Mathematics Anxiety: Separating the math from the anxiety

Ian M. Lyons[a] and Sian L. Beilock[a]

[a]University of Chicago, Department of Psychology

Address correspondence to Sian L. Beilock, Department of Psychology, 5848 South University Avenue, The University of Chicago, Chicago, Illinois 60637. Send email to beilock@uchicago.edu.

ABSTRACT

Anxiety about math is tied to low math grades and standardized test scores, yet not all math-anxious individuals perform equally poorly in math. We used fMRI to separate neural activity during the *anticipation* of doing math from activity during math *performance* itself. For higher (but not lower) math anxious individuals, increased activity in fronto-parietal regions when simply anticipating doing math mitigated math-specific performance deficits. This network included bilateral inferior frontal junction (IFJ), a region involved in cognitive control and reappraisal of negative emotional responses. Furthermore, the relation between fronto-parietal *anticipatory* activity and highly math anxious individuals' math-deficits was fully mediated (or accounted for) by activity in caudate, nucleus accumbens, and hippocampus *during* math performance. These subcortical regions are important for coordinating task demands and motivational factors during skill execution. Individual differences in how math-anxious individuals recruit cognitive control resources prior to doing math and motivational resources during math performance predict the extent of their math-deficits. This work suggests that educational interventions emphasizing control of negative emotional responses to math stimuli (rather than merely additional math training) will be most effective in revealing a population of mathematically competent individuals, who might otherwise go undiscovered.

INTRODUCTION

Although basic math skills are important for everyday life, many people report feeling anxious when faced with the prospect of doing math. *Mathematics anxiety* is characterized by feelings of tension, apprehension, and fear about performing math and is associated with delayed acquisition of core math and number concepts and poor math competence (Richardson & Suinn, 1972). Math-anxiety is clearly an impediment to math achievement (Ashcraft & Ridley, 2005; National Mathematics Advisory Panel, 2008). However, not all people high in math-anxiety perform equally poorly in math. This variation in math performance among the highly math-anxious provides an opportunity (1) to understand the reasons for the negative relation typically seen between math-anxiety and math competence, and (2) to shed light on how this relation is ameliorated.

In the current work, highly math-anxious individuals (HMAs) and a group of low math-anxious controls (LMAs) were identified via a common self-report measure of math-anxiety (SMARS; Alexander & Martray, 1989). All participants performed a mental arithmetic task and difficulty matched word-verification task during fMRI acquisition. Crucially, before each set of problems, individuals were presented with a cue (a simple colored shape) identifying the nature of the upcoming task (math or word). This paradigm, allowed us to separate neural activity underlying the *anticipation* of doing math from that of math performance itself.

As a preview, we found that HMAs' overall performance was characterized by a math-specific deficit: HMAs showed significantly poorer math performance relative to a non-math, difficulty-matched task (LMAs performed the same on both tasks – Figure 1a). Importantly, some HMAs showed more of a math-deficit than others and these math-deficits were not related to level of reported math anxiety. So, why might this variation in math-deficits among HMAs come about?

One possibility is that some HMAs are better at math than others. Such a finding would fall in-line with assertions that math-anxiety is simply a proxy for poor math competence (Fennema, 1989);

that is, people with the lowest math competence are most anxious about their lack of math proficiency. If so, then activity in neural regions important for numerical calculation (e.g., the intraparietal sulci, IPS; Simon et al., 2002) should relate to the extent of HMAs' math-deficits. Moreover, this activity, which supports the successful retrieval and implementation of mathematical computations, should occur during math performance itself.

Another possibility is that the anxiety felt by highly math-anxious individuals changes how HMAs approach math, and this in turn affects math competence. If so, then neural activity related to HMAs' math-deficits may be apparent before the math task even begins. For example, the extent to which HMAs recruit regions related to attentional control and the reappraisal of negative emotions in response to a cue indicating they are about to do math (e.g., regions such as mid-posterior DLPFC; Ochsner et al., 2004; Bishop, 2007) might predict their math-deficits. The more activity in these regions, the better HMAs may be at controlling their negative response to the upcoming math task. This, in turn, may allow for the coordination of the appropriate neural resources required to successfully perform the math task.

In sum, for the first time with math-anxious individuals, we were able to separate the neural correlates underlying the *anticipation* of doing math from those of math performance itself. Our findings shed light on the factors driving the comorbidity of math anxiety and poor math competence, knowledge needed to develop the appropriate educational interventions to ameliorate this relation.

MATERIALS & METHODS

Subjects

Subjects were 32 right handed University of Chicago students (mean age=20.47 years, range=18-25). Math-anxiety groups were determined using the Short Math-anxiety Rating Scale (SMARS). Across all prescreened subjects (N=108), math-anxiety ratings were in keeping with published norms (our sample: mean=32.11, SD=15.39; Alexander & Martray, 1989: mean=36, SD=16). To maximize group differences, upper and lower quintiles of the prescreening group* were selected to participate in the fMRI portion of the study: HMA mean=49.56, SD=6.98; LMA mean=15.00, SD=5.78. Of the 32 subjects thus scanned, four were removed from the dataset; two were removed due to major scanner artifact, and two were removed because of excess motion. The remaining 28 subjects were divided into high (HMA: 8 female) and low (LMA: 7 female) math-anxious groups. Note that HMAs and LMAs did not differ in terms of either trait-anxiety (Spielberger et al., 1970) (p=.5438) or working memory capacity (Unsworth et al., 2005) (p=.9134).

fMRI Task Procedure

Subjects performed two different tasks (math and word task) in blocked fashion. Two levels of difficulty were included for each task: hard and easy. Note, however, that the Group×Task interaction did not obtain for easy problems [F<1] (see Supplementary Information for further details), a result in keeping with previous work showing that HMAs' performance is primarily impacted on math problems with high working memory demands (Ashcraft & Krause, 2007). Thus, brain~behavior analyses were limited to hard problems. Critically, before each block of trials, a cue was provided that indicated which type of task was about to follow (note that the cue did not identify task difficulty). Cues were either a yellow circle or a blue square. Which cue indicated which task was randomized across subjects. Cues predicted the corresponding task with 100% accuracy.

To increase their functional saliency, cues always preceded blocks of 4 trials each. For example, if a flashing square indicated math trials, this cue was succeeded by a block of 4 math trials. The 4 trials in a block were separated by a fixed inter-stimulus-interval of 1500msec. Each trial had a 5.5 second cut-off. This cut-off was well above average task RTs (see Supplementary Information). Cues were presented for 1500msec. Fixation time between cue offset and trial-block onset was jittered

* 27 of the original 108 subjects were deemed unsuitable for scanning due to handedness, outlier responses on one of the other self-report measures, neurological abnormalities, or safety concerns.

between 2500 and 6500msec. This was done to separate neural signals generated by the cues and signals generated by performing the actual tasks. Fixation time between trial-block offset and onset of the cue for the following block was 18sec to allow for HRF resolution between blocks. This rest period was used to model baseline activity (this was done for each group separately). Participants completed 32 blocks of each task type over the course of 8 functional runs, with 4 blocks of each task presented in each run.

