Developmental Change in the Acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-Year-Olds and Adults

Justin Halberda and Lisa Feigenson
Johns Hopkins University

Behavioral, neuropsychological, and brain imaging research points to a dedicated system for processing number that is shared across development and across species. This foundational Approximate Number System (ANS) operates over multiple modalities, forming representations of the number of objects, sounds, or events in a scene. This system is imprecise and hence differs from exact counting. Evidence suggests that the resolution of the ANS, as specified by a Weber fraction, increases with age such that adults can discriminate numerosities that infants cannot. However, the Weber fraction has yet to be determined for participants of any age between 9 months and adulthood, leaving its developmental trajectory unclear. Here we identify the Weber fraction of the ANS in 3-, 4-, 5-, and 6-year-old children and in adults. We show that the resolution of this system continues to increase throughout childhood, with adultlike levels of acuity attained surprisingly late in development.

Keywords: analog magnitudes, mathematics, estimation, numerical, intraparietal sulcus, approximation

The ability to nonverbally represent number is shared across species and across development (for reviews see Dehaene, 1997; Feigenson, Dehaene, & Spelke, 2004). The foundational Approximate Number System (ANS) that underlies this ability produces abstract number representations (Barth, Kanwisher, & Spelke, 2003) that support arithmetic computation across the life span (Barth et al., 2003; Barth et al., 2006; McCrink & Wynn, 2004). The ANS is activated when adults perform symbolic number tasks (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003; Piazza, Pinel, LeBihan, & Dehaene, 2007) and may even provide a foundation for more sophisticated mathematics (Gilmore, McCarthy, & Spelke, 2007). However, the ANS represents number only approximately (Dehaene, 1997; Gallistel & Gelman, 2000), and the imprecision of its numerical representations is radically greater in infants than in adults (e.g., Pica, Lemer, Izard, & Dehaene, 2004; Xu & Spelke, 2000).

Given that the ANS is thought to play an important role in math learning (Booth & Siegler, 2006; Jordan, Kaplan, Locuniak, & Ramineni, 2007), it is surprising that no research to date has explored the full developmental trajectory of its representational acuity. Here we tested ANS acuity in 3-, 4-, 5-, and 6-year-old children and in adults, using psychophysical modeling to determine the finest numerical discriminations possible at each age. Our findings reveal that the ANS does not attain full acuity until quite late in development, long after children have begun formal instruction in mathematics.

The ANS differs from counting in that it produces inexact number representations (Gallistel & Gelman, 1992; but see Zorzi & Butterworth, 1999). A hallmark of the ANS is that the imprecision of its representations grows with the target numerosity, such that the ability to nonverbally discriminate two quantities depends on their ratio (Moyer & Landauer, 1967). This ratio dependence is observed when adults estimate numbers of items (Halberda, Sires, & Feigenson, 2006; Whalen, Gallistel, & Gelman, 1999), produce target numbers of actions (Cordes, Gelman, & Gallistel, 2001; Whalen et al., 1999), judge the more numerous of two arrays (Barth et al., 2003), and estimate the results of arithmetic events (Pica et al., 2004). Because of the inexactness of ANS representations, two quantities cannot be distinguished when the distance between them is too small. The finest numerical ratio that adults can consistently discriminate has been identified as 7:8. This limit can also be described as a Weber fraction that measures the smallest numerical change to a stimulus that can be reliably detected. The Weber fraction is equal to the difference between the two numbers divided by the smaller number; for example, 7:8 → (8 − 7)/7 = .14. When asked to indicate the more numerous of two simultaneously presented arrays containing 20–80 dots, French adults’ Weber fraction is .12 and Amazonian adults’ Weber fraction is .17; thus on average these adults could discriminate ratios differing by about 7:8 (Pica et al., 2004).

Ratio-dependent numerical performance also reveals that preverbal infants use the ANS, albeit with drastically less acuity than adults. Six-month old infants discriminate arrays of 4 versus 8, 8 versus 16, and 16 versus 32 dots, all of which instantiate a 1:2 ratio, but fail to discriminate 8 versus 12 and 16 versus 24, which instantiate a 2:3 ratio (Xu, 2003; Xu & Spelke, 2000; Xu, Spelke, & Goddard, 2005). The same pattern obtains in audition: 6-month-olds discriminate...
is important to note that this type of cross-sectional design might obscure possible distinctions between a group trend and an individual child’s developmental trajectory (e.g., individual children might show discontinuous changes in numerical acuity, which would be revealed only in a longitudinal analysis of changes in individual performance over time). Our present goal was to determine whether ANS acuity continues to develop during the early school years and to estimate the group trend in this developmental trajectory.

