Supporting Information

Halberda et al. 10.1073/pnas.1200196109

SI Scatter Plot

An interactive version of this plot with enhanced details and discussion can be viewed at http://www.halberdalab.net/pnas2012/.

Large samples of 10,000 data points offer the opportunity to interact with data in new ways, and interactive data plots allow the reader to explore patterns within the dataset on their own. Movie S1 moves through 74 y of the lifespan (ages 11-85 y) in 8 s. Each person from Experiments 1 and 2 (N = 13,554) emerges as the plot arrives at their age at the time of testing, and then their dot fades away as the next age emerges. The color of each dot indexes the person's self-reported school mathematics ability and the dot appears at the intersection of their personal Weber fraction (w) and Response Time (RT) during the Approximate Number System (ANS) dots test. Plus symbols and ovals represent group means and the 10th – 90th percentile ranges respectively for groups of subjects with similar school mathematics ability (e.g., the red plus sign and oval represent individuals who reported being in the top 20% in school mathematics relative to their peers while the yellow plus sign and oval represent individuals who reported being in the bottom 20% relative to their peers). The plus sign is placed at the group mean and the 10th to 90th percentile range determines the size and shape of the ovals surrounding each mean, with the color matching the mathematics ability score for that group. For the ovals, the e.g., left side of the oval is the position of the 90th percentile in w (i.e., good performance) for that group and the right side of the oval is the 10th percentile (i.e., poor performance) for that group. The top and bottom of the ovals represent the 10th and 90th percentile positions for RT.

Viewing the means and ranges over the lifespan reveals 3 major transitions in the development of ANS precision. These transitions can be seen in Movie S1 and with enhanced detail at http://www.halberdalab.net/pnas2012/.

In the first movement (~10-16 y) the data cloud emerges as RTs move down toward 600 ms.

In the second movement (~16-30 yrs.) the data cloud shifts towards the lower left of the graph (i.e., better ANS performance and lower w). Best performance occurs at ~30 y.

In the third movement (~30-81 y) the data cloud dissolves and becomes erratic, and performance drifts up and out from the lower left hotspot (towards worse ANS performance). Some of this erratic behavior is caused by low sample size at the older ages.

These transitions in ANS precision across the lifespan are made visible when pooling evidence across thousands of individuals.

The relationship between ANS precision (w, RT) and school mathematics ability is also apparent in the positions of the group means and ovals, as individuals with better school mathematics ability (i.e., red plus sign) tend to be to the lower left (i.e., better ANS performance) relative to individuals with poorer school mathematics ability (e.g., green or yellow plus sign) especially in the years of peak performance (~15-50 y).

The interactive version of this plot can be viewed at http://www.halberdalab.net/pnas2012/, where readers can explore the raw data on their own, e.g., moving through ages by controlling the cursor, adding or removing data points based on mathematics ability, removing subject means, and saving various versions of movies for offline viewing.

SI Decile Plot

An interactive version of this plot with enhanced details and discussion can be viewed at http://www.halberdalab.net/pnas2012/.

The relationship between w, RT, and mathematical ability across the lifespan can also be experienced when interacting with pooled data across time. In the 4-dimensional Movie S2 (w, RT, mathematics ability, age), changes in age occur as the viewer watches the data unfold in time. Local averaging of individuals with similar w and RT scores (determined by decile rank) organize individuals in the form of a 10×10 heat-map with color indicating the average self-reported mathematical ability for each of the 100 w/RT decile subgroups (i.e., each person is placed into a single cell of the 10×10 grid as a function of their personal w and RT rank relative to their age group). Cells with fewer than five subjects are shown in black (occurring at the older and youngest ages). The color of each decile pair is determined by taking an average of the performance of individuals within each cell and placing the highest and lowest averages within each age at the extreme color values (i.e., red, yellow). The remaining averages fall between these extreme values on a linear scale. Blending of each cell between ages occurs in five equal steps while applying a running average with a window size of ±2 years around each discrete age time slice.