Math Task. Subjects verified whether arithmetic problems had been correctly solved. All problems were of the form $(a*b)-c = d$; where $a \neq b$, $c>0$, $d>0$; for hard math problems, $5 \leq a \leq 9$, $5 \leq b \leq 9$ ($a*b \geq 30$), and $15 \leq c \leq 19$; for foil problems, $d \pm 2$. In addition, subtracting c from $a*b$ always involved a borrow operation. Subjects pressed a key with either their left or right index finger to indicate whether the problem had been solved correctly (which hand indicated which response was randomized across subjects; response randomization was independent of cue randomization).

Word Task. Subjects verified whether a word, if reversed, spelled an actual English word (all subjects were fluent English speakers). For half of the words, two internally adjacent letters were switched such that reversing the whole letter string would result in a non-word. For example, if subjects saw the string `tneimrepxe`, reversing this string would generate `expermient`. Because the `i` and the `m` have been switched, this is not an English word, and subjects should respond 'no'. For hard word problems, strings were always seven letters in length. Subjects pressed a key with either their left or right index finger to declare whether the string, when reversed, produced a real word or not. To avoid subject confusion, the hand that indicated 'not correct' for the math problems also was used to indicate 'not correct' for the word problems.

We chose this task because we were primarily concerned with creating a control task that was decidedly non-mathematical in nature but similar to the math-task in terms of response and general task-set properties (both were two-alternative, forced-choice verification-tasks). Crucially, the control-task needed be as difficult as the hard math-task – from at least one objective source (behavioral data) – in the low math-anxious (LMA) group. We also ensured that performance did not significantly differ across the HMA and LMA groups on this task (see behavioral results below).

After completion of all scans, subjects were verbally probed to see (1) if they actually knew which cue prompted which task: all subjects responded correctly to this question; and (2) if they had inferred the true focus of the study (math-anxiety): none reported having done so.

fMRI Data Acquisition and Analysis

Acquisition and Preprocessing

MRI data were acquired using a 3 Tesla Philips Achieva scanner with an 8-channel Philips Sense head-coil. A $T2^*$-weighted echo-planar imaging sequence was used to acquire functional images covering the whole brain (32 axial slices) with a repetition time (TR) of 2000ms and an echo time of 25ms (ascending acquisition; FOV: 240x240x127.5mm; 80x80x32 matrix; flip angle: 80°). In-plane resolution was 3x3mm and the slice thickness was 3.5mm (0.5mm skip). Signal from the orbital frontal cortex (OFC) and surrounding tissue was recovered using additional volume shimming with a box of 60x60x60mm centered on the OFC area. This method utilizes multiple 'pencil beam' acquisitions to compute shim values, using a Pencil Beam-Volume (PBV) shimming algorithm provided by Philips. Whole-brain high-resolution anatomical images were acquired in the axial plane (300 slices; slice thickness: 1.2mm, -.6mm gap; x-y dimensions: 1.04x1.04; FOV: 250x250x180mm, 240x240x300 matrix) with a standard Philips T1-weighted SENSE-Ref sequence.

All preprocessing steps and whole-brain data analyses were conducted using BrainVoyager QX (version 1.10.4, Brain Innovation, The Netherlands). Functional images were first slice-time corrected and then motion corrected using sinc interpolation. A high-pass GLM (Fourier basis set) temporal filter removed fluctuations <2 cycles, which also removed linear temporal drift. Each functional run was then manually aligned to the subject's 3D anatomical image, both of which were then transformed into Talairach space. Resulting volumetric time-series files were then spatially smoothed with a 6mm full-width at half-maximum Gaussian kernel.

Data were next submitted to a random effects GLM (Friston et al., 1994) with 6 main predictors of interest: math-cue, word-cue, hard and easy math task blocks, and hard and easy word task blocks. 7 predictors of no interest (motion parameters and errant button presses) were included in the model. In each voxel and for each subject, parameter estimates (hereafter βs) for each subject and each condition were thus generated. All second level analyses (see below) were conducted using these voxelwise βs.

Second Level Analyses

For all whole-brain analyses below, unless otherwise stated, statistical maps were thresholded at $p<.005$, and subsequently cluster-level corrected for multiple comparisons using a Monte-Carlo simulation procedure (Forman et al., 1995) with a family-wise false-positive rate $\alpha=.05$. With respect to ROI analyses, for each subject and predictor, ROI-level βs were determined by averaging βs from all voxels comprising the ROI volume in question (for that subject and that predictor). Once extracted, ROI βs were submitted for analysis in Matlab.

In light of recent debate (Vul et al,. 2009; Kriegeskorte et al., 2009) regarding the reporting of correlational values (upon which some of our analyses rest), we believe we would be remiss if we did not mention that *r*-values, like any other summary statistic, carry a certain degree of imprecision, which is exacerbated in cases involving relatively few degrees of freedom. Therefore, in all tables where we report correlation estimates expressed either in terms of standard deviations (*r*-values) or arbitrary units (β-values), we provide standard errors of that estimate as well (denoted in parentheses beside the actual estimate). Figures depicting fitted regression lines also contain 95% confidence-interval ranges for these fitted estimates.

RESULTS

Behavioral Results: Math-anxiety and individual performance differences. Our primary behavioral measure of interest was task error-rates (proportion wrong), which coincides with the fact that performance accuracy is what students are graded on in a majority of math performance situations[†].

As mentioned above (also Figure 1a), HMAs produced significantly more errors than LMAs on the math task [$t(26)=3.50$, $p=.0017$], but the two groups did not differ on the difficulty-matched word task [$t(26)= -0.02$, $p=.9923$]. Furthermore, HMAs performed more poorly on the math relative to the word task [$t(13)=3.80$, $p=.0022$]; LMAs did not [$t(13)=0.61$, $p=.5535$]. [Group×Task interaction: $F(1,26)=16.15$, $p=.0004$, $\eta^2=.383$.]

Importantly, some HMAs showed a larger math-deficit (math errors minus word errors) than others. Note, however, that a self-report measure of math-anxiety (Alexander & Martray, 1989) was not correlated with math-deficits in the HMA group ($p=.4634$).

Analysis 1: <u>Cue</u>-Activity ~ Math-Deficit Correlations. To determine why some HMAs showed larger math performance-deficits than others, we began by relating math-deficits (difference between math and word error-rates) to neural activity evoked when faced with the prospect of doing math. For HMAs, this math-deficit was a distribution of positive values; for LMAs it centered on zero. The LMA group served mainly as a control group in demonstrating that our effects were specific to HMAs, as LMA's had no math-deficit to explain. Thus, we tested whether the observed correlation for HMAs was specific to that group – i.e., whether the slope of predicted math-deficits from cue-activity depended on math-anxiety group. We regressed behavioral math-deficits on neural activity during math-cue presentation (controlling for word-cue activity).

