Supplemental Materials

Positive attitude towards math supports early academic success: behavioral evidence and neural mechanisms

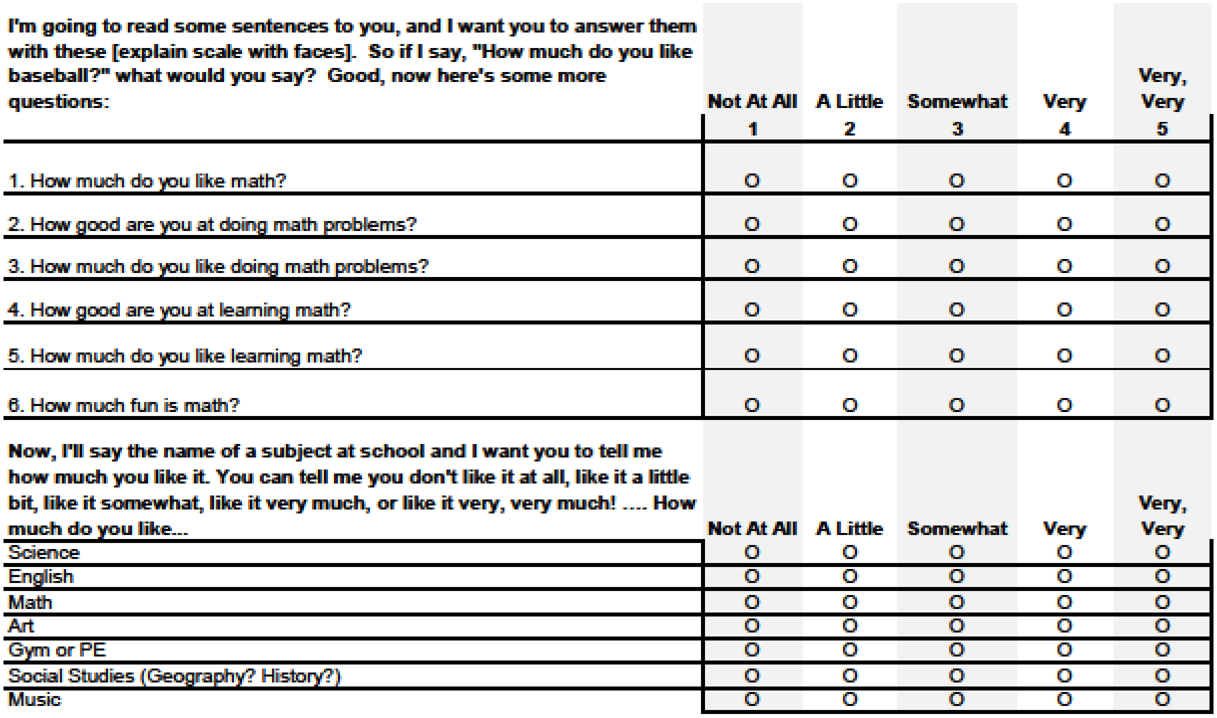
Lang Chen[1,], Se Ri Bae[1], Christian Battista[1], Shaozheng Qin[1], Tianwen Chen[1], Tanya M. Evans[1],*

Vinod Menon[1,2,3,]*

Table of Contents

I. Supplemental Tables S1-10

Table S1. Questionnaire assessing positive attitude towards math (PAM) and general attitude towards academics (GAA).



I'm going to read some sentences to you, and I want you to answer them with these [explain scale with faces]. So if I say, "How much do you like baseball?" what would you say? Good, now here's some more questions:	**Not At All 1**	**A Little 2**	**Somewhat 3**	**Very 4**	**Very, Very 5**
1. How much do you like math?	○	○	○	○	○
2. How good are you at doing math problems?	○	○	○	○	○
3. How much do you like doing math problems?	○	○	○	○	○
4. How good are you at learning math?	○	○	○	○	○
5. How much do you like learning math?	○	○	○	○	○
6. How much fun is math?	○	○	○	○	○

Now, I'll say the name of a subject at school and I want you to tell me how much you like it. You can tell me you don't like it at all, like it a little bit, like it somewhat, like it very much, or like it very, very much! .... How much do you like...	**Not At All**	**A Little**	**Somewhat**	**Very**	**Very, Very**
Science	○	○	○	○	○
English	○	○	○	○	○
Math	○	○	○	○	○
Art	○	○	○	○	○
Gym or PE	○	○	○	○	○
Social Studies (Geography? History?)	○	○	○	○	○
Music	○	○	○	○	○

Notes: Questions 1 through 6 assess PAM, and the last 7 questions assess GAA, which includes a combination of subjects: Science, English, Art, Gym or PE, and Social Studies. Note that the question “How much do you like math?” was originally repeated in the GAA but was excluded from data analysis in this study.

Table S2. Factor analysis of positive attitude survey.

Survey Questions	2-Factor model		3-Factor model		
			PAM		
	PAM	GAA	Interest	Self-perceived Ability	GAA
How much do you like math?	**0.82**	0.07	**0.82**	0.19	0.05
How much do you like learning math?	**0.80**	0.07	**0.84**	0.10	0.06
How much fun is math?	**0.73**	0.10	**0.78**	0.06	0.09
How much do you like doing math problems?	**0.81**	0.04	**0.81**	0.18	0.03
How good are you at doing math problems?	**0.55**	-0.10	0.22	**0.8**	-0.11
How good are you at learning math?	**0.54**	0.12	0.24	**0.76**	0.11
How much do you like science?	0.05	**0.56**	-0.14	0.43	**0.56**
How much do you like English?	0.20	**0.45**	0.21	0.04	**0.45**
How much do you like art?	-0.21	**0.45**	-0.12	-0.21	**0.45**
How much do you like gym?	0.03	**0.44**	0.01	0.08	**0.44**
How much do you like social studies?	0.20	**0.68**	0.16	0.15	**0.68**
How much do you like music?	-0.05	**0.71**	0.07	-0.24	**0.71**

Correlation	PAM	GAA		PAM		GAA
				Interest	Self-perceived Ability	
PAM	\		Interest	\		
GAA	0.12	\	Self-perceived Ability	**0.37*****	\	
			GAA	0.12	0.07	\

Notes: In the 2-Factor model, Factor 1 reflects positive attitude towards math (PAM) while Factor 2 reflects attitude towards general academics (GAA). In the 3-Factor model, Factor 1 reflects the strong interest towards math component while Factor 2 reflects the self-perceived ability component of PAM, and Factor 3 reflects GAA.

Table S3. Correlation matrix of behavioral measures.