The aforementioned data show that ANS acuity increases over the life span, with participants exhibiting a Weber fraction of 1.0 at 6 months, 0.5 at 9 months, and 0.14 in adulthood. However, these studies leave a large gap in our understanding of the development of nonverbal enumeration. It is known that preschool- and early school-age children (the ages during which formal instruction in mathematics typically begins) show performance controlled by numerical ratio, making more errors with close than with distant numerical comparisons (Barth, LaMont, Lipton, & Spekkel, 2005; Huntley-Fenner, 2001; Huntley-Fenner & Cannon, 2000; Starkey & Cooper, 1995; Temple & Posner, 1998). Estimates of numerical acuity in 5-, 6-, and 8-year-olds using a numerical bisection task suggest Weber fraction values in the vicinity of .26, intermediate between the acuity of adults and that of infants (Droit-Volet, Clément, & Fayol, 2003; Jordan & Brannon, 2006). An analysis of children’s reaction times to determine the larger of two Arabic numbers is also consistent with the proposal that the ANS affects performance such that numerically near digits take longer to discriminate (Sekuler & Mierkiewicz, 1977). This effect changes with age throughout the school years, suggesting the possibility of developmental changes in ANS acuity (Sekuler & Mierkiewicz, 1977).

Yet, no study to date has identified the finest numerical discriminations children can make, as none has used psychophysical methods similar to those used with adults to identify children’s Weber fraction. For this reason, the developmental trajectory of changes in ANS acuity between 9 months and adulthood remains undescribed. For example, it is not known whether numerical acuity rapidly asymptotes during the first year of life (similar to stereoaucuity; Held, Birch, & Gwiazda, 1980), gradually increases throughout early childhood (similar to executive function; Diamond, 2002; Zelazo, Craik, & Booth, 2004), or shows a discontinuous change when children master verbal counting at around age 4 (similar to a vocabulary spurt; Carey, 1978; Goldfield & Reznick, 1990). Furthermore, changes in ANS acuity have implications for formal instruction in mathematics. Many math curricula aim to tap children’s intuitions regarding “possible” or “impossible” solutions to quantitative problems, encouraging children to estimate numerical magnitudes before arriving at an exact answer (Johnson, 1979; Levin, 1981). Given the widespread nature of such teaching tools, it is surprising that children’s ANS acuity during the early years of mathematics instruction has not been determined.

We presented 3-, 4-, 5-, and 6-year-old children and adults with a nonverbal number discrimination task that did not permit counting and that included controls for continuous variables that often correlate with number. Varying the numerical ratio between stimulus arrays allowed us to determine the Weber fraction of the ANS for each age group and thereby to describe the function by which the ANS reaches the adult state of representational precision. This cross-sectional approach allowed us to determine whether numerical acuity continues to develop over the ages tested. However, it is important to note that this type of cross-sectional design might obscure possible distinctions between a group trend and an individual child’s developmental trajectory (e.g., individual children might show discontinuous changes in numerical acuity, which would be revealed only in a longitudinal analysis of changes in individual performance over time).
Grover would be the numerically correct choice by a ratio of 2:1. For example, on an area anticorrelated surface area such that the array with the smaller number of items had more total surface area. For example, on an area anticorrelated surface area such that the array with the smaller number of items had more total surface area.

These trials equated the total summed horizontal and total summed vertical extent of the items in Big Bird's and Grover's arrays. This procedure equated perimeter and anticorrelated area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials).

Displays were controlled either for average item size (area correlated trials) or summed continuous extent (area anticorrelated trials). For each ratio presented, on half of the trials the larger numerosity had more total surface area (area correlated trials), and on the other half the smaller numerosity had more total surface area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials). Area anticorrelated trials controlled for the total summed perimeter of the items and their total area (area anticorrelated trials).