No significant regions were found for LMAs. For HMAs, as math-deficits decreased (i.e., less of a difference between math and word error rates), differences in cue activity (math–word) increased in several regions. These regions were bilateral inferior frontal junction (IFJ), bilateral inferior parietal

[†] Note that behavioral results were similar for response-times (see Supplementary Information); importantly, there were no (speed-accuracy) trade-offs between response-times and error-rates (all ps≥.374).

lobe (IPL; this cluster spanned the junction between angular and supramarginal gyri), and left anterior inferior frontal gyrus (IFGa). Table 1 (top) summarizes region details; regions and regression diagnostics are shown in Figure 2. Note that the relation between cue-activity and math-deficits remained highly significant (*p*s<.005) even when controlling for ratings of math-anxiety (SMARS). Thus, for HMAs, it is not necessarily the level of one's self-reported math anxiety *per se* that predicts one's math-deficit, rather it is one's ability to call upon fronto-parietal regions before the math task has even begun – regions hypothesized to be involved in both cognitive control (Brass et al., 2005; Derrfuss et al., 2005; 2009) and regulating negative emotional responses (Ochsner et al., 2004; Bishop, 2007). A significant Group×Slope interaction obtained in all regions (all *p*s≤.0330), indicating that the relation between math-deficits and fronto-parietal cue-activity was specific to HMAs.

Because math-deficits were regressed on the difference between math-cue and word-cue activity (βs), it is important to determine whether math- or word-cue activity drove the observed effect in each region. Table 2 ('Correlation Splits' columns) shows HMA correlation results for each cue β separately (i.e., the correlation between raw subject βs and performance deficits)[‡]. The correlations in IFJ and IPL regions were driven primarily by math-cue activity. By contrast the correlation in left IFGa was driven primarily by word-cue activity. This latter result is not surprising given that this region is canonically associated with semantic language processing (Vigneau et al., 2006).

To help understand these results better, we divided the HMA group into two subgroups using a median split based on the difference between math-cue and word-cue βs in each of these regions[§].

[‡] LMAs were not considered as there were no significant cue~performance relations identified for this group.

[§] Note that this clearly constitutes a form of 'double-dipping' (Kriegeskorte et al., 2009) because these regions were localized based on the relation between math-deficits and the difference between math- and word-cue activity in HMAs. However, this section is meant merely as a means thinking about the data in more intuitive manner, rather than as a definitive analysis unto itself. For simplicity, IFJ and IPL splits were based on average ranking across hemispheres.

It is worth noting that mean activity for HMAs in these regions was at or near baseline in all regions (Table 2: HMA Raw βs: Math βs). This means that about half of HMAs should have shown activation differences above baseline and the remaining half somewhere below baseline (recall that baseline activity was calculated for each group separately based on the 18sec fixation period between blocks). Relative to the lower-half of HMAs, the upper-half tended to show greater math-cue activity [**IFJ**: upper-half: M=.268, lower-half: M= -.259, $t(12)=3.01, p=.0109$; **IPL**: upper-half: M=.155, lower-half: M=-.269, $t(12)=2.63, p=.0218$; **LIFGa**: upper-half: M=.176, lower-half: M=-.201, $t(12)=1.97, p=.0724$] but not word-cue activity [**IFJ**: upper-half: M=-.157, lower-half: M=-.039, $t(12)=-0.74, p=.4758$; **IPL**: upper-half: M=.069, lower-half: M=.068, $t(12)=0.01, p=.9906$; **LIFGa**: upper-half: M=-.051, lower-half: M=.116, $t(12)=-0.66, p=.5194$]. In terms of performance, using a median split based on the average rank across all 5 regions, the upper-half of HMAs showed 17.3% (hard)math-errors, while the lower-half showed 32.1% errors [$t(12)=2.83, p=.0152$]. The two sub-groups did not differ with respect to (hard)word-errors [upper-half: 14.4%, lower half: 10.3%, $t(12)=0.86, p=.4091$]. Thus, HMAs who activated this frontoparietal network more for the math-cue (but not the word-cue) also showed nearly complete elimination of their math-deficits (2.9%, on average), whereas those who did not show this activation had much greater math-deficits (21.8%, on average).

In sum, increased fronto-parietal activity in response to the prospect of doing math predicted reduction in the magnitude of math-deficits in a manner that also depended on whether one was high in math-anxiety. Presumably, then, there should be some neural region(s) whose activity during actual math-performance directly mediates this relationship.

Analysis 2: <u>Task</u>-Activity ~ Math-Deficit Correlations. We next investigated task-activity during math *performance* itself that might mediate (or account for) the relation between fronto-parietal cue-activity and HMAs' math-deficits. This analysis was conducted in the same manner as Analysis 1, with the exception that math-deficits were regressed on the difference between math and word *task* (rather than *cue*) activity. Table 1 (bottom) summarizes region details; Figure 3 shows regions and regression diagnostics.

Again, no significant regions were found for LMAs. For HMAs, two regions showed a negative relation with math-deficits: right dorsomedial caudate (overlapping with right nucleus accumbens) and left hippocampus. As the difference between math and word task-activity increased (math–word), HMAs' math-deficit decreased. Interestingly, hippocampus and dorsomedial caudate are highly interconnected and are thought to form a tightly coupled network that plays a role in flexible, on-line processing (White, 2009). Significant Group×Slope interactions indicated these relations were specific to HMAs (Table 1). Furthermore, as seen in Table 2, the correlations in both regions were driven primarily by a negative relation between math task-activity and math-deficits: as math task-activity increased, math-deficits decreased.

As with Analysis 1, to help understand these results better, we divided the HMA group into two subgroups, this time using a median split based on the difference between math-task and word-task βs in right caudate and left hippocampus. Relative to the lower-half of HMAs, the upper-half tended to show greater math-task activity [**caudate**: upper-half: M=.228, lower-half: M= -.144, $t(12)$=2.95,p=.0121; **hippocampus**: upper-half: M=.275, lower-half: M=-.018, $t(12)$=1.55,p=.1480] but not word-task activity [**caudate**: upper-half: M=-.008, lower-half: M=-.004, $t(12)$=-0.04,p=.9701; **hippocampus**: upper-half: M=.158, lower-half: M=.196, $t(12)$=-0.26,p=.7976]. In terms of performance, using a median split based on the average rank across the two regions, the upper-half of HMAs showed 18.0% (hard)math-errors, while the lower-half showed 31.4% errors [$t(12)$=2.41,p=.0327]. The two sub-groups did not differ with respect to (hard)word-errors [upper-half: 13.4%, lower half: 10.9%, $t(12)$=0.61,p=.5528]. Thus, HMAs who activated this network more for the math-task (but not the word-task) showed nearly complete elimination of their math-deficits

(4.6%, on average), whereas those who did not show this activation difference had much greater math-deficits (20.5%, on average).

