Developmental Science

Developmental Science (2016), pp 1–15 DOI: 10.1111/desc.12450

PAPER

Functional brain organization of working memory in adolescents varies in relation to family income and academic achievement

Amy S. Finn,[1] Jennifer E. Minas,[2] Julia A. Leonard,[2] Allyson P. Mackey,[2] John Salvatore,[2] Calvin Goetz,[2] Martin R. West,[3] Christopher F.O. Gabrieli[3,4] and John D.E. Gabrieli[2,3]

1. *Department of Psychology, University of Toronto, Canada*
2. *Department of Brain and Cognitive Sciences and McGovern Institute for Brain Research, Massachusetts Institute of Technology, USA*
3. *Graduate School of Education, Harvard University, USA*
4. *Transforming Education/National Center on Time & Learning, Boston, USA*

Abstract

Working memory (WM) capacity reflects executive functions associated with performance on a wide range of cognitive tasks and education outcomes, including mathematics achievement, and is associated with dorsolateral prefrontal and parietal cortices. Here we asked if family income is associated with variation in the functional brain organization of WM capacity among adolescents, and whether that variation is associated with performance on a statewide test of academic achievement in mathematics. Participants were classified into higher-income and lower-income groups based on family income, and performed a WM task with a parametric manipulation of WM load (N-back task) during functional magnetic resonance imaging (fMRI). Behaviorally, the higher-income group had greater WM capacity and higher mathematics achievement scores. Neurally, the higher-income group showed greater activation as a function of WM load in bilateral prefrontal, parietal, and other regions, although the lower-income group exhibited greater activation at the lowest load. Both groups exhibited positive correlations between parietal activations and mathematics achievement scores, but only the higher-income group exhibited a positive correlation between prefrontal activations and mathematics scores. Most of these findings were maintained when higher- and lower-income groups were matched on WM task performance or nonverbal IQ. Findings indicate that the functional neural architecture of WM varies with family income and is associated with education measures of mathematics achievement.

Research highlights

- Higher family income was associated with greater working memory capacity and higher scores on a statewide test of math achievement.
- Higher family income was associated with greater activation of the fronto-parietal executive network for demanding working memory conditions.
- Higher scores on the statewide test of math achievement were associated with greater parietal activation during working memory performance.
- Exploring ways in which working memory ability can be enhanced in students from lower-income backgrounds may help reduce the income–achievement gap.

Introduction

Children from lower-socioeconomic status (SES) environments, relative to higher-SES environments, perform worse on many measures of cognitive ability, including working memory (WM) capacity, which determines how much goal-relevant information can be held and manipulated in mind (Hackman, Gallop, Evans & Farah, 2015; Noble, Norman & Farah, 2005). Children

Address for correspondence: Amy S. Finn, University of Toronto, Department of Psychology, 100 St George Street, Room 4002, Toronto, ON M5S 3G3, Canada; e-mail: finn@psych.utoronto.ca

from lower-SES backgrounds also perform less well on measures of academic achievement, such as school-based standardized tests, a pattern that is referred to as the income–achievement gap (Brooks-Gunn, Guo & Furstenberg, 1993; Duncan, Yeung, Brooks-Gunn & Smith, 1998; Noble *et al*., 2005). Here we asked whether the functional brain organization of WM capacity (1) differs in relation to SES, and (2) relates to performance on a statewide measure of academic achievement in mathematics.

WM capacity is hypothesized to be a major determinant of both cognitive performance and educational outcomes (Gathercole, Pickering, Knight & Stegmann, 2004). Greater WM capacity is associated with superior reading comprehension, problem solving, and inhibitory control (Conway, Kane & Engle, 2003). Greater WM capacity is also associated with better performance on tests of academic achievement, especially in mathematics (Finn, Kraft, West, Leonard, Bish *et al*., 2014; Gathercole *et al*., 2004). WM capacity is a predictor of performance on multiple measures of mathematics ability (LeBlanc & Weber-Russell, 1996; Passolunghi & Siegel, 2001; Raghubar, Barnes & Hecht, 2010; Zheng, Swanson & Marcoulides, 2011) and limited WM capacity could serve as a bottleneck that constrains mathematics performance (Swanson & Beebe-Frankenberger, 2004). In addition, mediation analyses show that executive functions, including measures of working memory, partially mediate the relation between SES and longitudinal change in math achievement (Lawson & Farah, 2015).