The relationship among w, RT, mathematics ability and age appears as a rainbow in the Movie S2, where the bottom left corner (i.e., better ANS precision with faster RT and lower w) tends towards the red end of the spectrum (i.e., better mathematics ability) and the top right corner (i.e., poorer ANS precision with longer RT and larger w) tends towards the green end of the spectrum (i.e., worse mathematics ability). As age progresses the colors in the Movie S2 scintillate somewhat, but the gradient from better mathematics ability to poorer mathematics ability remains visible across the lifespan. People who report being better in school mathematics (i.e., red-to-blue) tend to do better on the ANS test (i.e., have faster RT and lower w in the bottom left corner of the graph). People who report being worse in school mathematics (i.e., yellow-to-green) tend to do worse on the ANS test (i.e., have slower RT and higher w in the top right corner of the graph).

The interactive version of this plot can be viewed at http://www.halberdalab.net/pnas2012/.

SI Results

Correlation Between ANS Precision and School Mathematics Ability. We explored the stability of the relationship between ANS precision (w, RT) and school mathematics ability in several ways: (*i*) in the age decile plots of Fig. 2*C*, (*ii*) in Movie S1 and Movie S2, and (*iii*) in regression analyses that control for age (main text). A fourth way of displaying this result is in figures that plot the β-weights derived from linear regressions of ANS precision (w, RT) with self-reported school mathematics ability. These values, displayed in Fig. S1, indicate the slope of the linear regression relating ANS precision (w, RT) to self-reported school mathematical ability. The size of each dot is proportional to the number of individuals of that age in the dataset and the color of each dot is determined by the statistical significance for that value (i.e., red, $P < .05$; blue, $P > .05$). With the exception of the volatility that occurs for ages above 60 y where sample sizes are small for each individual year, the value of the β-weights for w and RT remain fairly stable. This further suggests a relationship between ANS precision (w, RT) and school mathematics ability that is stable across ages.

Correlation Size and Correlation Source. The precise size of the correlation between ANS precision (*w*, RT) and school mathematics ability at various ages remains to be definitively determined. The measures used here (e.g., self-report and a brief online ANS assessment) likely include measurement noise that limits our ability to estimate the true size of the correlation in the population. For this reason, we focused on the stability of this relationship across the lifespan rather than on its size.

One important open question concerns the source of the observed correlation. Actual academic performance emerges from a combination of many factors. ANS precision may be related to some of these factors more than others. If the source of the correlation found here derives from ANS precision (*w*, RT) correlating with mathematics-specific intellectual ability, the self-report measure used in the present study–which measures one's self-perceived ability and not one's actual school performance–may limit our ability to estimate the true effect size. We explored this concern in two ways: (*i*) by comparing ANS precision (*w*, RT) to scores on a nationally standardized test of academic achievement [the Scholastic Assessment Test (SAT) college entrance exam], and (*ii*) by investigating correlations between SAT scores and participants' responses on our self-reported measures of school abilities.

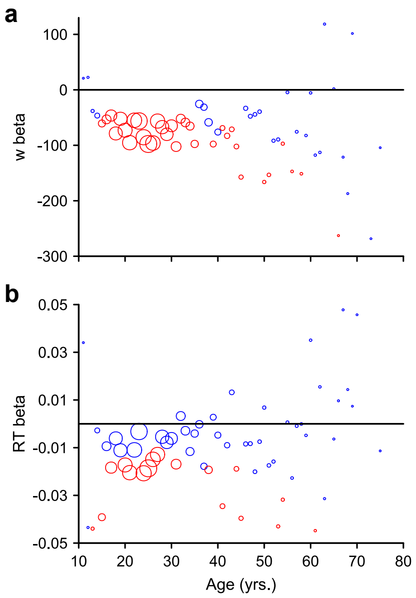


Fig. S1. Supplementary Results. (*a* and *b*) β-weights derived from linear regressions of ANS precision (*w*, RT) on self-reported school mathematics ability for each year. The size of each dot is proportional to the number of individuals of that age in the dataset and the color of each dot is determined by the statistical significance for that β-weight (i.e., red: $P < .05$; blue: $P > .05$).