Table 2

<table>
<thead>
<tr>
<th>Objects Used in Experiment 1, Reported According to the Label Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagels</td>
</tr>
<tr>
<td>Balls</td>
</tr>
<tr>
<td>Bananas</td>
</tr>
<tr>
<td>Bicycles</td>
</tr>
<tr>
<td>Blocks</td>
</tr>
<tr>
<td>Boots</td>
</tr>
<tr>
<td>Bottles</td>
</tr>
<tr>
<td>Bows</td>
</tr>
<tr>
<td>Buckets</td>
</tr>
<tr>
<td>Bunnies</td>
</tr>
<tr>
<td>Butterflies</td>
</tr>
<tr>
<td>Cars</td>
</tr>
</tbody>
</table>

Possible array were not all larger than those in the more numerous array (see Figure 1).

Unequal numbers of trials were presented from each ratio bin in order to focus on more difficult ratios (ratio bin 1:2 = 2 trials, 2:3 = 2, 3:4 = 2, 4:5 = 2, 5:6 = 10, 6:7 = 10, 7:8 = 10, 8:9 = 14, 9:10 = 1). These numbers were initially chosen for adult participants and so focus more heavily on difficult ratios (e.g., 5:6–9:10). To ensure consistency, we used the same ratio distribution for all ages.

Display time was adjusted for each age group through pilot testing and was chosen to be long enough to allow participants to view both arrays but short enough to prevent serial counting. Display time was 2,500 ms for 3-year-olds, 1,200 ms for 4-, 5-, and 6-year-olds, and 750 ms for adults. The winning side (Big Bird or Grover), ratio presented, trial type (area correlated, area anticorrelated), item type, and absolute number of items presented varied randomly across trials.

Results

To examine whether performance differed across age groups or depended on numerical ratio, we entered each participant’s percentage correct for each ratio bin (1:2, 2:3, etc.) into a 5 (age group) × 2 (sex) × 2 (trial type: area correlated, area anticorrelated) × 9 (ratio)

---

1. Our method for anticorrelating area and equating the total summed perimeter of the items in Big Bird’s and Grover’s sets relies on the geometric fact that the areas of similar polygons are to each other as the squares of any two corresponding segments. This fact allows us to rely on the horizontal and vertical extent of the image files rather than measuring each minute curve in the perimeter of each picture. Because any curved area can be estimated by a series of connected polygons and because the images used within each trial were scaled variants of a single image, the total summed perimeter was precisely equated on each trial and the ratio of areas was exactly opposite the ratio of the number of items.

2. Children’s verbal counting ability was assessed for the purposes of a separate study using Wynn’s Give A Number task (Wynn, 1992) and revealed that none of the 3-year-old children understood the Cardinal Word Principle (Gelman & Gallistel, 1978) and so could not have used counting to determine the numerosities of the sets irrespective of display time.
Table 3
Controls for Continuous Extent on Area Anticorrelated Trials

<table>
<thead>
<tr>
<th>Whole-number ratio</th>
<th>Numerical ratio (N_2/N_1)</th>
<th>Total perimeter</th>
<th>Ratio of total area (N_1/N_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:2</td>
<td>2</td>
<td>Equated</td>
<td>2</td>
</tr>
<tr>
<td>2:3</td>
<td>1.5</td>
<td>Equated</td>
<td>1.5</td>
</tr>
<tr>
<td>3:4</td>
<td>1.33</td>
<td>Equated</td>
<td>1.33</td>
</tr>
<tr>
<td>4:5</td>
<td>1.25</td>
<td>Equated</td>
<td>1.25</td>
</tr>
<tr>
<td>5:6</td>
<td>1.2</td>
<td>Equated</td>
<td>1.2</td>
</tr>
<tr>
<td>6:7</td>
<td>1.16</td>
<td>Equated</td>
<td>1.16</td>
</tr>
<tr>
<td>7:8</td>
<td>1.14</td>
<td>Equated</td>
<td>1.14</td>
</tr>
<tr>
<td>8:9</td>
<td>1.13</td>
<td>Equated</td>
<td>1.13</td>
</tr>
<tr>
<td>9:10</td>
<td>1.11</td>
<td>Equated</td>
<td>1.11</td>
</tr>
</tbody>
</table>

repeated measures analysis of variance. These data are presented in Figure 2, which plots percentage correct for each age group as a function of ratio ( numerosity of larger set/numerosity of smaller set ). There was a significant age group effect, with participants performing better with increasing ratio, \(F(4, 70) = 28.891, p < .001\); a significant ratio effect, with participants performing better with increasing ratio, \(F(8, 560) = 20.199, p < .001\); and a marginally significant effect of sex, with girls and women performing slightly better than boys and men overall, \(F(1, 70) = 3.122, p = .082\).