	Age	FSIQ	CBCL AnxDep	CBCL Withdraw Dep	CBCL Anxiety	WM Digit Rec	WM Block Rec	WM Count Rec	WM Back Rec	SEMA	GAA	PAM	WIAT Word Read	WIAT Read Comp	WIAT Num Ops	WIAT Math Reason
Age	1															
FSIQ	-0.015	1														
CBCL AnxDep	0.059	-0.081	1													
CBCL WithdrawDep	-0.022	-0.001	***0.554***	1												
CBCL Anxiety	0.026	-0.066	***0.785***	***0.498***	1											
WM Digit Rec	-0.088	***0.265***	-0.027	-0.050	-0.001	1										
WM Block Rec	-0.125	***0.280***	-0.036	-0.007	-0.007	**0.137**	1									
WM Count Rec	0.036	***0.468***	-0.009	-0.081	-0.027	***0.283***	***0.314***	1								
WM Backward Rec	0.012	***0.340***	0.075	-0.001	0.014	***0.293***	***0.290***	***0.416***	1							
SEMA	-0.014	***-0.320***	0.055	0.101	0.015	***-0.195***	***-0.188***	***-0.246***	-0.070	1						
GAA	-0.036	-0.012	-0.064	**-0.154**	-0.009	0.073	0.072	-0.004	0.028	-0.083	1					
PAM	***-0.296***	***0.199***	-0.029	0.026	-0.022	0.052	***0.250***	***0.202***	0.083	***-0.315***	**0.148**	1				
WIAT Word Read	-0.041	***0.533***	0.100	0.053	0.053	***0.309***	***0.183***	***0.400***	***0.275***	***-0.206***	-0.034	0.060	1			
WIAT Read Comp	0.087	***0.573***	-0.022	-0.072	-0.060	***0.332***	**0.132**	***0.333***	***0.247***	***-0.183***	-0.004	0.099	***0.634***	1		
WIAT Num Ops	-0.030	***0.443***	-0.025	-0.005	-0.024	**0.144**	***0.370***	***0.442***	***0.281***	***-0.203***	-0.012	***0.282***	***0.373***	***0.349***	1	
WIAT Math Reason	-0.065	***0.705***	-0.035	0.004	-0.062	***0.310***	***0.321***	***0.511***	***0.304***	***-0.396***	-0.088	***0.301***	***0.573***	***0.507***	0.642	1

Notes: FSIQ = Weschler Abbreviated Scale of Intelligence (WASI) Full Scale IQ, CBCL = Child Behavior Checklist (AnxDep = Anxious/Depressed Subscale, WithdrawDep = Withdrawn/Depressed Subscale) WM = Working Memory Test Battery for Children (Rec = Recall, Backward = Backward digit), SEMA = Scale for Early Math Anxiety; GAA = General attitude towards general academics, PAM = Positive attitude towards math; WIAT = Weschler Individual Achievement Test – Second Edition (Word Read = Word Reading, Read Comp = Reading Comprehension, Num Ops = Number Operations, Math Reason = Math Reasoning). ***Bold*** = $p < .001$; ***Bold*** = $p < .01$; **Bold** = $p < .05$.

Table S4. Regression analysis with composite and subcomponent scores for positive attitude towards math (PAM).

	PAM composite score						**Strong interest subcomponent score**						**Self-perceived ability subcomponent score**					
	Numerical Operations			Math Reasoning			Numerical Operations			Math Reasoning			Numerical Operations			Math Reasoning		
	β	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*
Age	-0.00	0.06	-0.02	-0.04	0.07	-0.87	-0.02	0.06	-0.28	-0.05	0.05	-1.11	-0.03	0.06	-0.51	-0.06	0.05	-1.25
IQ	0.28	0.07	3.93***	0.54	0.05	10.44***	0.28	0.07	4.02***	0.55	0.05	10.49***	0.27	0.07	3.79***	0.54	0.05	10.26***
Gender	0.17	0.13	1.35	0.37	0.09	3.93***	0.20	0.13	1.56	0.38	0.09	4.10***	0.15	0.13	1.16	0.35	0.10	3.71***
Anx/Dep	-0.11	0.10	-1.12	-0.01	0.08	-0.18	-0.11	0.10	-1.09	-0.01	0.08	-0.16	-0.11	0.10	-1.09	-0.01	0.08	-0.17
Wdrawn/ Dep	-0.03	0.07	-0.36	-0.02	0.05	-0.46	-0.02	0.07	-0.32	-0.02	0.05	-0.42	-0.01	0.07	-0.10	-0.01	0.05	-0.26
Anx	0.05	0.10	0.51	-0.01	0.07	-0.16	0.05	0.10	0.53	-0.01	0.07	-0.15	0.02	0.10	0.23	-0.03	0.07	-0.38
GAA	-0.01	0.06	-0.18	-0.09	0.05	-1.72	-0.01	0.06	-0.10	-0.08	0.05	-1.65	0.00	0.06	0.04	-0.07	0.05	-1.56
WM composite	0.33	0.08	4.32***	0.22	0.06	3.99***	0.33	0.08	4.35***	0.23	0.06	4.03***	0.34	0.08	4.38***	0.23	0.06	4.05***
Math anxiety	0.01	0.07	0.12	-0.12	0.05	-0.23*	-0.02	0.07	-0.23	-0.14	0.05	-2.66**	-0.01	0.07	-0.20	-0.13	0.05	-2.54*
Attitude[#]	**0.22**	0.07	3.23**	**0.14**	0.05	2.62**	**0.19**	0.07	2.74**	**0.11**	0.05	2.17*	**0.17**	0.07	2.47*	**0.11**	0.05	2.16*

Notes: * < .05; ** < .01; ***< .001. [#]This attitude measure stands for either PAM composite score or the two subcomponent scores. Anx/Dep = Child Behavior Checklist: Anxiety/Depression Subscale; Wdrawn/Dep = Child Behavior Checklist: Withdrawn/Depression Subscale; Anx = Child Behavior Checklist: Anxiety Subscale; GAA = General attitude towards academics; WM = working memory.

Table S5. Regression analysis of hippocampal activation for positive attitude in Cohort 1.

	Right hippocampus			**Left hippocampus**		
	β	*S.E.*	*t-value*	*β*	*S.E.*	*t-value*
Age	-0.09	0.15	-0.63	-0.08	0.15	-0.53
IQ	0.11	0.16	0.67	0.06	0.17	0.32
Gender	0.40	0.28	1.40	-0.06	0.30	-0.20
WM composite	0.13	0.15	0.86	0.11	0.15	0.74
Math anxiety	0.01	0.16	0.08	0.10	0.17	0.58
PAM	0.35	0.15	2.31*	0.37	0.16	2.35*

Notes: * < .05. S.E. = standard error; IQ = intelligence quotient; WM = working memory; PAM = Positive attitude towards math.

Table S6. General Linear Model (GLM) results with Positive Attitude Towards Math (PAM) scores as covariate in Cohort 1.

GLM (PAM covariate)			Peak MNI Coordinates		
Region	Cluster size (voxels)	Peak z-score	x	y	z
Positive effect of PAM (Task - Rest)					
With small volume correction					
L Hippocampus	111	2.57	-36	-14	-26
R Hippocampus	55	2.83	38	-22	-16
Whole-brain GLM					
Dorsal Cerebellum	2197	4.71	10	-64	-28
R Lingual Gyrus		3.18	8	-52	-4
L Anterior FG	217	3.42	-40	-8	-34
L SFG/SMA	159	3.61	-12	4	66
Negative effect of PAM (Task - Rest)					
No significant voxels					

Notes: L = left; R = right; MNI = Montreal Neurological Institute; FG = Fusiform Gyrus; SFG = Superior Frontal Gyrus; SMA = Supplementary Motor Area.