Despite clear evidence that WM capacity is an essential determinant of cognitive ability, is influenced by SES, and is important for academic achievement in mathematics, there is no evidence about how the functional brain organization of WM capacity varies with SES or how it relates to academic achievement in mathematics. WM depends on a neural system that includes the dorsolateral prefrontal cortex (DLPFC), parietal cortices, and the basal ganglia (Awh & Vogel, 2008; Curtis & D'Esposito, 2003; Goldman-Rakic, 1987). Magnetic resonance imaging (MRI) studies have found that lower SES is associated with reduced gray matter (Hanson, Hair, Shen, Shi, Gilmore *et al*., 2013; Jednoróg, Altarelli, Monzalvo, Fluss, Dubois *et al*., 2012), including prefrontal, temporal, and parietal cortices, and hippocampus (Hanson, Chandra, Wolfe & Pollak, 2011; Lawson, Duda, Avants, Wu & Farah, 2013; Luby, Belden, Botteron, Marrus, Harms *et al*., 2013; Noble, Houston, Brito, Bartsch, Kan *et al*., 2015; Noble, Houston, Kan & Sowell, 2012). Further, reduced cortical thickness has been associated with lower scores on statewide academic achievement tests (Mackey, Finn, Leonard, Jacoby-Senghor, West *et al*., 2015). Although no functional imaging study has examined the physiological bases of reduced WM capacity in lower-SES children, functional MRI (fMRI) and electroencephalography (EEG) studies have reported altered activations in relation to SES for language processing (Raizada, Richards, Meltzoff & Kuhl, 2008), stimulus–response rule learning (Sheridan, Sarsour, Jutte, D'Esposito & Boyce, 2012), and attention (Kishiyama, Boyce, Jimenez, Perry & Knight, 2009; Stevens, Lauinger & Neville, 2009).

We therefore examined, for the first time, how the functional brain organization of WM capacity (1) differs in children from higher- versus lower-SES environments, and (2) relates to statewide measures of academic achievement in mathematics. The participants were 7th and 8th graders attending public schools, and familial SES environment was operationalized by whether or not the children qualified for free or reduced lunch on the basis of family income. We hypothesized that children from higher-SES backgrounds would score higher on the statewide test of academic achievement in mathematics, and exhibit a greater WM capacity. Prior studies in adults have reported that individuals with lower WM ability tend to have disproportionately poor performance at higher WM loads and a reduction in DLPFC activation at these higher loads (e.g. (Cubillo, Smith, Barrett, Giampietro, Brammer *et al*., 2014; Jansma, Ramsey, van der Wee & Kahn, 2004; Perlstein, Carter, Noll & Cohen, 2001)), so we hypothesized that lower-SES children would exhibit reduced activation in DLPFC, and perhaps parietal regions also associated with WM, at higher WM loads. Although no prior study has related WM activations to achievement tests in mathematics, we hypothesized that the same DLPFC and parietal regions known to be sensitive to WM demands would also be related to such academic outcomes.

Method

Participants

Written consent was obtained from students and parents. Participants were recruited from urban and suburban middle schools. From among 83 participants, results are reported from 67 middle school students; the other 16 participants were excluded for one or more reasons (achievement test score unavailable ($n = 1$); unable to complete the task ($n = 1$); parental report of severe language delay ($n = 1$); excessive motion during scanning defined as movement on more than 10% of trials (or 30 total volumes; $n = 7$); did not understand or could not perform the task, as indicated by being unable to detect at least half of targets on the two easiest conditions ($n = 5$) or having more false alarms than hits on any of the conditions ($n = 3$).

Among the 67 participants, 36 students qualified for free or reduced-price lunch at any time in the 3 years previous to participation (nine excluded participants fit this criteria), indicating that their family had an income at or below 185% of the poverty level ($42,200 or less per year for a family of four at the time these data were collected). This lower-income group had a mean age of 14.33 years (*SD*: 0.66 years), was 64% female, 50% African-American, 3% Asian, and 11% White; 53% were Hispanic. Thirty-one students did not qualify for free or reduced lunch in the 3 years previous to participation (seven excluded participants fit this criteria). This higher-income group had a mean age of 14.48 years (*SD*: 0.40 years), was 45% female, 10% African-American, 19% Asian, and 58% White; 3% were Hispanic (Table 1).