In experiment 2, we had participants report their scores from the most widely used college entrance exam in the United States (i.e., the SAT). Participants also reported their confidence in their memory for these scores. Inclusion of all participants irrespective of confidence (n = 458) revealed a relationship between performance on the mathematics portion of this examination (i.e., SAT quantitative) and ANS precision (*w*, RT), r = -.24 (main text, *Results and Discussion*). An analysis restricted to only those subjects who reported being 100% confident in their memory for this score (n = 139) returned a similar result ($P_w = 3.04 \text{ x } 10^{-3}$, $P_{RT} = 1.95 \text{ x } 10^{-1}$, $r_{w\&RT} = -.27$) suggesting that this relationship was fairly robust across various levels of confidence.

As a validation of our self-report measures, we explored correlations between self-reported school abilities and the content-appropriate subtests of the SAT–an objective measure of these abilities. Highlighted cells in Table S1 display the correlations between our self-report measures of school abilities and the appropriate SAT subtest (SAT quantitative, SAT verbal, SAT writing). Fewer individuals have a SAT writing score because the writing component was only recently added to this exam. Of particular interest, self-reported school mathematics ability correlated most strongly with SAT quantitative performance while SAT writing and SAT verbal correlated most strongly with self-reported writing class ability (the three highlighted cells in Table S1). Assuming that SAT scores were recalled correctly, this suggests that our self-report measures were accurately assessing the targeted content areas of interest.

Table S1. Correlations of SAT performance and self-reported school abilities

		Self-report		
SAT	*n*	Math	Science	Writing
Quantitative	458	.67	.42	-.01
Verbal	454	.12	.23	.41
Writing	220	.23	.32	.45

Values displayed are *r* values from linear correlations.

The relationship between our self-report measures and the appropriate subtests of the SAT was specific to the content areas of interest as revealed by partial correlations that controlled for other measures. In Table S2, each of the three highlighted correlations from Table S1 is reanalyzed as a partial correlation controlling for the SAT and self-report measures unrelated to that content area. The logic of these partial correlations controls for performance in other content areas as well as factors such as general IQ, test-performance anxiety and response biases (e.g., evaluating oneself as "above average" across all measures irrespective of content). Results revealed a specific relationship between each SAT subtest and the appropriate content area of the self-report questionnaire. This further validated our self-report questionnaire measures.

Table S2. Partial correlations of SAT performance and self-reported school abilities

Correlated variable		Controlled variable	
SAT–Self-report	*n*	SAT / Self-report	r_p
Quantitative–Math	214	Verbal, Writing / Writing	.65
Verbal–Writing	445	Quantitative / Math, Science	.45
Writing–Writing	220	Quantitative / Math, Science	.44

Values displayed are partial correlations controlled for other noted variables

Self-reported academic abilities have been shown to correlate with academic performance most accurately when respondents judge their abilities relative to other people's abilities under conditions of anonymity with an expectation that their abilities will later be assessed by an exam (1, 2). This was the approach taken in our experiments, combined with the added benefit of an explicitly stated norming group.

Of potential theoretical interest, self-report and actual academic performance may be dissociable and track somewhat independent aspects of our psychology (3). These may each capture unique variance in ANS precision (*w*, RT) as separable components. In support of this suggestion, in experiment 2, for subjects that entered an SAT quantitative score (n = 458), we found that partial correlations including both SAT quantitative (i.e., the mathematics subtest of the college entrance exam) as a measure of academic performance and self-reported school mathematics ability (available from the online questionnaire) as predictors of ANS precision (*w*, RT) returned significant coefficients for both SAT quantitative and self-report. Specifically, SAT quantitative correlated most strongly with *w* while self-reported mathematics ability correlated most strongly with RT (Table S3). This suggests that both self-perceived abilities and actual academic performance may be important contributors to an overall correlation between ANS precision (*w*, RT) and formal mathematics ability.

Exploring the relationships between ANS precision (w, RT) and subcomponents from the many factors that give rise to school mathematics performance (e.g., intellectual ability, mathematics anxiety, self-perceived abilities, school and home environment etc) is an important avenue for future exploration. The current results reveal a relationship between ANS precision (w, RT) and school mathematics ability that remains stable throughout adolescence and adulthood, but the causal relationships at play remain to be explored. The large individual differences and prolonged development of the number sense (throughout the school-age and the adult years), paired with its consistent and specific link to mathematics ability, are encouraging for the potential impact of educational interventions that target the number sense across a wide range of ages.