Table S7. Region of Interest (ROI) analysis of affective-motivational system response correlated with Positive Attitude towards Math (PAM).

			Cohort 1		Cohort 2	
Source	ROI	Coordinates (MNI)	Beta value Mean (SD)	Correlation with PAM (*r*)	Beta value Mean (SD)	Correlation with PAM (*r*)
Gur et al. (2002)	Left Amygdala	(-20,-2,-18)	0.09 (0.75)	0.19	-0.12 (0.69)	0.11
	Right Amygdala	(21, -2, -9)	-0.09 (0.77)	-0.25	-0.19 (0.64)	0.20
Di Martino et al. (2008)	Left VS (C)	(-9, 9, -8)	0.10 (0.49)	0.12	-0.15 (0.52)	0.09
	Right VS (C)	(9, 9, -8)	0.00 (0.75)	-0.22	-0.13 (0.43)	0.06
	Left VS (NAc)	(-10, 15,0)	-0.08 (0.51)	0.04	-0.18 (0.52)	0.11
	Right VS (NAc)	(10,15, 0)	-0.05 (0.75)	-0.19	-0.20 (0.49)	-0.02

Notes: Theory-driven analyses using a priori ROIs from previous literature did not reveal any significant relationships between positive attitude towards math (PAM) scores and regions involved in affective-motivational system (i.e., amygdala and ventral striatum). None of the ROIs showed a significant correlation with PAM scores even at an uncorrected level of $p < .05$. MNI = Montreal Neurological Institute; SD = standard deviation; VS = Ventral Striatum; C = Caudate; NAc = Nucleus accumbens.

Table S8. General Linear Model (GLM) results with Postive Attitude Towards Math (PAM) scores as covariate of interest and correct retrieval rates as control covariate in Cohort 1.

GLM (PAM)			MNI Coordinates		
Region	Cluster size (voxels)	Peak z-score	X	y	z
Positive effect of PAM (Task - Rest)					
Dorsal Cerebellum	1319	4.29	10	-64	-28
L SFG/SMA	153	3.48	-10	4	64
Negative effect of PAM (Task - Rest)					
R nucleus accumbens	164	3.31	4	18	-6

Notes: MNI = Montreal Neurological Institute; L = left; R = right; SFG = superior frontal gyrus; SMA = supplementary motor area.

Table S9. Power analysis of Structural Equation Modeling (SEM) direct and indirect effect in both cohorts.

	Cohort 1 (n =47)	Cohort 2 (n =28)
Retrieval → Math	0.761	0.348
Hippo. Act. → Retrieval	0.918	0.712
PAM → Hippo. Act.	0.809	0.962
PAM → Hippo. Act. → Retrieval	0.594	0.558

Notes: PAM = Positive attitude towards math; Hippo = Hippocampus; Act = Activity.

Table S10. Comparison of model fit parameters of original and alternative partially-mediated and fully-mediated Structural Equation Modeling (SEM) models.

	chi-square (*df*)	CFI	TLI	RMSEA	SRMR	AIC	BIC
Partially-mediated (Original)	0.875 (1)	1.000	1.033	0.000	0.033	527.874	551.926
Partially-mediated (Alternative)	NA (0)	1.000	1.000	0.000	0.000	528.999	554.901
Fully-mediated (Original)	3.141 (3)	0.994	0.987	0.032	0.057	***526.14***	***546.492***
Fully-mediated (Alternative)	2.266 (2)	0.988	0.965	0.053	0.050	527.265	549.467

Notes: The bold/italic numbers highlight models with the lowest (better) AIC and BIC scores.

II. Supplemental Figures S1-S6

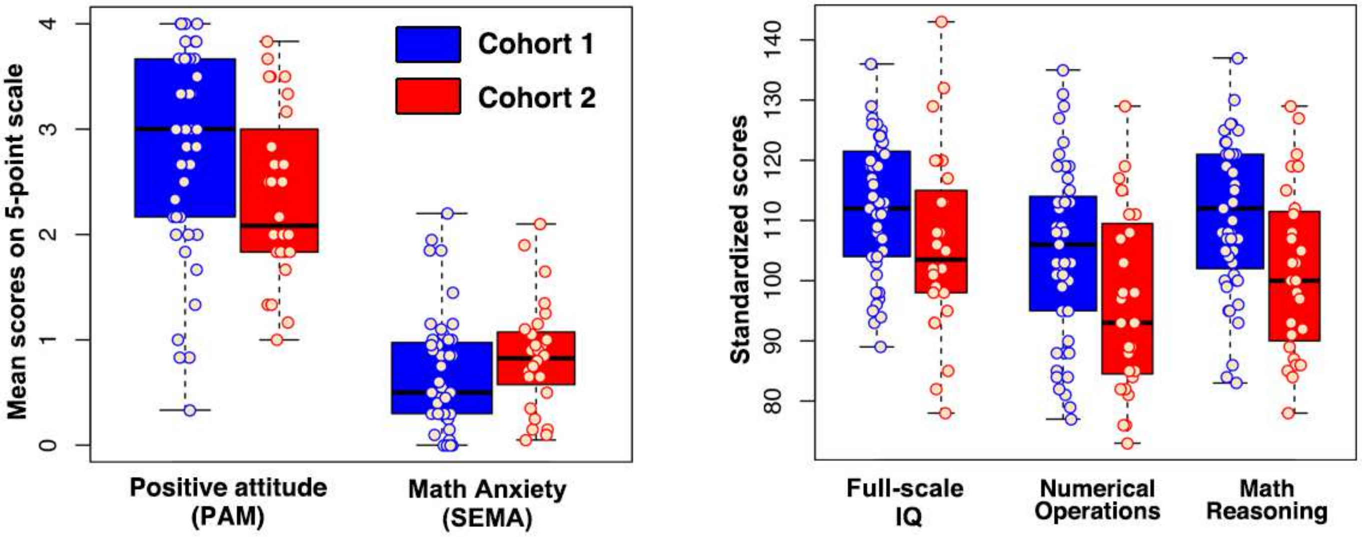


Figure S1. Comparisons of behavioral measure distributions in Cohorts 1 and 2. Overall distributions on PAM, SEMA, Full Scale IQ, Numerical Operations and Math Reasoning were not significantly different ($p > .05$) across cohorts. PAM = Positive attitude towards math, SEMA = Scale for early math anxiety.