Regarding the effect of total area on participants’ numerical discriminations, we observed a significant trial type effect, \(F(1, 70) = 6.138, p < .05\). We investigated this effect via planned \(t\) tests comparing each age group’s percentage correct on area correlated and area anticorrelated trials with chance (50%). Collapsing across all ratios revealed that all age groups performed significantly above chance on both trial types, as summarized in Table 2, but performance was slightly better on area correlated trials. All age groups based their responses on number, not area, though area had some effect on judgments.

We also observed a significant Trial Type \(\times\) Ratio interaction, as the difference in performance on area correlated and area anticorrelated trials was larger for easier ratios than for harder ratios, \(F(8, 560) = 2.62, p < .01\). A significant Age Group \(\times\) Ratio \(\times\) Trial Type interaction revealed that older children and adults showed less differentiation in performance between area correlated and area anticorrelated trials as a function of ratio than did younger children. In particular 3-, 4-, and 5-year-olds performed better on area correlated trials than area anticorrelated trials with 1:2 ratio comparisons, whereas 6-year-olds and adults showed little or no difference in performance with this comparison ratio, \(F(32, 560) = 1.575, p < .05\). Any effect of area was minor in the experiment, however, as basing answers on area would have resulted in performance that was significantly below chance, and this was never observed.

Participants’ performance varied as a function of ratio (see Figure 2). If participants were using the ANS to determine the more numerous array, then percentage correct as a function of ratio (collapsed across trial type) should be well fit by a computational model of the ANS. Pica et al., (2004) examined performance on a task similar to ours in adults and children from both a developed country (France) and an indigenous culture (the Munduruku of Amazonia) whose language lacks exact large-number words such as “seven.” Pica and colleagues found that both groups’ performance was well fit by a psychophysics model (Green & Swets, 1966; Moyer & Bayer, 1976) that models ANS representations as “noisy” Gaussian random variables and numerical discrimination as the subtraction of the two Gaussian random variables that represent the numerosities of the two arrays. This model has a single free parameter, the Weber fraction \(w\), which determines the increase in percentage correct with increasing ratio. We rely on this same psychophysics model, which has been argued to be a parsimonious, psychologically plausible model of numerical performance (Pica et al., 2004). This model has received further support from neuronal data of monkeys performing a numerical discrimination task (Nieder & Miller, 2004).

As has been observed in previous numerical discrimination tasks (Pica et al., 2004), participants’ performance can fail to reach 100% correct because of a tendency to guess randomly on some trials. As seen in Figure 2, this was the case for our 3-, 4-, and 5-year-old children. The simplest way to account for this tendency computationally is to include a parameter that is a constant probability of guessing randomly on any particular trial (for additional discussion of this approach see the supporting online materials of Pica et al., 2004). This parameter lowers the model’s asymptotic performance while retaining an accurate estimate of the Weber fraction. We included this parameter in our model as follows: where \(p_{\text{guess}}\) is the probability of guessing randomly, \(P_{\text{correct}}\) is the probability of being incorrect given the model, and chance is .5 multiplied by 100 to return a percentage.

\[
\text{percent correct} = \left[1 - p_{\text{guess}}\right] \left[1 - P_{\text{correct}}\right] + p_{\text{guess}} \times .5 \times 100
\]

As seen in Figures 3a and 3b and Table 5, the psychophysics model provides an accurate fit to our data from 6-year-olds and adults (adult \(R^2\) value = .93). Performance for 3-, 4-, and 5-year-olds deviates from this model, however, because younger children were at chance for more difficult ratios. Figures 3d and 3e show that data from 3- and 4-year-olds exhibit a sigmoidal shape. The likely cause is that the children ceased to rely on ANS representations for the hardest ratios and instead guessed randomly (e.g., 3-year-olds asymptote at 53% for comparisons of ratio 1.25 and lower). This pattern might also simply be a failure to discriminate the numerosities involved on these more difficult ratio comparisons, but performance well above the least squares fit psychophysics model for easier ratios near the elbow of the function (e.g., ratios 1.25, 1.33, and 1.5 for 4-year-olds) suggests that the failures at more difficult ratios are aberrant and a result of “giving up” on

\[
\frac{1}{2} - \text{erfc} \left( \frac{\left| n_1 - n_2 \right|}{\sqrt{2w^2 n_1^2 + n_2^2}} \right)
\]

3 Results similar to those reported here were obtained from analyses of reaction time, which was also recorded, and these data show no evidence of speed–accuracy trade-off differences across ages (e.g., 3-year-olds were both slower and less accurate than 4-year-olds). For purposes of psychophysical modeling, we focus our analyses on percentage correct.