Analysis 3: Mediation Framework. In a mediation analysis, one asks whether the direct effect from an independent variable (IV) to a dependent variable (DV) can be accounted for by the indirect influence of a mediating variable (Figure 4). Significance estimates for indirect effects were obtained via the bootstrapping method described in Preacher and Hayes (2008). We treated *cue*-activity (math minus word) as our IV (note that cue-activity temporally preceded both task-activity and behavioral performance), *task*-activity as our mediating factor(s) (math minus word), and math-deficits as the DV we were interested in explaining. In sum, we asked whether task-activity (in Analysis 2 regions) mediated the cue-activity~math-deficit relation reported in Analysis 1.

To assess the criterion that the IV is related to the mediator(s), we first tested whether cue-activity in the regions identified in Analysis 1 correlated with task-activity in the regions identified in Analysis 2. Right caudate task-activity was positively correlated with cue-activity in all five fronto-parietal regions identified in Analysis 1 (ps≤.0058). All effects were specific to HMAs (ps≤.0337). Left hippocampus task-activity was also positively correlated with cue-activity in all five fronto-parietal regions identified in Analysis 1 (ps≤.0037). These correlations were specific to HMAs for all but left IFJ (p=.2468; all other ps≤.0156).

Both right caudate and left hippocampus were included as simultaneous mediators to examine their combined mediating effect[**]. We treated each of the five cue regions (left and right IFJ, IPL and left IFGa) as separate IVs. In addition, we constructed a model that assumed the net signal from all 5 regions served as a unified (averaged) signal. This was because, for HMAs, cue activity (math-cue – word-cue) was highly correlated between these regions (mean r = .809; range: .745 to .898). In this

[**] Note that task activity in each region alone provided a significant mediating influence; however, since these two regions were highly correlated with one another in HMAs (r=.919), for simplicity we treat them here as a single influence (see Figure 4 for details).

way, 6 separate models were run: one for each of the 5 cue regions and one for the cue network average. The combined mediating effect of right caudate and left hippocampus was significant in all 6 models (*p*s≤.0206). In sum, hippocampus and caudate/NAc task-activity mediated the fronto-parietal cue-activity~math-deficit relation seen in Figure 2. Table 3 summarizes the mediation results.

DISCUSSION

Our central aim was to identify neural areas – during both the anticipation and performance of math – that predict variation in the math-deficits exhibited by highly math-anxious individuals (HMAs). In doing so, we reveal (1) neural evidence that the negative relation typically seen between math-anxiety and math competence arises even before math performance begins, and (2) how some HMAs may overcome this math-specific deficit.

Math-deficits were predicted by cue-activity in a network of inferior fronto-parietal regions (IFJ, IPL, left IFGa), and in a manner statistically specific to HMAs – a relation driven primarily by math-cue activity in the IFJ and IPL regions (Tables 1 and 2). One possibility is that the reduced math-deficits exhibited by some HMAs are the result of these individuals' ramping up cognitive control resources when anticipating math in a manner that allows them to change the way they approach performing the upcoming math task.

The bilateral IFJ regions in particular correspond to an area of the cortex associated with high-level cognitive control processes such as task- or set-shifting and inhibition of prepotent responses (Brass et al., 2005; Derrfuss et al., 2005; 2009). A prominent theory regarding the negative impact of anxiety on cognitive task performance (attentional control theory) (Eyesenck et al., 2007) suggests that anxiety compromises performance via limiting the successful operation of attention-shifting and inhibitory processes. For instance, anxious individuals, when in the context of emotion-inducing stimuli, tend to exhibit poorer control of saccades (Ansari et al., 2008; Wieser et al., 2009), poorer task-switching ability during mental-arithmetic performance (Derakshan et al., 2009), and poorer performance in an emotional Stroop paradigm (Reinholdt-Dunne et al., 2009). This evidence is also consistent with the suggestion that math-anxiety involves a reduction in control-related working-memory capacity (Hopko et al., 1998; Ashcraft & Krause, 2007; Beilock, 2008). Our results suggest that, in the context of math-anxiety, some HMAs are able to overcome such attentional deficits by ramping up these resources before the math itself begins – a process which may also allow them to reappraise their approach to the upcoming math task as well.

Consistent with this view, in a meta-analysis, Bishop et al. (2007) identified a cluster of activations centered on bilateral IFJ associated with interpretation of potentially threat-related stimuli. Similarly, Ochsner et al. (2004) found that regulation of negative emotions via reappraisal was related to increased activity in bilateral DLPFC (note that in many studies cited, activations overlapping with the IFJ region reported here are sometimes labeled 'DLPFC'). Overlap between activations seen in Ochsner et al. and the IFJ regions in the current work was observed in the Ochsner et al. condition where participants were explicitly instructed to *reduce* the interpreted negative emotional content of visual images. Thus, it may be that HMAs who most successfully reappraised their negative emotional response to the prospect of doing math are most successful at reducing math-deficits. Importantly, according to this interpretation, such a relationship should not be seen for LMAs because they do not have a negative emotional response in anticipation of math that requires reinterpreting – exactly what we found.

In addition, it is worth noting that cue-activity was not related to math-deficits in regions typically associated with anxiety responses [e.g., amygdala (LeDoux, 2008), hypothalamus (Dedovic et al., 2009), insula (Domschke et al., 2010)]. Moreover, task-activity in regions associated with arithmetic calculation did not predict math-deficits in HMAs either [e.g., the IPS (Simon et al., 2002)]. Thus, our data are most consistent with the notion that HMAs' math-deficits are determined primarily by how well they respond to and perhaps reinterpret their anxiety response, rather than by the magnitude of those anxiety responses per se, or their math skills.

Cue-activity was positively correlated with task-activity in two subcortical regions: right caudate nucleus and left hippocampus – again, in a manner again specific to HMAs. Task-activity in these regions was in turn correlated with performance, such that HMAs who showed relatively greater math- than word-task activity showed the smallest math-deficits. In addition, this task-activity fully mediated the relation between cue-activity (in regions noted above) and math-deficits in HMAs. Dorsomedial caudate and the hippocampus (especially the fimbria-fornix substructure) are highly interconnected regions and have been shown to cooperate in flexible stimulus-response learning in rats (White, 2009). Interestingly, right caudate and left hippocampus have been found to be functionally interconnected in humans in the context of active episodic and semantic memory retrieval (Burianova & Grady, 2007) – which is highly consistent with the strong degree of functional connectivity found between HMA right caudate and left hippocampus in our data. Left mid-posterior hippocampus in particular is thought to be central to maintaining information in one's episodic working memory buffer (Rudner et al., 2007; Berlingeri et al., 2008).Together, these regions are thought to play a central role in integrating higher-level cognitive control of goal-driven actions (i.e., through selection of appropriate action schemas) with motivational and reward context (Grahn et al., 2008).