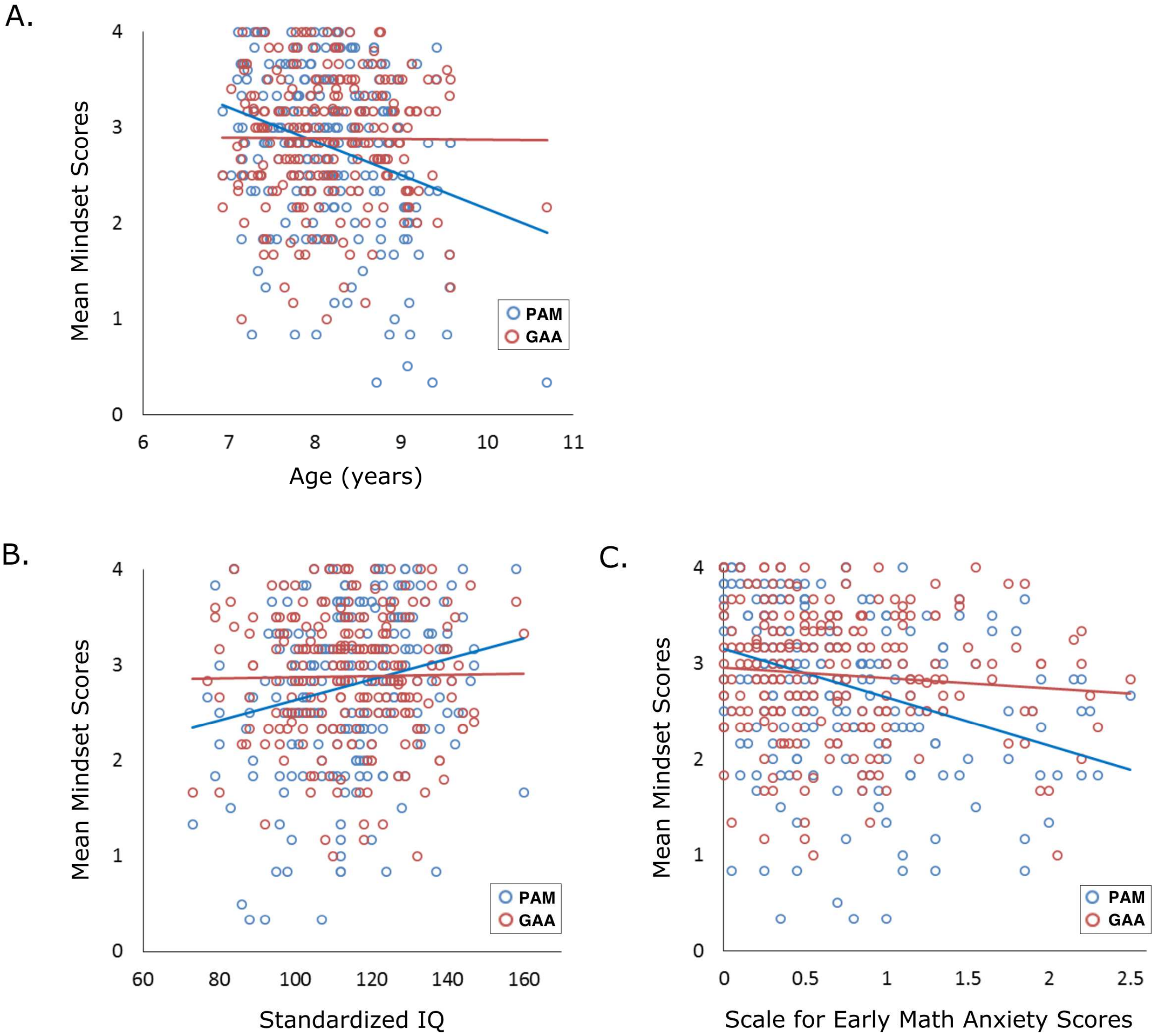


Figure S2. Distinct correlation patterns of PAM and GAA. PAM scores were negatively correlated with **(A)** age, *r* (237) = - .28, p < .001, 95% C.I. = [-0.39, -0.15], positively correlated with **(B)** standardized IQ, *r* (238) = .21, p < .01, 95% C.I. = [0.08, 0.33], and negatively correlated with **(C)** math anxiety, *r* (237) = - .33, p < .001, 95% C.I. = [-0.44, -0.21]. There were no linear trends for GAA. PAM = Positive attitude towards math; GAA = General attitude towards academics.