4 In the psychophysics model, each numerosity is represented as a Gaussian random variable (i.e., \(\times 2\) and \(\times 1\)) with means \(n_1\) and \(n_2\) and standard deviations equal to the Weber fraction \(w\) \(n\). Subtracting the Gaussian for the smaller set from that for the larger set returns a new Gaussian that has a mean of \(n_2 - n_1\) and a standard deviation of \(w \sqrt{n_1^2 + n_2^2}\) (simply the difference of two Gaussian random variables). Percentage correct is then equal to 1 – error rate, where error rate is defined as the area under the tail of the resulting Gaussian curve computed as follows:
the task. These data points artificially pull the performance of the psychophysics model down as it attempts to reduce the error between the fit and these data points, resulting in a Weber fraction that likely underestimates the actual acuity of these participants (witness the low $R^2$ values for these fits in Table 5).

To gain further accuracy in estimating the Weber fraction at each age (especially for younger children), we also modeled performance for each age group using a sigmoidal function fit by the Levenberg-Marquardt algorithm. We relied on the following sigmoidal equation:

$$y = \frac{\text{lower} + \frac{\text{upper} - \text{lower}}{1 + e^{-\text{infection} - \text{rate}}}}{}$$

This model should not be viewed as a psychologically plausible model of the representations involved, unlike the psychophysics model we use (Moyer & Bayer, 1976; Pica et al., 2004), but rather as a method for obtaining the best possible fit to our data in order to estimate the Weber fraction while controlling for participants’ tendencies to guess randomly. The sigmoid model has a greater

Table 4
Percentage Correct Compared With Chance (50%) Across All Trials for Each Age Group

<table>
<thead>
<tr>
<th>Age group</th>
<th>Area correlated</th>
<th></th>
<th>Area anticorrelated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentage correct</td>
<td></td>
<td>Percentage correct</td>
</tr>
<tr>
<td></td>
<td>$M$</td>
<td>SE</td>
<td>$t$</td>
</tr>
<tr>
<td>3 years</td>
<td>61.7</td>
<td>3.1</td>
<td>3.741</td>
</tr>
<tr>
<td>4 years</td>
<td>64.4</td>
<td>3.0</td>
<td>4.838</td>
</tr>
<tr>
<td>5 years</td>
<td>70.3</td>
<td>2.7</td>
<td>7.525</td>
</tr>
<tr>
<td>6 years</td>
<td>79.8</td>
<td>3.5</td>
<td>8.537</td>
</tr>
<tr>
<td>Adults</td>
<td>87.7</td>
<td>1.2</td>
<td>32.039</td>
</tr>
</tbody>
</table>
number of free parameters than the psychophysics model and thus will nearly always provide a more accurate fit than the psychophysics model, but with reduced parsimony. From the sigmoidal equation for each age group, the Weber fraction can be estimated as the inflection point of the sigmoid (psychologically equivalent to the midpoint between successful discrimination [upper asymptote] and subjective-equality [lower asymptote]). As seen in Figures 3d and 3e and Table 5, the sigmoidal function provides an accurate fit to our data from younger children (e.g., 4-year-olds $R^2$ value = .98). Table 5 lists the estimated Weber fraction for each age group from both the psychophysics model and the sigmoid model along with the $R^2$ values for the fit. These numbers are also translated into a nearest whole number ratio (e.g., Weber fraction of .25 = 4:5 ratio). As Table 5 shows, estimated Weber fraction decreased with age, confirming that ANS acuity increased across the age groups we tested. Whereas 3-year-old children can accurately discriminate numerosities differing by a 3:4 ratio, 6-year-olds have sufficient acuity to discriminate numer-