Anatomically the caudate head receives primarily glutamatergic inputs from prefrontal areas such as DLPFC and anterior cingulate cortex (ACC), as well as dopaminergic inputs arising from nearby nucleus accumbens and ventral midbrain regions (Utter & Basso, 2008). Dopamine levels in dorsal caudate have also been shown to correlate with delay activity (i.e., while holding information in mind) in left IFJ during a Sternberg delayed-recognition working-memory task (Landau et al., 2009). One interpretation of the mediating role observed for caudate in the current data set is that

HMAs who show the smallest math-deficits do so because they are able to dynamically reorganize their approach to doing math. If HMAs' default response is to avoid math or just wait for it to be over, a more approach-oriented response is likely needed to successfully engage in the actual cognitive requirements of math performance (Markman et al., 2006). Consistent with this interpretation, the caudate region we identified also overlaps with nucleus accumbens (Figure 3). Nucleus accumbens is believed to be central for motivating behavior (both appetitive and aversive) and regulating effortful functioning (Salamone, 1994; Nicola, 2005; Salamone et al., 2007).

To summarize, we demonstrate that the mechanisms associated with the reduction – and even elimination – of math-specific performance deficits in HMAs are initiated before actual math processing occurs. However, it is not variation in HMAs' math-anxiety per se that best predicts the extent of their math-specific-deficits. Rather, our data indicate that the extent of HMAs' math-deficits is associated with the ability to ramp up cognitive control resources in response to *the prospect* of doing math, which leads to a reorganization of task priorities, due at least in part to motivational factors. To our knowledge, this work serves as the first evidence from cognitive neuroscience in support of the idea that education interventions which emphasize the *control* of negative emotional responses to math stimuli may reveal a population of potentially numerically competent individuals (see McClandiss, 2010, in support of an educational neuroscience approach more generally). In particular, the fact that PFC regions typically associated with cognitive control and working-memory processes (Kane & Engle, 2002; Brass et al., 2005) were found to be related to performance during the cue but not the actual task is highly consistent with the process model of emotion regulation proposed by Gross (1998; 2002). In this model, emotional control processes that act early on the arousal of negative affective responses (e.g., reappraisal) are more effective at mitigating these responses and limiting concomitant performance decrements than explicit suppression of these responses later in the affective process (e.g., during online performance).

Interestingly, recent evidence has shown that a similar approach highlighting the reappraisal of negative emotional reactions has proven effective, for example, in allaying the negative impact on math-performance of the fear of confirming negative stereotypes about one's academic abilities (i.e., stereotype threat; Johns et al., 2008).

The current findings therefore suggest that best educational practices for enhancing math competency in HMAs is not to generate costly math courses specifically for the highly math-anxious (Gresham, 2007); nor is the best method likely to be one that focuses solely on eliminating one's initial anxiety response (for a review of these and other approaches, see especially Hembree, 1990). Instead, classroom practices that help students learn how to marshal cognitive control resources and effectively check one's math-related anxiety response once it occurs – but before it has a chance to reduce actual math performance – will likely be the most successful avenue for reducing anxiety-related math-deficits.

ACKNOWLEDGMENTS

Research supported by NSF CAREER DRL-0746970 and the NSF Spatial Intelligence Learning Center to Sian Beilock.

REFERENCES

Alexander L, Martray C (1989) The development of an abbreviated version of the Mathematics Anxiety Rating Scale. Meas Eval Couns Dev, 22:143-150.

Ansari TL, Derakshan N, Richards A (2008) Effects of anxiety on task switching: evidence from the mixed antisaccade task. Cogn Affect Behav Neurosci, 8:229-38.

Ashcraft MH, Krause JA (2007). Working memory, math-performance, and math-anxiety. Psychon Bull Rev, 14:143-8.

Ashcraft MH, Ridley KS (2005) Math-anxiety and its cognitive consequences: A tutorial review. In Campbell JID (Ed.):Handbook of Mathematical Cognition (315-327) New York: Psychology Press.

Beilock SL (2008) Math performance in stressful situations. Curr Dir Psychol Sci, 17:339-343.

Berlingeri M, Bottini G, Basilico S, Silani G, Zanardi G, Sberna M, Colombo N, Sterzi R, Scialfa G, Paulesu E (2008). Anatomy of the episodic buffer: a voxel-based morphometry study in patients with dementia. Behav Neurol, 19:29-34.

Bishop SJ (2007) Neurocognitive mechanisms of anxiety: an integrative account. Trends Cogn Sci, 11:307-16.

Brass M, Derrfuss J, Forstmann B, von Cramon DY (2005) The role of the inferior frontal junction area in cognitive control. Trends Cogn Sci, 9:314-6.

Burianova H, Grady CL (2007). Common and unique neural activations in autobiographical, episodic, and semantic retrieval. J Cogn Neurosci, 19:1520-34.

Dedovic K, Duchesne A, Andrews J, Engert V, Pruessner JC (2009). The brain and the stress axis: the neural correlates of cortisol regulation in response to stress. Neuroimage, 47:864-71.

Derakshan N, Smyth S, Eysenck MW (2009) Effects of state anxiety on performance using a task-switching paradigm: an investigation of attentional control theory. Psychon B Rev, 16:1112-7.

Derrfuss J, Brass M, Neumann J, von Cramon DY (2005) Involvement of the inferior frontal junction in cognitive control: meta-analyses of switching and Stroop studies. Hum Brain Mapp, 25:22–34.

Derrfuss J, Brass M, von Cramon DY, Lohmann G, Amunts K (2009) Neural activations at the junction of the inferior frontal sulcus and the inferior precentral sulcus: interindividual variability, reliability, and association with sulcal morphology. Hum Brain Mapp, 30:299-311.

Domschke K, Stevens S, Pfleiderer B, Gerlach AL (2010) Interoceptive sensitivity in anxiety and anxiety disorders: an overview and integration of neurobiological findings. Clin Psychol Rev, 301-11.

Eysenck MW, Derakshan N, Santos R, Calvo MG (2007) Anxiety and cognitive performance: attentional control theory. Emotion, 7:336-53.

Fennema, E. (1989). The study of affect and mathematics: A proposed generic model for research. In D. B. McLeod & V. M. Adams (Eds.), Affect and mathematical problem solving: A new perspective (pp. 205-219). New York: Springer-Verlag.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med, 33:636-647.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1994) Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp, 2:189-210.

Grahn JA, Parkinson JA, Owen AM (2008) The cognitive functions of the caudate nucleus. Prog Neurobiol, 86:141-55.

Gresham G (2007) A study of mathematics anxiety in pre-service teachers. Early Child Educ J, 35:181-8.

Gross JJ (1998) Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression and physiology. J Pers Soc Psychol, 74:224–37.

