<td>−.21**</td>
<td></td>
</tr>
</tbody>
</table>

Note. ANS = approximate number system. Subscripts reflect the time of measurement. Correlations significant at \( p < .001 \) would remain statistically significant with Bonferroni corrections.

\* \( p < .10; \) † \( p < .05; \) ** \( p < .01; \) *** \( p < .001. \)
Cross-Lagged Panel Models for Area-Controlled and Size-Controlled ANS Trials

Finally, we repeated our analyses while separately indexing children’s performance on the two trial types in the ANS precision task (i.e., area-controlled and size-controlled trials). Models were estimated with residual covariances that included children’s age, IQ, and inhibitory control as covariates. The model without stationarity constraints for area-controlled trials had excellent model fit, $\chi^2(6) = 4.74, p = .58; \text{RMSEA} = .00 \ [95\%\ CI = .00, .08]$; $\text{CFI} = 1.000; \text{SRMR} = .01$. Stationarity constraints significantly decreased model fit, $\Delta \chi^2(4) = 14.48, p = .01$. However, when the autoregressive pathway for ANS precision was allowed to vary over time (identical to the specification described for the model shown in Figure 1), no significant decrease in model fit was observed, $\Delta \chi^2(3) = 6.94, p = .07$, and overall model fit was good, $\chi^2(9) = 11.68, p = .23; \text{RMSEA} = .04 \ [95\%\ CI = .00, .10]; \text{CFI} = .997; \text{SRMR} = .02$. Thus, we modeled accuracy on area-controlled trials as stationary with one modification.

The model for size-controlled trials had adequate fit to the data, $\chi^2(6) = 13.39, p = .04; \text{RMSEA} = .08 \ [95\%\ CI = .02, .14]; \text{CFI} = .992; \text{SRMR} = .02$. Adding stationarity constraints did not significantly decrease model fit, $\Delta \chi^2(4) = 4.29, p = .37$, and resulted in an overall well-fitting model, $\chi^2(10) = 17.64, p = .06; \text{RMSEA} = .06 \ [95\%\ CI = .00, .11]; \text{CFI} = .992$.

![Figure 1](image-url) Cross-lagged panel model of associations between math ability and approximate number system (ANS) precision over time, including covariates. Although not shown, ANS precision and math ability at T2 and T3 were regressed on child age, IQ, and inhibitory control. Math ability and ANS precision scores were converted to z-scores prior to analyses to ease interpretation of these estimates. Single-headed arrows represent regression pathways, whereas double-headed arrows indicate correlations. Unconnected arrows represent residuals rather than raw values. Path estimates (regression coefficients or correlations) are shown on each arrow, with standard errors in parentheses. Time points are indicated by the variable subscript.

† $p < .10; *p < .05; **p < .01; ***p < .001.$
SRMR = .02. Thus, we modeled accuracy on size-controlled trials as fully stationary (unlike the model shown in Figure 1, where ANS precision autocorrelations were independently estimated).

Estimates from the best fitting models for each trial type are shown in Figure 2. Importantly, models for both trial types again suggested significant bidirectional relations between math ability and ANS precision. Interestingly, autoregressive pathways were only marginally significant on size-controlled trials, suggesting that performance on these trials was less stable over time compared to area-controlled trials. As such, our data support the claim that the ANS and math are significantly related, irrespective of trial type (i.e., regardless of whether cumulative surface area is correlated with number or controlled).

**Discussion**

In this study, we tested for predictive relations both from ANS ability to math ability and from math ability to ANS ability in 3- to 5-year-old children. To accomplish this, we relied on a cross-lagged panel analysis to examine bidirectional associations between the ANS and math abilities. Our results revealed stability in both ANS precision and math ability...
during early childhood, suggesting that children maintain their relative ranking compared to the sample over early childhood—children with math or ANS scores above their peers at one time point tended to have relatively higher scores on the same measure later on. These results are consistent with ANS precision and math ability having stable developmental trajectories over time. Furthermore, earlier ANS precision was related to later math ability after controlling for early math ability; and earlier math ability related to later ANS precision after controlling for early ANS precision suggesting possible causal links from enhanced ANS to improving math and from enhanced math to improving ANS.

Our results differ from those of He et al. (2016), in which the observed associations were unidirectional from the ANS to math skills. However, our sample was considerably younger than theirs, raising the possibility that pathways from math to the ANS may be more pronounced in early childhood. Our data informally support this hypothesis, as models without controls showed that pathways from math to ANS precision were significant from T1 to T2 but not T2 to T3, tenuously suggesting that this directional pathway might weaken with development. Additionally, previous work suggests that the ANS is more strongly related to children’s informal math skills (e.g., comparing quantities) than formal math skills (e.g., recalling math facts; Libertus et al., 2013b). Our math measure included a broader range of math abilities, including both formal and informal math abilities, compared to that of He et al. —a difference that again may have contributed to the divergent findings. Finally, He et al. did not include control variables in their models. Including such covariates may be crucial for detecting the potentially weaker link in older children.