Procedure

Participants came to the Massachusetts Institute of Technology for testing and scanning. After consenting, and to obtain a measure of working memory outside of the scanner, they completed the Count Span task (Conway, Kane, Bunting, Hambrick, Wilhelm *et al.*, 2005) in a quiet testing room. They then practiced scanning (mock scanner), and were scanned while completing an N-back task (Owen, McMillan, Laird & Bullmore, 2005). After completing the scan, participants completed the Test of Nonverbal Intelligence (TONI) (Brown, Sherbenou & Johnsen, 2010).

Functional MRI data were acquired on a Siemens MAGNETOM Trio 3T MR Scanner at the Massachusetts Institute of Technology. Information about the structural and functional sequences is detailed in the Supplemental Material.

Apparatus and stimuli

N-back task

Participants were presented with letters (one at a time) and asked to indicate (with a button press) if the presented letter was the same as a letter that was presented a certain number (N) of screens previously (Figure 1a). Prior to beginning a block of trials, participants were instructed to look for a match, 1-, 2-, or 3-back, or to detect the letter ‘W’ on each trial (0-back), and asked to press a button only if there was a match (and not to press otherwise). Letters were drawn from a pool of eight (b, f, h, j, q, m, r, w) and were presented in both upper and lower case. Each letter was presented in the middle of an array of empty circles for 500 ms and was followed by 2500 ms of fixation before the next letter was presented. Participants were given 2900 ms from letter onset to respond to each trial (Figure 1a). Thirty percent of all trials were targets (or repetitions requiring an affirmative response). Each block lasted 45.5 seconds, contained 10 letters, and began with an initial fixation of 500 ms, and an instruction screen that indicated block-type for 3000 ms. Each block was followed by 12 seconds of rest, during which the screen showed a fixation cross on a black screen. Participants completed two 9.1-minute runs of the N-back task, with each run containing 12 blocks. The experiment lasted just over 18 minutes. For each load (0–3), performance was measured by calculating *d* prime, and latency by averaging response times for accurate responses.

Count Span task

To enhance construct validity and evaluate the possible effect of scanning on behavioral measurement during working memory performance, we included an additional working memory task that was collected outside of the scanner. This Count Span task was administered using Psychopy software (Peirce, 2007). In this task, based on previous work (Conway *et al.*, 2005; Cowan, Elliott, Saults, Morey, Mattox *et al.*, 2005), participants viewed an array with blue circles, blue triangles, and red circles, and were instructed to count only the blue circles (targets) and to press the space bar to move on (a trial would time out after 5 seconds if there was no response). After one or more arrays, participants entered the number of targets presented in each display in the order they were presented. Loads ranged from 1 to 6 consecutive arrays with three instances of each load presented in random order. A participant’s Count Span score was the highest load (from 1 to 6) at which 2 of 3 trials were

Table 1 *Age, gender, race and ethnicity by income group*

Income group	Age (yrs)	Female	African-Amer	Asian	White	Hispanic	Grade (at scan)	
Lower	14.33	64%	50%	3%	11%	53%	89%	8th grade (11% 7th)
Higher	14.48	45%	10%	19%	58%	3%	100%	8th grade (0% 7th)