Table 5

<table>
<thead>
<tr>
<th>Age group</th>
<th>w</th>
<th>$R^2$</th>
<th>Nearest whole number fraction</th>
<th>w</th>
<th>$R^2$</th>
<th>Nearest whole number fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 years</td>
<td>.525</td>
<td>.502</td>
<td>2:3</td>
<td>.333</td>
<td>.829</td>
<td>3:4</td>
</tr>
<tr>
<td>4 years</td>
<td>.383</td>
<td>.632</td>
<td>3:4</td>
<td>.240</td>
<td>.976</td>
<td>4:5</td>
</tr>
<tr>
<td>5 years</td>
<td>.229</td>
<td>.785</td>
<td>4:5</td>
<td>.225</td>
<td>.938</td>
<td>5:6</td>
</tr>
<tr>
<td>6 years</td>
<td>.179</td>
<td>.846</td>
<td>6:7</td>
<td>.199</td>
<td>.940</td>
<td>5:6</td>
</tr>
<tr>
<td>Adults</td>
<td>.108</td>
<td>.926</td>
<td>9:10</td>
<td>.097</td>
<td>.988</td>
<td>10:11</td>
</tr>
</tbody>
</table>

Note. $R^2$ values represent the agreement between the modeled fit and the data for the entire function (see Figure 3).
osities differing by a 5:6 ratio, and adults’ acuity in our sample was as high as 10:11. Using a numerical discrimination task similar to ours, Pica et al., (2004) estimated untrained French adults’ Weber fraction to be .12 or a 9:10 ratio. The extent of individual differences in Weber fraction in adults and children remains to be determined, but our estimate and that of Pica et al., suggest that the average acuity in adults from educated numerate cultures is in the range of 9:10 or 10:11.

In Figure 4 we have plotted estimates from both models alongside estimates from the developmental literature on infants’ numerical acuity (Lipton & Spelke, 2003; Xu & Spelke, 2000). These estimates have been modeled by least squares fit to determine the developmental trajectory of the increasing acuity of the ANS (i.e., decreasing Weber fraction). These suggest a logarithmic decrease in Weber fraction throughout childhood, with adultlike levels of numerical acuity being attained sometime during the preteen years.

Discussion

Although children between 3 and 6 years old have already begun formal instruction in mathematics, the present results show that the acuity of the ANS is still developing during this time. Indeed, the sharpening of the ANS does not appear to be complete until early adolescence. Given the central role this system plays in supporting mathematical intuitions, this protracted period of development highlights the importance of coming to understand the effects of changes in ANS acuity on math learning and achievement (Booth & Siegler, 2006; Jordan et al., 2007).

What causes the ANS to increase in acuity? Although some of the sharpening of this system may be due to simple maturation of the neural circuitry subserving the ANS, recent evidence suggests that experience can also affect its development. Practice at numerical discrimination appears to increase acuity in children with math learning disabilities (Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). After approximately 8 hr of practice spanning 5 weeks on a computer game designed to engage the ANS, 7- to 9-year-old children showed improvement in both symbolic and nonsymbolic numerical tasks. Continued engagement in numerical discrimination throughout childhood may therefore contribute to the increase in acuity we observed in normally developing children. This suggests the intriguing possibility that subpopulations that do not frequently engage in numerical discrimination might show increasingly reduced acuity with age relative to individuals undergoing formal education. Existing data are consistent with this hypothesis. Although Pica et al., (2004) stressed the similarity in the numerical discrimination abilities of their French and Munduruku participants, Munduruku adults’ Weber fraction was estimated to be .17 (approximately 6.7), whereas French adults’ was estimated to be .12 (approximately 9.10). This divergence could plausibly be the result of differences in day-to-day engagement in numerical discrimination. Training studies in normally developing children and adults will be helpful in determining whether practice with numerical discrimination leads to increases in acuity.

Changes in working memory capacity and executive functioning may also affect numerical discrimination. There are improvements in both spatial and verbal working memory in the preschool and early elementary school years (for a review, see Cowan, 1997), as