Table S3. Correlations of SAT quantitative, self-reported school mathematics ability, and ANS precision

Correlated variable	*n*	Controlled variable	r_p	*P*
SAT quantitative–w	458	Self-report Math, RT	−.16	6×10^{-4}
Self-report Math–RT	458	SAT quantitative, w	−.13	6×10^{-3}

r_p values are correlations controlled for other noted variables.
P values represent the probability of obtaining the observed correlation in a sample of data by random chance when there is truly no relation in the population.

Descriptives. Fig. S2 presents histograms of self-reported school abilities and performance metrics from the ANS dots test from experiment 1 (N = 10,548) (results from experiment 2 were similar).

Fig. S2 a–c presents participants' self-reported school mathematics, science, and writing ability. These responses were made on a continuous scale with a moveable slider on a bar that ranged from 0 (below average) to 100 (above average). The default setting for the slider on each scale was 50 (i.e., average ability). Many participants chose to leave some of the sliders at this setting. Over concerns that some participants may have skipped the questionnaire without considering the sliders, we decided to remove any participant who left two or more sliders at the default setting. This removed 317 individuals from our initially larger sample of 10,907 individuals in experiment 1. This was a small percentage of our subjects (i.e., <3%) and all results remained similar and significant if these subjects were retained.

The trend towards subjects reporting that they have "above average" abilities is salient in Fig. S2 a–c. Any uniform score inflation across the slider values (i.e., everyone over-estimating their abilities) would not have affected our regression analyses because we explored individual differences in abilities relative to peers and not absolute values for those abilities. However, some individuals may overestimate, some underestimate, while others may accurately report their abilities (1). Over- and underestimation are interesting behaviors in their own right and future investigations of how such behaviors relate to ANS precision may be fruitful. The presence of a variety of such deviations spread throughout our sample may have hampered our ability to detect a relationship between ANS precision (w, RT) and self-reported abilities because of noise in the self-report estimate of school abilities.

Figure S2 d and e display the scores for w and RT observed in the sample. The range of observed w and RT scores suggests large individual differences (e.g., a w of 0.11 predicts that the person would correctly identify the color that had more dots on 75% of the trials involving 18 blue versus 20 yellow dots; while an individual with a w of 0.5 would need a larger difference of 18 blue versus 27 yellow dots in order to obtain this level of performance). Individual differences in w and RT scores appear largely uncorrelated (r = -.11; *Results and Discussion*) suggesting they may index independent abilities.

1. Ackerman PL, Wolman SD (2007) Determinants and validity of self-estimates of abilities and self-concept measures. *J Exp Psychol Appl 13*:57–78.
2. Mabe PA, III, West SA (1982) Validity of self-evaluation of ability: A review and metaanalysis. *J Appl Psychol 67*:280–296.
3. Greven CU, Harlaar N, Kovas Y, Chamorro-Premuzic T, Plomin R (2009) More than just IQ: School achievement is predicted by self-perceived abilities—but for genetic rather than environmental reasons. *Psychol Sci 20*:753–762.