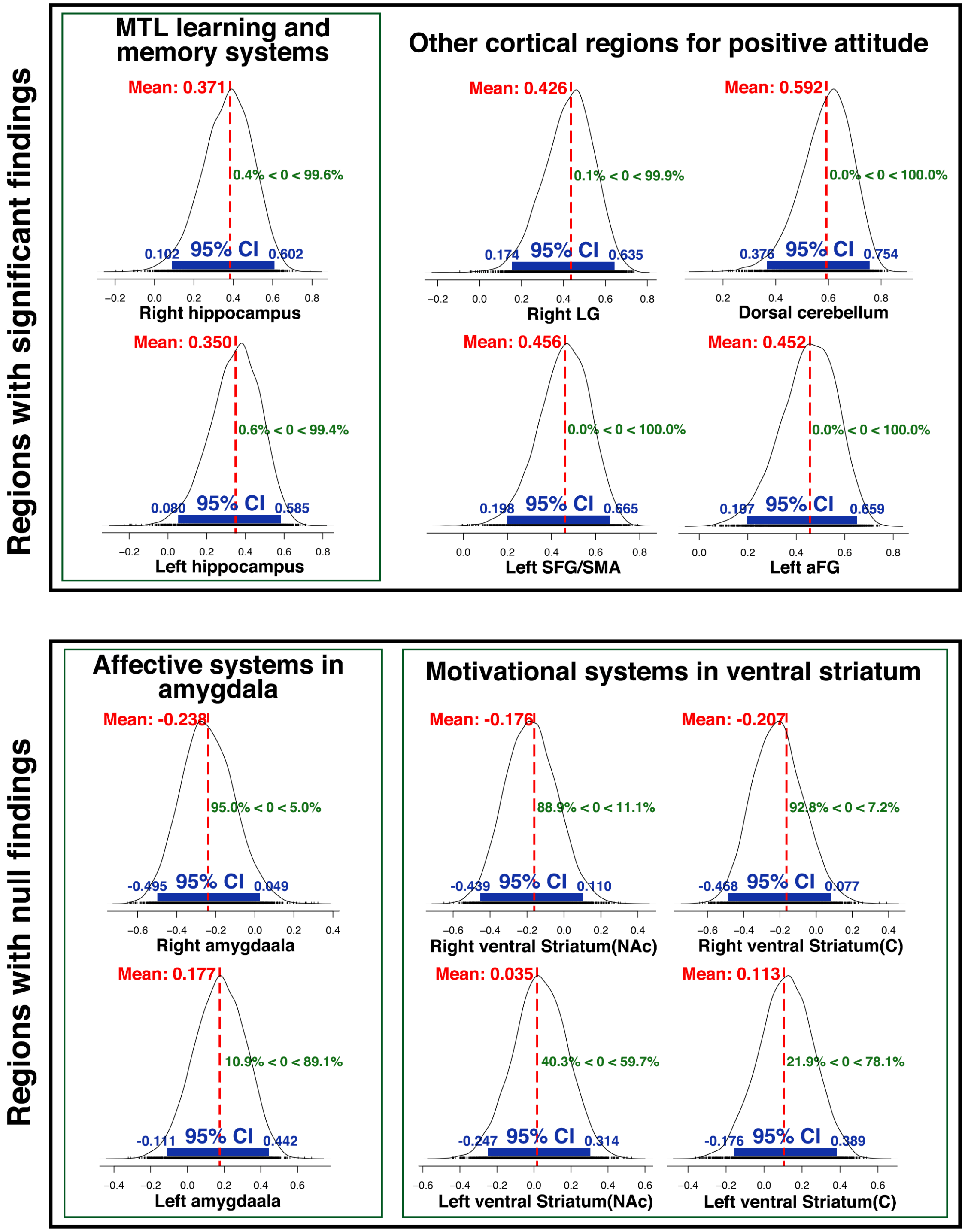


Figure S3. Density graphs of Markov-Chain Monte Carlo (MCMC) simulation for correlations between PAM scores and beta values from ROI analysis in respective brain regions for Cohort 1. In all simulations, a total of 3 chains, each with 5,000 permutations, were used. Estimates presented here were based on all chains. Blue bars on each graph display 95% confidence intervals of the Pearson's correlation ρ estimates based on observed correlation between beta values from ROI analysis in each region and PAM scores. Dashed red lines denote mean value of ρ. LG = lingual gyrus, SFG/SMA = superior frontal gyrus and supplemental motor area;

aFG = anterior fusiform gyrus; NAc = nucleus accumbens; C = caudate.

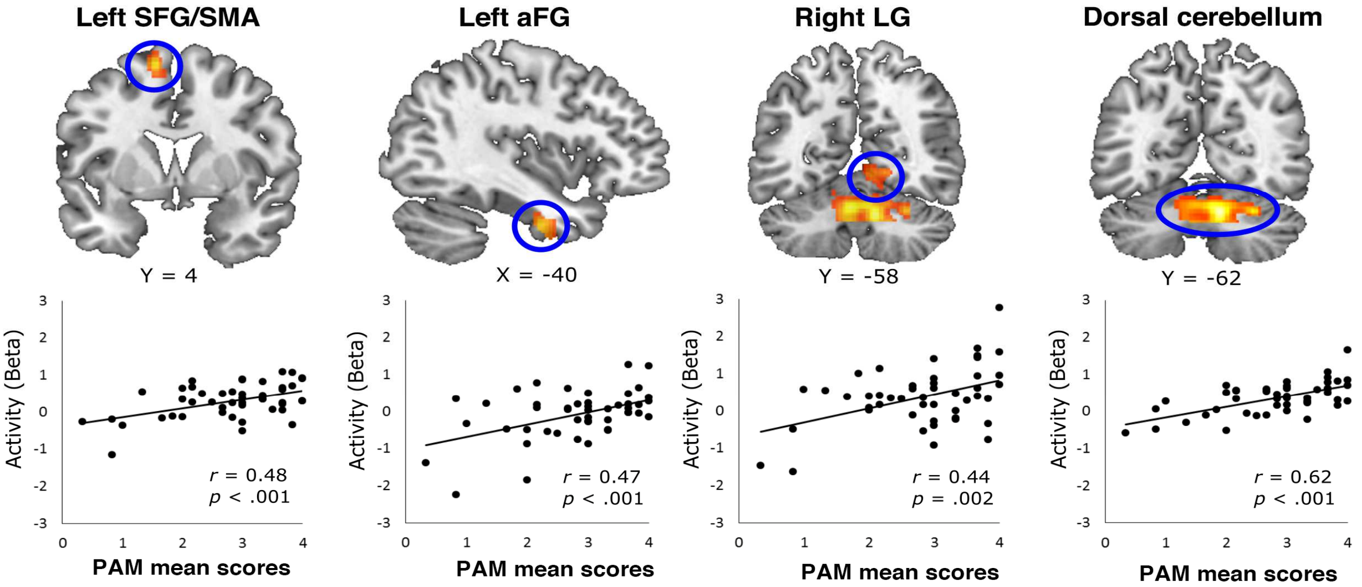


Figure S4. Whole-brain GLM results of positive attitude in Cohort 1. Brain regions positively correlated with PAM scores during arithmetic problem solving. SFG/SMA = Superior frontal gyrus/supplementary motor area; aFG = Anterior fusiform gyrus; LG = Lingual gyrus.