Gross JJ (2002). Emotion regulation: Affective, cognitive, and social consequences. Psychophysiology, 39:281–91.

Hembree R (1990) The nature, effects, and relief of mathematics anxiety. J Res Math Educ, 21:33-46.

Hopko DR, Ashcraft MH, Gute J, Ruggiero KJ, Lewis C (1998) Mathematics anxiety and working memory: support for the existence of a deficient inhibition mechanism. J Anxiety Disord, 12:343-55.

Johns M, Inzlicht M, Schmader T (2008). Stereotype Threat and Executive Resource Depletion: Examining the Influence of Emotion Regulation. J Exp Psychol Gen, 137:691-705.

Kane MJ, Engle RW (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. Psychon Bull Rev, 9:637-71.

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci, 12:535-540.

Landau SM, Lal R, O'Neil JP, Baker S, Jagust WJ (2009) Striatal dopamine and working memory. Cereb Cortex, 19(2):445-54.

LeDoux JE (2000). Emotion circuits in the brain. Annu Rev Neurosci, 23:155-84.

Markman AB, Maddox WT, Worthy DA (2006) Choking and excelling under pressure. Psychol Sci, 17:944-8.

McCandliss BD (2010) Educational Neuroscience: The early years. Proc Nat Acad Sci USA, 107:8049-50.

National Mathematics Advisory Panel (2008) Foundations for Success: The Final Report of the National Mathematics Advisory Panel (U.S. Department of Education: Washington, DC).

Nicola SM, Taha SA, Kim SW, Fields HL (2005) Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues. Neuroscience, 135:1025-33.

Ochsner KN, Ray RD, Cooper JC, Robertson ER, Chopra S, Gabrieli JD, Gross JJ (2004) For better or for worse: neural systems supporting the cognitive down- and up-regulation of negative emotion. NeuroImage, 23:483-99.

Preacher KJ, Hayes AF (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. Behav Res Meth, 40:879-91.

Reinholdt-Dunne ML, Mogg K, Bradley BP (2009) Effects of anxiety and attention control on processing pictorial and linguistic emotional information. Behav Res T, 47:410-7.

Richardson FC, Suinn RM (1972) The Mathematics Anxiety Rating Scale: Psychometric data. J Couns Psychol, 19:551-554.

Rudner M, Fransson P, Ingvar M, Nyberg L, Rönnberg J (2007). Neural representation of binding lexical signs and words in the episodic buffer of working memory. Neuropsychologia, 45:2258-76.

Salamone JD (1994) The involvement of nucleus accumbens dopamine in appetitive and aversive motivation. Behav Brain Res, 61:117-33.

Salamone JD, Correa M, Farrar A, Mingote SM (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. Psychopharmacology, 191:461-82.

Simon O, Mangin JF, Cohen L, Le Bihan D, Dehaene S (2002). Topographical layout of hand, eye, calculation, and language-related areas in the human parietal lobe. Neuron, 33:475-87.

Spielberger CD, Gorsuch, R.L, Lushene, R.E. (1970) Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press.

Unsworth N, Heitz RP, Schrock JC, Engle RW (2005) An automated version of the operation span task. Behav Res Meth, 37:498-505.

Utter AA, Basso MA (2008) The basal ganglia: An overview of circuits and function. Neurosci Biobehav R, 32:333-42.

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O (2006) Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. NeuroImage, 30:1414-1432.

Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. Perspect Psychol Sci, 4:274-290.

White NM (2009) Some highlights of research on the effects of caudate nucleus lesions over the past 200 years. Behav Brain Res, 199:3-23.

Wieser MJ, Pauli P, Mühlberger A (2009) Probing the attentional control theory in social anxiety: an emotional saccade task. Cogn Affect Behav Neurosci, 9:314-22.

TABLES

Table 1

Region	Tal. Coords. x	y	z	Volume (mm^3)	Correlations HMAs	LMAs	Group × Slope
R.IFJ	44	13	31	2124	-.762 (.116) *p*=.0015	-.151 (.271) *p*=.6064	*p*=.0192
L.IFJ	-45	7	36	1308	-.778 (.109) *p*=.0011	.168 (.270) *p*=.5667	*p*=.0034
R.IPL	43	-44	37	951	-.816 (.093) *p*=.0004	-.279 (.256) *p*=.3337	*p*=.0011
L.IPL	-38	-48	33	740	-.798 (.101) *p*=.0006	.087 (.275) *p*=.7679	*p*=.0006
L.IFGa	-40	43	-4	942	-.720 (.133) *p*=.0037	-.064 (.276) *p*=.8267	*p*=.0330
R.Caud & NAc	9	14	2	1420	-.882 (.061) *p*<.0001	.451 (.221) *p*=.1059	*p*<.0001
L.Hipp	-23	-23	-5	800	-.849 (.077) *p*=.0001	.189 (.267) *p*=.5186	*p*=.0005

Table 1 Caption

ROI Details for Analyses 1 (top) and 2 (bottom). For Analysis 1, *Cue*-activity (math>word) was correlated with math deficits (math>word error-rates). For Analysis 2, *Task*-activity (math>word) was correlated with math deficits (math>word error-rates). Note that all ROIs were localized for HMAs. The right-most three columns reflect correlation details for each region. The first two of these columns report mean *r*-values (i.e., averaged across all voxels in the ROI), corresponding standard errors (in parentheses), and *p*-values (bottom row) for each group. The final column reports the *p*-value associated with the interaction term testing for whether the observed correlation was in fact specific to the HMAs.

Table 2

Region	Correlation Splits Math β	Correlation Splits Word β	HMA Raw βs Math β	HMA Raw βs Word β	LMA Raw βs Math β	LMA Raw βs Word β
R.IFJ	-.726 (.131) *p*=.0033	-.051 (.277) *p*=.8626	.001 (.113)	-.127 (.096)	.152 (.104)	-.162 (.112)
L.IFJ	-.625 (.169) *p*=.0170	.117 (.274) *p*=.6896	.008 (.103)	-.068 (.088)	.130 (.082)	-.123 (.105)
R.IPL	-.729 (.130) *p*=.0031	-.170 (.269) *p*=.5601	-.086 (.101)	-.016 (.075)	-.150 (.106)	-.177 (.098)
L.IPL	-.696 (.143) *p*=.0057	-.103 (.274) *p*=.7259	-.029 (.100)	.153 (.074)	.012 (.122)	.102 (.088)
L.IFGa	-.032 (.277) *p*=.9129	.584 (.183) *p*=.0282	-.013 (.106)	.031 (.123)	-.060 (.094)	.012 (.095)
R.Caud & NAc	-.630 (.167) *p*=.0158	.116 (.274) *p*=.6935	.042 (.079)	-.006 (.094)	.230 (.066)	.135 (.048)
L.Hipp	-.723 (.132) *p*=.0035	-.063 (.276) *p*=.8302	.128 (.100)	.177 (.070)	.011 (.089)	-.009 (.080)