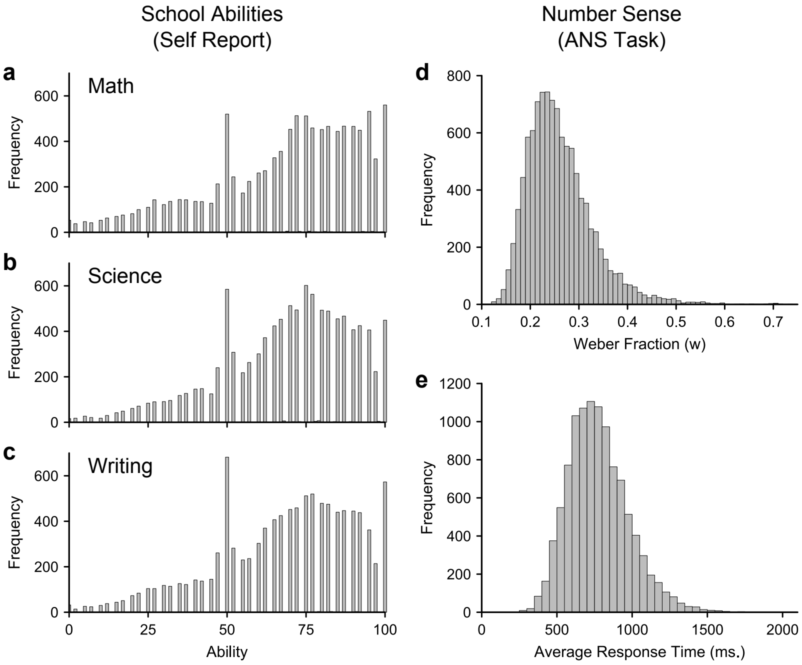
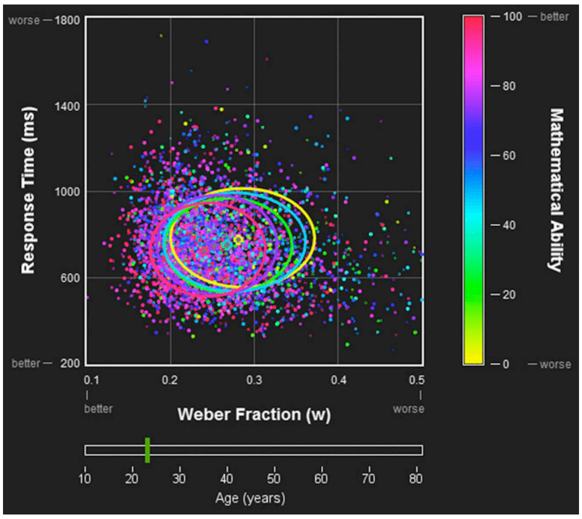
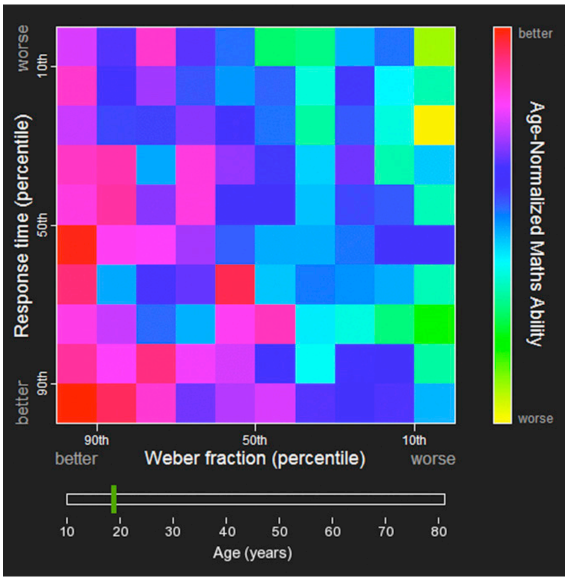


Fig. S2. Supplementary Results. (a-c) Frequency distributions of the value set by participants in experiment 1 for self-reported academic abilities. (d-e) Histograms of w and RT scores on the ANS dots test from experiment 1.



Movie S1. Interactive scatter plot. Each person from experiments 1 and 2 (*n* = 13,554) emerges as the plot arrives at their age at the time of testing, and then their dot fades away as the next age emerges. The color of each dot indexes the person's self-reported school mathematics ability and the dot appears at the intersection of their personal w and RT during the ANS dots test. Plus symbols (+) and ovals represent group means and the 10th to 90th percentile ranges, respectively, for groups of subjects with similar school mathematics ability (e.g., the red plus sign and oval represent individuals who reported being in the top 20% in school mathematics relative to their peers whereas the yellow plus sign and oval represent individuals who reported being in the bottom 20% relative to their peers).

Movie S1



Movie S2. Interactive decile plot. This movie presents local averaging of individuals with similar *w* and RT scores (determined by decile rank) organizing individuals in the form of a 10×10 heat-map with color indicating the average self-reported mathematical ability for each of the 100 *w*/RT decile subgroups (i.e., each person is placed into a single cell of the 10×10 grid as a function of their personal *w* and RT rank relative to their peers). Cells with fewer than five subjects are shown in black (occurring at the older and youngest ages). The color of each decile pair is determined by taking an average of the performance of individuals within each cell and placing the highest and lowest averages within each age at the extreme color values (i.e., red, yellow). The remaining averages fall between these extreme values on a linear scale. Blending of each cell between ages occurs in five equal steps while applying a running average with a window size of ±2 y around each discrete age time slice.

Movie S2