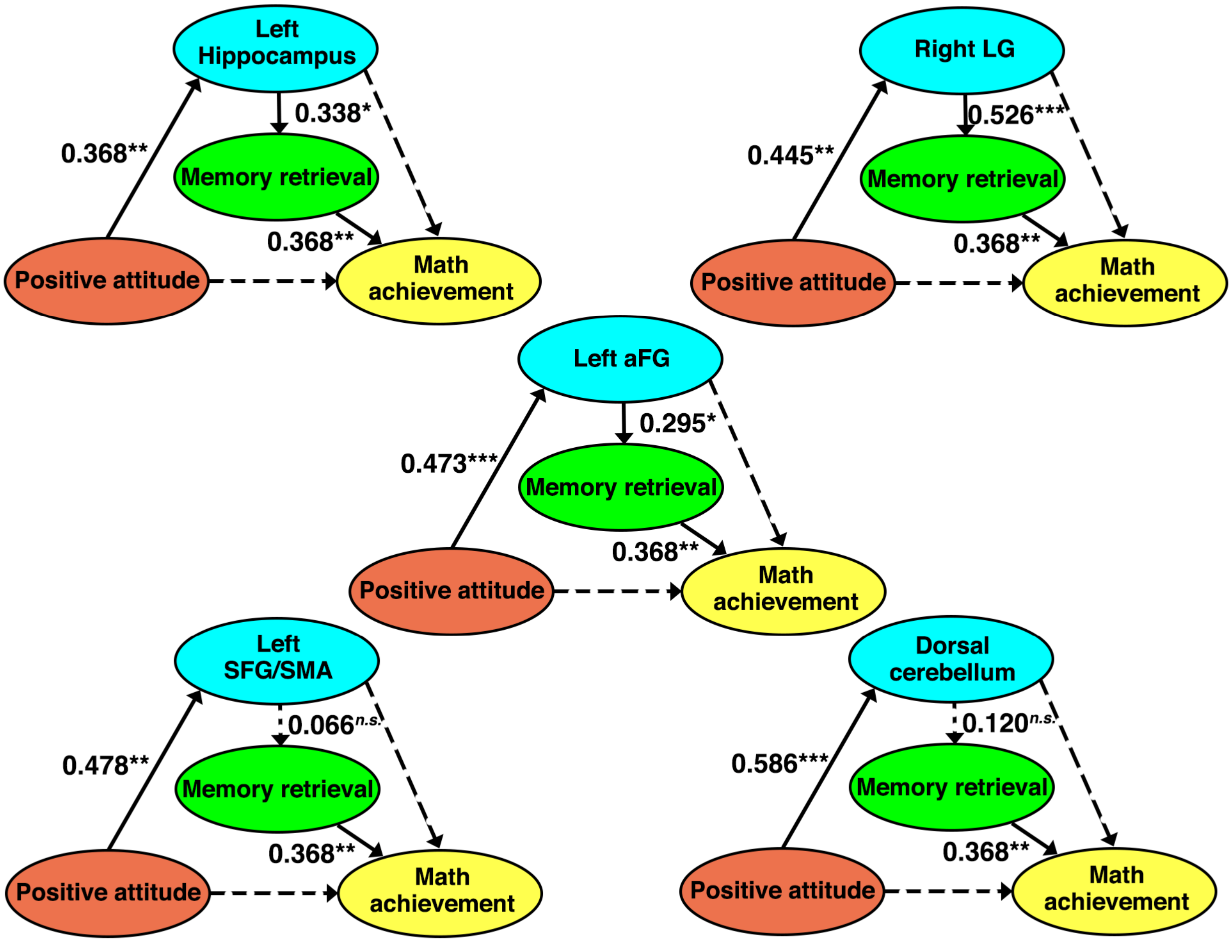


Figure S5. Full-retrieval-based SEM models with activation in left hippocampus, right lingual gyrus (LG), left anterior fusiform gyrus (aFG), left superior frontal gyrus/supplemental motor area (SFG/SMA) and dorsal cerebellum in Cohort 1. For other regions showing significant correlation with PAM score, we applied the same full-retrieval-based SEM model with beta averages from ROI analyses. Results revealed that left hippocampus, right lingual gyrus and left aFG fit the full mediation models whereas left SFG/SMA and dorsal cerebellum did not. Solid lines show significant path coefficients in the model, whereas dashed lines denote non-significant paths in the SEM hypothesized in other models. * $p < .05$, ** $p < .01$, *** $p < .001$.

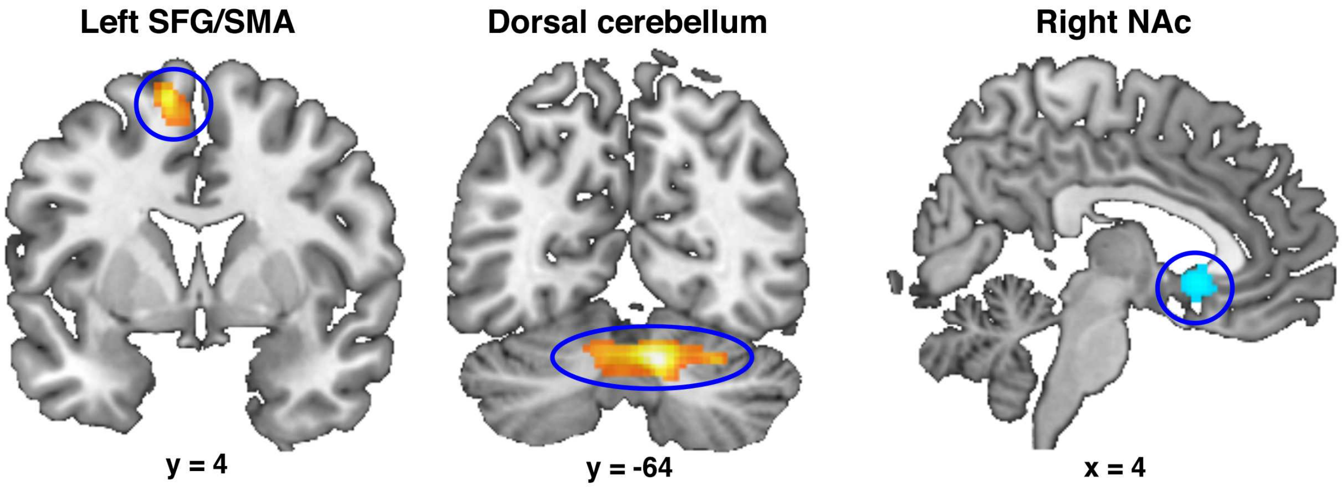


Figure S6. General Linear Model (GLM) results with Positive Attitude towards Math (PAM) (retrieval rate as control covariate) in Cohort 1. The superior frontal gyrus/supplementary motor area (SFG/SMA) and dorsal cerebellum showed significant positive correlations with PAM scores when retrieval rates were regressed out in the GLM analysis. The right nucleus accumbens (NAc) showed a negative correlation with PAM scores.

III. Supplemental Methods

Methods S1: Supplemental Information for Cohort 1

Factor analysis The exploratory factor analysis (EFA) was conducted in R (version 3.2.1; R Core Team, 2015) with the nFactors package (version 2.3.3; Raiche & Magis, 2010). Six questions related to the positive attitude towards math (PAM) and 6 questions assessing general attitude towards academics (GAA) were used for the EFA. Based on the Scree plot and Optimal Coordinate (Ruscio & Roche, 2012), a three-component solution was chosen. As described in the main text, these three components corresponded to the two subcomponents of PAM (strong interest and self-perceived ability), and GAA. The three-component solution explained 53% of the variance in the data. In order to demonstrate the validity of combining the two components of PAM, we conducted an additional EFA with a forced two-component solution. This analysis revealed that all questions on PAM clustered as one component whereas all questions on GAA clustered as another (for item loadings, see Supplementary Table 2). This two-component solution explained 43% of the variance in the data. Our results demonstrated that PAM and GAA are distinct conceptual constructs, with PAM consisting of two correlated subcomponents.

Scanner task Participants in Cohort 1 performed a block-design fMRI experiment which consisted of one run of addition problems with four task conditions: (1) Complex addition problems, (2) Simple addition problems, (3) Number identification, and (4) Passive fixation. In the current study, we focused on the contrast between "Complex addition" and "Fixation" conditions to maximize brain responses for arithmetic processing (see Figure 1A). In the Complex addition task condition, participants were presented with an equation involving two addends and asked to indicate whether the answer shown was correct or incorrect (e.g., "5 + 2 = 7"). The first operand ranged from 2 to 9, and the second from 2 to 5. Double-digit addition problems, such as 5 + 5 = 10, were excluded. Answers were correct in 50% of the trials, and incorrect answers deviated by ±1 or ±2 from the correct sum. In the Passive fixation task, an asterisk symbol appeared at the center of the screen and participants were asked to focus their attention on it. Stimuli were presented in block fMRI design in order to maximize signal detection. In each run, there were four blocks of each condition (18 trials per condition) and the order of the blocks was randomized across participants. Each trial lasted 5.5 seconds with an inter-trial interval of 500 milliseconds followed by the stimuli displayed for 5 seconds. Trials were grouped together in sets of four to five trials. Further details of the experiment design can be found in Young et al.(Young, Wu, & Menon, 2012).

Whole-brain group analysis For the whole-brain group analysis in Cohort 1, contrast images of Complex addition relative to Fixation conditions were analyzed. Significant clusters were determined using a height threshold of $p < .01$ using a Monte Carlo determined cluster extent of 128 voxels in the whole-brain analysis after inclusive gray matter masking (Forman et al., 1995). We used PAM scores as a continuous variable to identify brain regions that showed increases and decreases in brain activation related to PAM during arithmetic problem solving.