As described in the introduction, a major concern in the extant literature is the degree to which nonsymbolic number comparison tasks such as the one used here reflect the ability to discriminate continuous magnitudes instead of number (Gebuis et al., 2016; Gevers et al., 2016; Leibovich et al., 2016; Szűcs et al., 2013; Tokita & Ishiguchi, 2010). Here, we found that bidirectional associations between numerical comparisons and math remained when examining trials in which cumulative surface area was either congruent with numerosity (size-controlled trials) or equated (area-controlled trials). If associations between nonsymbolic numerical comparisons and math abilities were attributable solely to children’s abilities to make discriminations about cumulative surface area, we would expect that only performance on trials in which numerosity and surface area are congruent would relate to children’s math abilities. Admittedly, other continuous magnitudes such as convex hull were not accounted for in our study, leaving open the possibility that judgments of these continuous dimensions may have influenced the observed results. However, since it is impossible to rule out the influence of all continuous magnitudes on task performance simultaneously, we argue that the observed these bidirectional associations remain theoretically interesting regardless of whether they reflect an association between purely numerical processing and math or a combination of numerical and nonnumerical processing and math. In other words, our findings suggest that children’s math skills and their ability to make numerical judgments are bidirectionally linked, even though it remains to be seen precisely how numerical judgments involve the use of nonnumerical cues.

If bidirectional associations do in fact describe causal pathways, what mechanisms could explain these effects? This key question remains an active topic of debate with several viable possibilities (Feigenson et al., 2013; Mussolin, Nys, Leybaert, & Content, 2016). One is that more precise approximate number representations may assist when children initially learn number words and symbols, such that children with more precise
ANS representations might obtain a deeper or earlier understanding of number words and digits. Although this might account for some of the observed links between the ANS and math in the preschool and elementary school years, this explanation has a harder time accounting for observed effects among older participants (Halberda et al., 2012; Libertus et al., 2012). Another possibility is that more precise ANS representations allow participants to “ballpark” the answers to symbolic math problems, potentially helping them avoid gross errors in calculation (Feigenson et al., 2013; Lourenco et al., 2012). This process would require not just representing and comparing abstract quantities (including numerical representations as well as continuous magnitudes) but also operating on them (e.g., performing nonsymbolic arithmetic), which was not addressed in this study but has been investigated elsewhere (e.g., Barth et al., 2006). A third possibility is that the ANS provides a sense of the meaning of arithmetic operations such as addition and subtraction, and this stronger intuitive sense of approximate arithmetic operations (rather than simply a more precise sense of approximate cardinalities) could assist in symbolic computation (Park & Brannon, 2014; Hyde et al., 2016). This explanation could also apply to relations between continuous magnitude representations and math. Finally, the ANS may provide a sense of confidence in number estimation, comparison, and arithmetic manipulation, with some individuals being both more accurate and more confident in their numerical judgments across development, leading to the observed effects on math ability (Odic, Hock, & Halberda, 2014; Wang et al., 2016).

Fewer mechanisms have been proposed regarding the link from math to the ANS. One possibility is that exposure to the precise system of symbolic numbers helps individuals attain increased precision in their nonsymbolic ANS representations. An alternative but not mutually exclusive possibility is that associations between the two systems stem from shared underlying neural resources. Dehaene’s neuronal recycling hypothesis suggests that culturally specific skills are built from existing brain circuitry that serves a similar but simpler purpose (Dehaene, 2005; Dehaene & Cohen, 2007). Under this hypothesis, the neural systems that support the ANS, which is evolutionarily ancient and present cross-culturally, may also support the formal math skills that are slowly acquired through instruction. That is, complex cultural constructs such as integer representations may “recycle” existing cortical structures that support their primitive underpinnings—in this case, the ANS. Dehaene (2005) further hypothesized that as the newer, learned function is acquired, the original function can be affected; in fact, if the functions are fairly similar, as in the case of math abilities based on symbolic representations and the ANS, refinement of the networks that occurs during math learning may also benefit other systems that utilize them. A recent longitudinal study revealed some support for these hypotheses, as children’s ANS precision showed a jump in precision that coincided with children’s learning the meaning of the number words (Shusterman, Slusser, Halberda, & Odic, 2016).

Our findings highlight several avenues for continued investigation. First, although cross-lagged panel models allow for the testing of directional hypotheses, these analyses do not address whether associations are actually causal in nature (see Berry & Willoughby, 2017; Rogosa, 1980). Even though experimental approaches offer stronger causal evidence, past studies examining the benefits of ANS training on math performance have some potential limitations that call for caution in the interpretation of the observed effects (Szücs & Myers, 2016); hence future studies that examine the impact of specific experience on both ANS function and math will be important. Second, it remains unknown how our results would generalize to a more diverse sample, as children in the present study had parents with fairly
high levels of education. Although little work has directly addressed how the ANS might operate across contexts, some evidence suggests that associations between math and ANS precision differ by socioeconomic status (Fuhs & McNeil, 2013; Valle Lisboa et al., 2017). In terms of practical importance, some work suggests that symbolic representations and comparisons are more consistently and strongly related to math abilities than the nonsymbolic representations focused on here (De Smedt, Noël, Gilmore, & Ansari, 2013; Fazio et al., 2014; Lyons et al., 2014; Sasanguie et al., 2013; Schneider et al., 2016). This is perhaps to be expected, as acquired facility with number symbols is likely to be more important to acquired symbolic mathematics than to other aspects of cognition—especially evolutionarily ancient ones.

Finally, in the present study we relied on standardized scores of children’s math ability and ANS precision. As such, our analyses describe changes in children’s performance relative to the other children assessed at that same time point. Standardizing ANS accuracy scores was necessary in our study, given differences in our approximate number task across testing sessions. Future studies could administer repeated measures of ANS precision using a single stimulus set to model growth in ANS performance directly.

In summary, we demonstrate that in early childhood, the ANS and symbolic math skills are bidirectionally linked. These results strengthen our understanding of the relation between an evolutionarily and developmentally foundational system for approximating quantities and the uniquely human ability to engage in symbolic math.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Institute of Child Health and Human Development [R01 HD057258].

References

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. Cognition, 98(3), 199–222. doi:10.1016/j.cognition.2004.09.011


Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science.* doi:10.1111/desc.12372