Table 2 Caption

ROI Verification Results for Analyses 1 (top) and 2 (bottom). For Analysis 1, *Cue*-activity (math *or* word) was correlated with math deficits (math>word error-rates). For Analysis 2, *Task*-activity (math *or* word) was correlated with math deficits (math>word error-rates). Columns labeled 'Correlation Splits' contain HMA correlation results for each task β separately (i.e., the correlation between raw, non-subtracted, subject βs and performance deficits). Cell values report mean *r*-values (i.e., averaged across all voxels in the ROI), corresponding standard errors (in parentheses), and *p*-values (bottom row) for each cue/task type. LMAs were not considered here as there was no correlation for LMAs in these regions to begin with. For Analysis 1, these data demonstrate that the correlations in IFJ and IPL regions were driven primarily by math-cue activity; by contrast the correlation in left IFGa was driven primarily by word-cue activity. For Analysis 2, correlations in both right caudate/nucleus accumbens and left hippocampus were primarily driven by math-task activity. The right-most four columns contain raw (non-subtracted) βs (and standard errors) for both groups and conditions.

Table 3

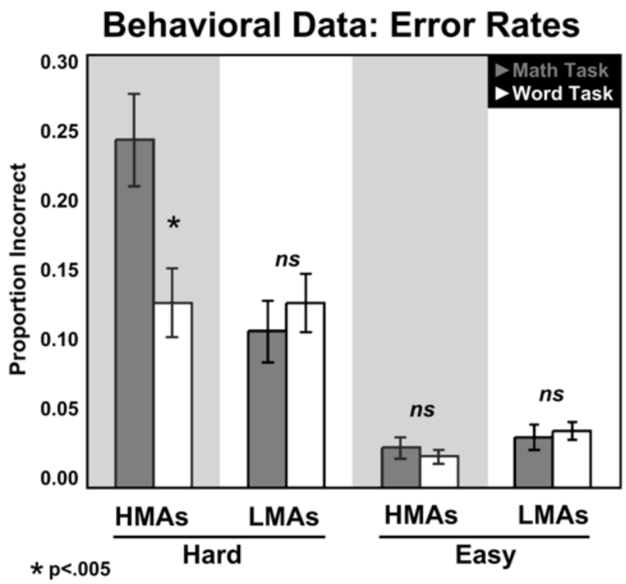
Independent Variable (Cue Activity)		Combined Mediators (Task Activity) L. Hipp & R.Caud/NAc	
Region	**Orig. IV~DV Effect (*c*)**	**Mediation Effect (*ab*)**	**Unmediated Effect (*c*')**
R IFJ	-.246 (.060) *p*=.0015	-.152 (.061) *p*=.0129	-.094 (.064) *p*=.1771
L IFJ	-.258 (.060) *p*=.0011	-.156 (.058) *p*=.0071	-.102 (.060) *p*=.1219
R IPL	-.355 (.073) *p*=.0004	-.210 (.077) *p*=.0066	-.145 (.088) *p*=.1307
L IPL	-.334 (.073) *p*=.0006	-.202 (.073) *p*=.0056	-.132 (.080) *p*=.1283
L IFGa	-.223 (.062) *p*=.0037	-.139 (.060) *p*=.0206	-.084 (.059) *p*=.1839
Network Average	-.318 (.064) *p*=.0003	-.182 (.068) *p*=.0071	-.154 (.079) *p*=.0802

Table 3 Caption

Mediation analysis for HMAs. The six models considered are listed in rows under Region. Original IV~DV effect (denoted *c* in Figure 5) refers to the original cue~performance relationship to be explained. Total mediation effect (product of *a* and *b* in Figure 4) refers to the IV~DV effect indirectly explained by the combined influence of mediator (*task* activity in both left hippocampus and right caudate/NAc. (Individual region contributions for left hippocampus and right caudate/NAc can be found in Supplementary Information.) Non-mediated IV~DV effect (denoted *c*' in Figure 4) refers to the remaining effect of the IV on the DV that cannot be explained by the mediators. Cell values: linear regression estimate, (std. err.), *p*-value.