Cross-validation and prediction analysis To investigate whether functional brain responses could reliably predict individual differences in PAM scores, we used a machine learning approach and conducted a balanced cross-validation analysis similar to previously published studies (Supekar, Iuculano, Chen, & Menon, 2015). First, we used averaged functional activation within the right hippocampal ROI as the independent variable and participants' PAM scores as the dependent variable. In order to provide evidence for the robust findings in both cohorts, we performed the analysis in both Cohort 1 and 2 using inclusive functionally-specified ROI masks. Data were then split into two folds such that the distributions of dependent and independent variables were balanced across the two folds. A linear regression model was built using data from one fold, leaving out the other fold as the test data. The PPM scores in the left-out fold were then predicted using this model. This procedure was repeated for the left-out fold, which would be the training data, and the Pearson correlation between predicted and actual PPM scores (r(pred, actual)) was computed and used as a measure of how well the independent variable predicted the dependent variable. The cross-validation was repeated 100 times to account for variance due to random split into training and test sets. The average r(pred, actual) was computed across the 100 repetitions. Finally, the statistical significance of model performance (i.e. r(pred, actual)) was assessed using nonparametric analysis. The empirical null distribution of r(pred, actual) was estimated by generating 1,000 surrogate datasets under the null hypothesis that there was no association between functional activation and PAM scores. Each surrogate dataset has the same size of the observed dataset and was generated by randomly permuting the participants' PAM scores. $r(\text{pred, actual})^{\text{null}}$ was computed using the actual non-permuted PPM scores and predicted scores using the same two fold balanced cross-validation procedure described previously. This procedure produces a null distribution of r(pred,actual) for the regression model. The statistical significance (p value) of the model was then determined by counting the number of $r(\text{pred, actual})^{\text{null}}$ greater than r(pred, actual) and then dividing that count by the number (1000) surrogate datasets.

Methods S2: Supplemental Information for Cohort 2

Subjects Participants were recruited from a wide range of schools in the San Francisco Bay Area using school mailings and postings at libraries and community groups. Participants were typically developing third grade children, aged 8-9 years. All participants were right-handed and had no history of psychiatric illness. The initial sample included 44 participants who completed an 8-week math training, as well as the same positive attitude assessment administered to the sample in the main text. Data from their initial visit, or pre-training, was used for the replication study. Four subjects were excluded for poor movement, 7 for low accuracy (< 60%) on the in-scanner addition task, and 5 for artifacts detected upon quality control inspection of brain images. The final sample consisted of 28 children (mean age = 8.70, SD = 0.48, 15 females) who had both complete functional imaging data and positive attitude assessment scores.

Materials and scanner task The in-scanner arithmetic verification task consisted of two runs of arithmetic problem solving during which the child indicated whether addition equations were correct (e.g., 3 + 4 = 7) or incorrect (e.g. 3 + 4 = 8). Problems were presented in a fast event-related fMRI design with 12 single-digit problems per run. In each run, problems were presented horizontally in green lettering on a black background. In 50% of the problems, the answers presented were correct (e.g., 2 + 4 = 6); in the remaining 50%, the answers presented deviated from the correct solution by ±1 or ±2 (e.g., 3 + 5 = 7). Arithmetic problems with 1 or 0 as operands were excluded. The larger operand was equally likely to appear in the first or second position. Each trial started with a fixation asterisk that lasted for 0.5 seconds. Then, the problem was presented for a maximum of 9.5 seconds, during which time the child could make the response. The participant used a response box and used their index or middle finger to indicate if the answer was correct or not. After the response, the problem disappeared from the screen and a black screen appeared until 9.5 seconds was reached. The task design also included a total of six 10-second rest periods, which occurred at jittered intervals during each run to achieve an optimal event-related fMRI design (Kao, Mandal, Lazar, & Stufken, 2009).

fMRI data processing For brain imaging data acquisition information, please see (Supekar et al., 2015). Data were preprocessed in SPM8 (http://www.fil.ion.ucl.ac.uk/spm/; Hall & Degenhardt, 2008). To allow for signal equilibration, the first five volumes were removed from the analysis. A linear shim correction was applied separately for each slice during reconstruction using a magnetic field map acquired automatically by the pulse sequence at the beginning of the scan (Lai & Glover, 1998). Images were realigned to correct for motion, corrected for errors in slice timing, and spatially transformed to standard stereotaxic space (based on the Montreal Neurological Institute coordinate system), resampled every 2 mm using sinc interpolation, and smoothed with a 6 mm full-width half-maximum Gaussian kernel to decrease spatial noise before statistical analysis. Translational movement in millimeters (x, y, z) and rotational motion in degrees (pitch, roll, yaw) were calculated based on the SPM8 parameters for motion correction of the functional images for each subject. To correct for deviant volumes resulting from spikes in movement, we used despiking procedures similar to those implemented in Analysis of Functional NeuroImages (Cox, 1996). Volumes with movement exceeding 0.5 voxels (1.562 mm) or spikes in global signal exceeding 5% were interpolated using adjacent scans.

ROI analysis We expected to find enhanced activation associated with positive attitude scores in the hippocampus. We used a one-tailed t-test with a threshold of p <.05 to identify clusters that positively correlated with PAM scores. The same small volume correction method was used to compute the threshold for hippocampal clusters. Due to the nature of event-related design, we were able to model the trials with correct and incorrect trials separately. Beta averages were extracted from a 2mm diameter ROI centered at the peak coordinate in the hippocampus. Correlations between PAM scores and hippocampal activation were calculated separately for correct and incorrect trials. Note that three children were 100% accurate in the addition task so the correlation between PAM scores and activation in incorrect trials had a sample size of 25.

For cortical regions, namely, right lingual gyrus, anterior fusiform gyrus/anterior temporal lobe, superior frontal gyrus/supplementary motor are, and cerebellum, an ROI with a 2 mm diameter centering peak MNI coordinates (see Table 2) from Cohort 1 was used to extract beta averages from the replication data. Pearson's correlation was then computed to identify any significant correlation between activation in these regions and positive attitude.

IV. Supplemental Results

Positive attitude towards math can be dissociated from attitude for other academic domains

We acquired comprehensive neuropsychological data and 5-point Likert scale questionnaires for positive attitude towards math (PAM) and general attitude towards academics (GAA) from 240 children ages 7-12 (Table 1 and Supplementary Table 1; details in Methods). Our first goal was to corroborate that PAM can be dissociated from GAA. Factor analysis revealed three factors (Supplementary Table 2 and Methods): one factor encompassing questions regarding GAA, and two factors encompassing strong interest and self-perceived ability in math, sub-components of PAM. The two sub-components of PAM were significantly correlated with each other, r (238) = 0.37, p <.001, 95% *C.I.* = [0.26, 0.48], but neither was correlated with GAA (both ps > .05). Based on the Fisher's Z test, the correlation between the two components of PAM was significantly higher than those with GAA (both ps < .01). These results provided proof of concept for PAM as a composite measure and supported findings from previous studies showing that positive attitude towards math can be dissociated from attitude towards general academics.