5 We have included logarithmic fits (in fact, negative In fits) because these are perhaps more familiar to some readers. However, the data are more accurately fit by a power function with an exponent of −.55 (e.g., for previous data points from the literature and points from the current sigmoid model the least squares power function fit is $14.871x^{-0.5535}$, $R^2 = .9488$). The power function has the psychologically sensible behavior of predicting asymptotic performance to be a Weber fraction $> 0$ (as opposed to a negative In function that would allow for psychologically meaningless negative values for the Weber fraction as seen in the trend lines of Figure 4). The power function is the most likely group trend given the current data and should be used in all future studies. The logarithmic trends in Figure 4 are given solely for ease of communication with readers more familiar with these functions.
well as in aspects of executive functioning such as inhibition and cognitive flexibility (Espy, Kaufmann, McDiarmid, & Glisky, 1999; Happaney, Zelazo, & Stuss, 2004). These factors correlate with math achievement in tasks that focus on symbolic and explicit mathematical reasoning (Espy et al., 2004; Gathercole & Pickering, 2000; McClelland, Acock, & Morrison, 2006; McClelland et al., 2007). In our task, although working memory and other executive functions were likely needed to maintain multiple arrays in memory and to focus on the dimension relevant to the task (i.e., number rather than area), differences in these factors were minimized by lengthening the display times for younger children. Nonetheless, an important future direction is to determine how developmental changes in these nonnumerical abilities affect deployment of the ANS.

Orthogonal to changes in ANS acuity, we also found that 3- and 4-year-old children exhibited an increased likelihood of guessing randomly, especially on harder numerical comparisons. Because we presented the identical task and identical numerical ratios to all of our participants, task difficulty varied considerably with age. For adults, most of the numerical discriminations we presented were easy. For 3-year-olds, most of the discriminations were hard. It is possible that the high rates of guessing shown by the youngest children resulted from general discouragement in the face of so many hard problems. Developmental change in executive function is another likely contributor to this difference across age groups (Happaney et al., 2004). Six-year-olds remained vigilant and attempted to respond correctly on the most difficult discrimination trials (such as 9:10), even though the acuity of their ANS supported a level of accuracy of only 62% correct. That they remained vigilant is seen in Figure 3, where percentage correct on difficult ratios for this age group did not deviate from the predictions of the psychophysics model. Future studies may rely on a staircase approach that tailors the ratios presented to focus on a level of difficulty appropriate for each child (e.g., between 1:2 and 5:6 for 3-year-olds). Such an approach may result in lower guessing rates for younger children. It may not, however, reveal considerably different estimates in ANS acuity from those reported here, because the sigmoid model used here allowed us to control for rates of random guessing while estimating acuity. Still, the present cross-sectional data provide developmental milestones of ANS acuity that can guide future investigations into nonnumerical contributors to children’s performance in numerical tasks.

Our task was designed to neutralize nonnumerical correlates of number by varying the degree to which number and both area and contour length were correlated across trial types. However, it is important to note that we used only a single task to estimate numerical acuity, and we did not measure any potential increases in acuity for any nonnumerical dimensions (e.g., changes in the ability to discriminate differences in surface area). Without such measures, it remains possible that the developmental changes identified in the present experiment are the result of more general developmental changes in magnitude representations, as opposed to changes that are specific to numerical representations. Previous research with both children and nonhuman animals suggests that the representation of numerical magnitude may rely on a format also shared by temporal magnitude (Brannon, Suanda, & Libertus, 2007; Feigenson, 2007; VannMarle & Wynn, 2006; Walsh, 2003) and perhaps by continuous extent (Brannon, Lutz, & Cordes, 2006; Feigenson, 2007; Walsh, 2003; Zorzi, Priftis, & Umilta, 2002).

Whether number, time, and space all share a common representational format remains a question of great theoretical interest. The developmental changes in number discrimination that we describe here provide what we hope will be a useful comparison point for future work investigating developmental changes in the discrimination of time, area, volume, and other magnitudes.

In summary, the ability to nonverbally approximate number plays a role in quantitative reasoning throughout the human life span, even after the ability to represent exact integers is attained. Cognitive psychology and cognitive neuroscience have made much recent progress in understanding this ability. However, little research has addressed the development of the ANS after infancy. Here we show that the acuity of the ANS continues to increase between ages 3 and 6 years and does not reach adultlike levels until some time during the preteen years. The protracted nature of ANS development, spanning the period when symbolic mathematical instruction begins, has implications both for math education and for our understanding of the interplay between individual experience and the “number sense.”

References
Barth, H., Kanwisher, N., & Spelke, E. S. (2003). The construction of large number representations in adults. Cognition, 86, 201–221.
Barth, H., LaMont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. S. (2006). Non-symbolic arithmetic in adults and young children. Cognition, 98, 199–222.
Barth, H., LaMont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. Proceedings of the National Academy of Sciences, 102, 14116–14121.


Received October 1, 2007

Revision received April 3, 2008

Accepted April 8, 2008