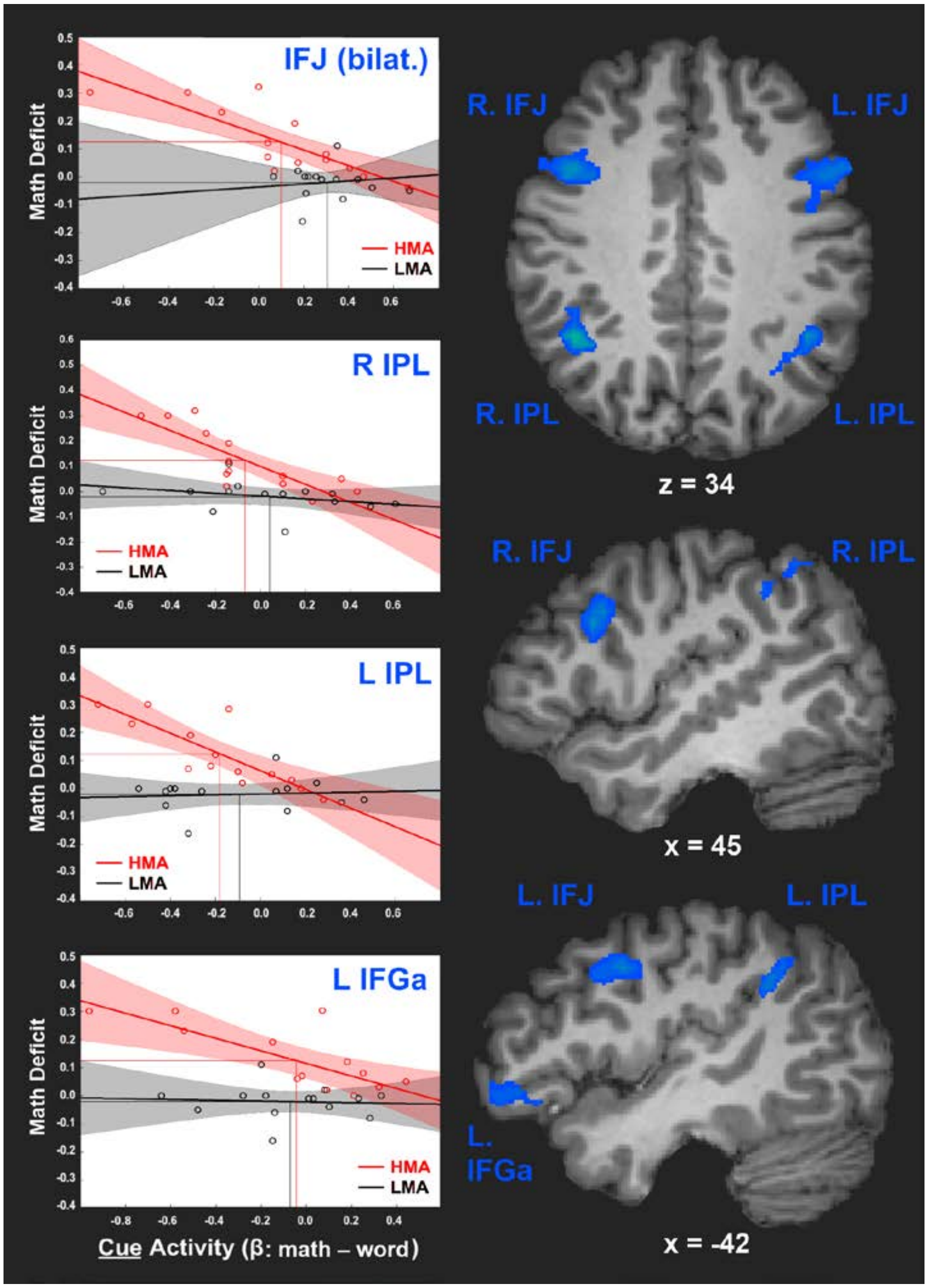
FIGURES

Figure 1



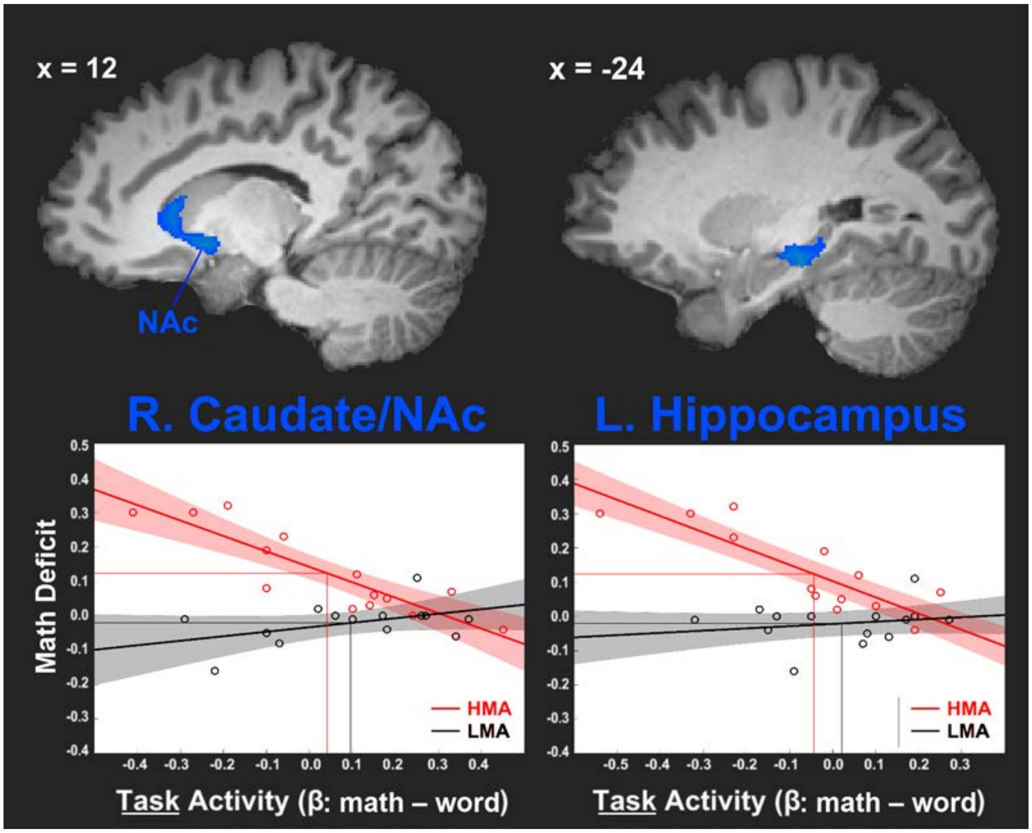
Behavioral Results: Error bars indicate standard errors of the mean. Note that only HMAs showed a significant difference between tasks (more errors on the math task).

Figure 2



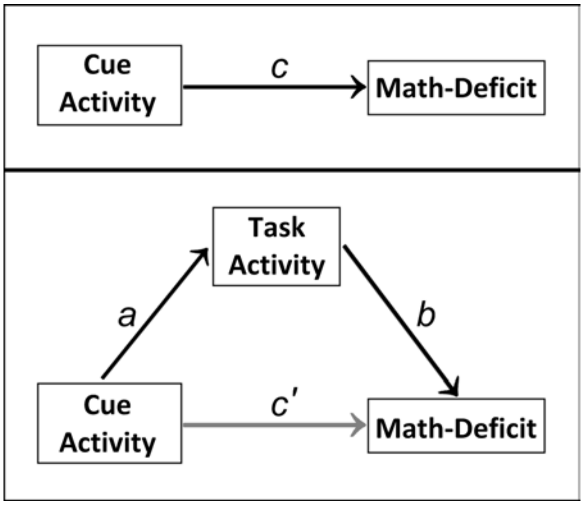
Depiction of regions and scatter plots from Analysis 1. No effects of hemisphere were found for IFJ regions; thus data are averaged across the two hemispheres in the scatter plot above. Shaded areas in the scatter plots represent 95% confident intervals (observed) for the fitted linear regression line. MC=math-cue; WC=word-cue. Y-axis: math-deficits [higher number = higher math (relative to word) error-rate].

Figure 3



Analysis 2 regions and scatter plots. Note that the right caudate region spans both the medial head of the caudate and nucleus accumbens (NAc). Shaded areas in the scatter plots represent 95% confident intervals (observed) for the fitted linear regression line. Y-axis: math-deficits [higher number = higher math (relative to word) error-rate].

Figure 4



Mediation Framework: We tested whether task-activity in regions identified in Analysis 2 mediated the relationship *c* between math-deficits and cue-activity in regions identified in Analysis 1. In this framework, one asks whether there is a significant indirect effect of the mediator (quantified as the product of the unstandardized path coefficients, *a* and *b*, in the figure above) that accounts for some portion of the direct effect *c* originally observed between the independent (in this case, cue-activity) and the dependent (in this case, math-deficit) variables. When more than one mediator is considered at a time, the total mediating influence is the sum of the products $a_i b_i$, for i=1…n mediators. The remaining (unmediated) direct effect is denoted *c*'. Full mediation occurs when *ab* is significant but *c*' is not; partial mediation refers to when both *ab* and *c*' remain significant. Note that in this framework, the model is constrained by the assumption that $c = ab+c'$. In other words, unlike in a standard multiple regression analysis, we are explicitly asking what portion of the IV~DV (cue-activity~math-deficit) relationship can be accounted for by the proposed mediating variable (task-activity).