Positive attitude towards math is also associated with enhanced cortical responses

In Cohort 1, we found that activation in left anterior fusiform gyrus (aFG; peak coordinate: [-40, -8, -34]), right lingual gyrus (rLG, [10, -54, -6]), left superior frontal gyrus/supplementary motor area (SFG/SMA, [-10, 4, 64]) and dorsal cerebellum [10, -64, -28] were significantly correlated with PAM scores (Supplementary Figure 4). However, these findings were not replicated in Cohort 2. Furthermore, no brain regions showed task-related activation that was negatively correlated with PAM in either cohort.

Neural correlates of positive attitude towards math beyond memory retrieval

We explored other neural correlates of PAM when retrieval rates were controlled for in Cohort 1. We found similar activation patterns in the cerebellum and left superior frontal gyrus/supplementary motor area (Supplementary Figure 6), indicating that activations in the left anterior fusiform gyrus and right lingual gyrus were implicated in retrieval use during complex arithmetic problem solving. Interestingly, the right subcallosal cortex, extending into nucleus accumbens, was negatively correlated with PAM scores when retrieval rate was controlled (Supplementary Figure 6).

Replication of hippocampal activation for positive attitude and brain-behavioral SEM models

With the exception of the right hippocampus, we failed to replicate results in other brain regions including the left hippocampus, right lingual gyrus, anterior fusiform gyrus, superior frontal gyrus, and cerebellum in Cohort 2. Therefore, we failed to provide reproducible evidence for activation in these brain regions associated with positive attitude.

Brain-behavioral SEM models were also produced with data from Cohort 2. Beta Average values extracted from 2mm-diameter ROIs of the two right hippocampal clusters were combined for the right hippocampal activation in the SEM models. Although the mediating role of right hippocampal activation and memory retrieval for positive attitude on math achievement was replicated, the overall pattern was slightly different. Amongst all three models, the partial-retrieval-based model showed the lowest AIC (298.39) and BIC (315.71) values, compared to the non-retrieval-based (AIC= 303.78, BIC= 320.60) and full-retrieval-based (AIC= 308.17, BIC= 322.82) models. In the partial-retrieval-based model, PAM had a significant and direct effect on right hippocampal activation ($\beta = 0.60$, $z = 4.00$, $p < .001$), which consequently had a significant influence on retrieval rate ($\beta = 0.45$, $z = 2.30$, $p < .05$). Scores on Numerical Operations were directly influenced by both retrieval rate ($\beta = 0.33$, $z = 2.03$, $p < .05$) and positive attitude ($\beta = 0.75$, $z = 4.31$, $p < .001$); the latter was not observed in Cohort 1. This partial-retrieval-based model best fit the data ($\chi^2 = 1.90$, $p = .168$), although other fit parameters were not quite satisfactory (CFI = 0.97, TLI = 0.81, RMSEA = 0.18, SRMR = .070). Nevertheless, it was the best model amongst the three.

Power analysis of the fMRI and SEM results

In order to demonstrate the use of an adequate sample size to acquire significant results in fMRI and SEM analyses, we first used power analysis to determine the minimum number of participants necessary to detect significant fMRI activation with power greater than 0.7 at a significance level of 0.05. Results indicated that the minimum number of subjects required for Cohort 1 was N = 30 for and N = 27 for Cohort 2. Our fMRI sample sizes (N = 47 for Cohort 1, N =28 for Cohort 2) exceed the number of necessary subjects. In the revised manuscript, we report Cohen's *d* for all correlation analyses. We demonstrate sufficient power in our key fMRI analyses with medium to large effect size (Cohen's *d* = 0.6 to 0.9).

Using the *R* package "*simsem*" (Pornprasertmanit, Miller, & Schoemann, 2013), we then ran a Monte Carlo simulation with 10,000 repetitions of the full-hippocampal-mediated model given the covariance-variance structure of the four measures used in the SEM model (i.e., PAM, right hippocampal activation, correct retrieval rate, and Numerical Operations). Based on the simulation, Cohort 1 (n=47) had decent power (0.6~0.9) to detect the non-zero effect (see Table S9). The same was observed in Cohort 2 (n=28) as well. Given the smaller sample size, some of the effects in Cohort 2 had lower power compared to those in Cohort 1. However, both cohorts demonstrated medium power (0.56~0.60) to detect a non-zero mediation path from PAM to Retrieval via Hippocampal activation.

Alternative SEM models

To further validate the SEM models and confirm the directionality of the effect from positive attitude to memory retrieval, we created two alternative models: (1) the alternative fully-mediated model and (2) the alternative partially-mediated model. These models differ from the original models in that they include a direct effect of retrieval on PAM. In both alternative models, we observed a non-significant influence of memory retrieval on PAM, $Z = 0.98$, $p > .250$. Additionally, both alternative models showed higher AIC and BIC values, indicating a less satisfactory model fit (see Table S10). Thus, the fully-mediated model remains the best-fit model, and evidence contradicting the directionality from PAM to retrieval is weak.

V. Supplemental References

Cox, R. W. (1996). AFNI: Software for Analysis and Visualization of Functional Magnetic Resonance Neuroimages. *Computers and Biomedical Research*, *29*, 162–173. http://doi.org/10.1006/cbmr.1996.0014

Forman, S. D., Cohen, J. D., Fitzgerald, M., Eddy, W. F., Mintun, M. A., & Noll, D. C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, *33*(5), 636–47. http://doi.org/10.1002/mrm.1910330508

Hall, W., & Degenhardt, L. (2008). SPM8 - Statistical Parametric Mapping. Retrieved from http://www.fil.ion.ucl.ac.uk/spm/software/spm8/

Kao, M.-H., Mandal, A., Lazar, N., & Stufken, J. (2009). Multi-objective optimal experimental designs for event-related fMRI studies. *NeuroImage*, *44*(3), 849–856. http://doi.org/10.1016/j.neuroimage.2008.09.025

Lai, S., & Glover, G. H. (1998). Three-dimensional spiral fMRI technique: A comparison with 2D spiral acquisition. *Magnetic Resonance in Medicine*, *39*(1), 68–78. http://doi.org/10.1002/mrm.1910390112

Pornprasertmanit, S., Miller, P., & Schoemann, A. R. (2013). package simsem: SIMulated structural 459 equation modeling (version 0.5-3) 2013.

R Core Team. (2015). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria*. http://doi.org/ISBN 3-900051-07-0

Raiche, G., & Magis, D. (2010). nFactors: Parallel Analysis and Non-Graphical Solutions to the Cattell Scree Test. R Package Version 2.3. 3. Available online at: ht tp://cran. r-project. org/web/packages/nFactors/index. html.

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, *24*(2), 282–292. http://doi.org/10.1037/a0025697

Supekar, K., Iuculano, T., Chen, L., & Menon, V. (2015). Remediation of Childhood Math Anxiety and Associated Neural Circuits through Cognitive Tutoring. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *35*(36), 12574–83. http://doi.org/10.1523/JNEUROSCI.0786-15.2015

Young, C. B., Wu, S. S., & Menon, V. (2012). The Neurodevelopmental Basis of Math Anxiety. *Psychological Science*, *23*(5), 492–501. http://doi.org/10.1177/0956797611429134