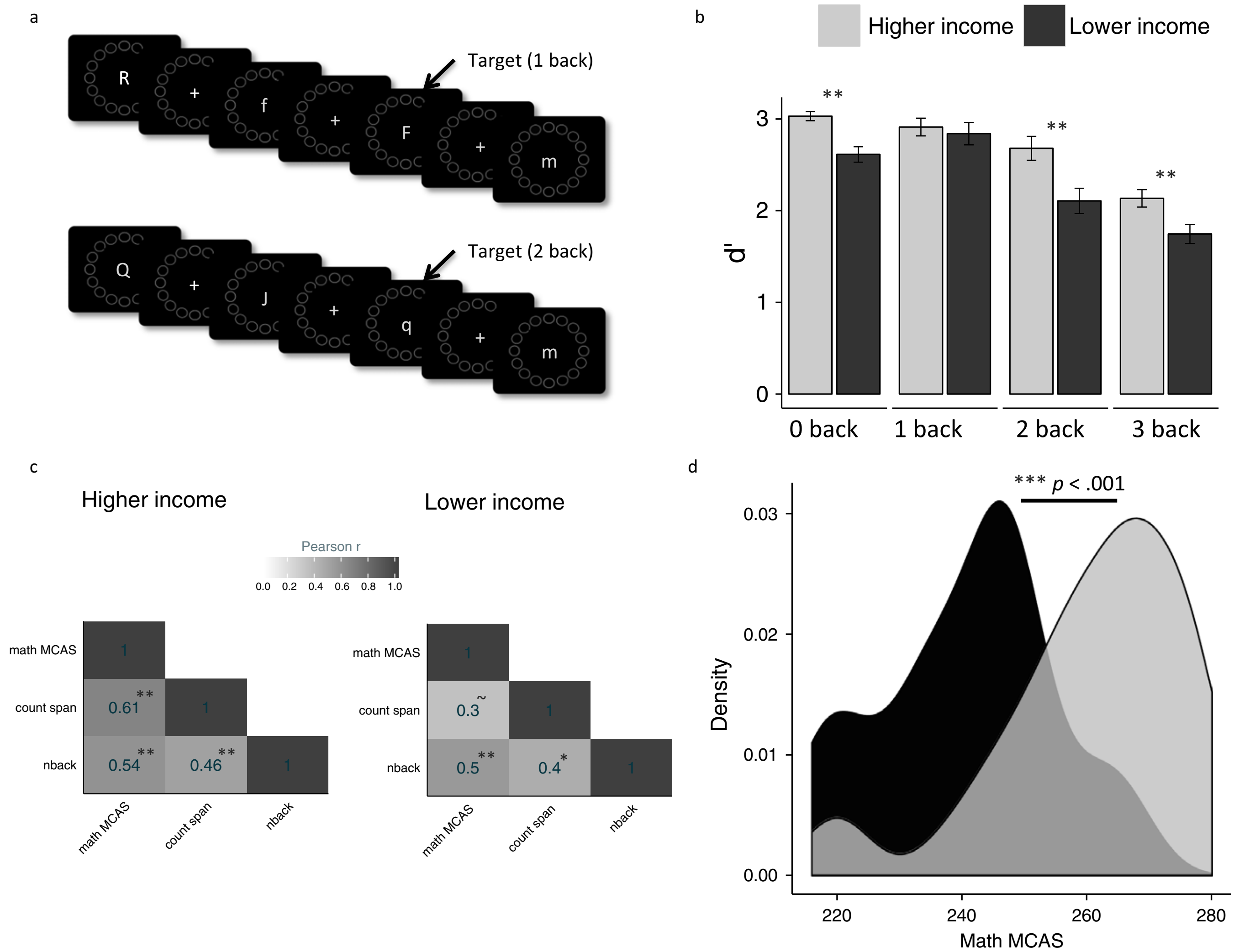

Figure 1 *Task design and behavioral performance. (a) The N-back task required participants to detect repeating stimuli N (0–3) items back. (b) The higher-income group (light gray bars) performed significantly better than the lower-income group (dark gray bars) on 0-back, 2-back and 3-back conditions, but groups did not differ significantly in the 1-back condition. (c) Performance on both N-back and Count-Span tasks correlated with mathematics achievement scores. (d) The higher-income group scored significantly higher than the lower-income group on the MCAS mathematics test (d). Error bars represent standard error of the mean.* ** = p < *.05 (Bonferroni adjusted alpha = .008),* * = p < *.1 (Bonferroni adjusted alpha = .017),* ~ = p < *.1* (p = *.084, uncorrected for relation between Count Span and MCAS math in the lower-income group).*

answered correctly, plus 0.5 if 1 of 3 trials at the next highest load was answered correctly (Daneman & Carpenter, 1980).

Test of Nonverbal Intelligence

In the Test of Nonverbal Intelligence (Version B; Brown *et al.*, 2010) participants were seated across from an experimenter and were asked to choose, by pointing, which of six pictures completed the missing piece of a puzzle. Choosing the correct response requires the integration of increasingly difficult information. Per instructions of this standardized test, the experimenter stopped after participants made three incorrect selections in a row. Reponses were scored (total number correct) and converted to age-normed scores provided by the publisher. Because this measure was collected last, data are missing from four lower-income students (and 0 higher-income students).

Standardized test scores and low-income status

Scores on the Massachusetts Comprehensive Assessment System (MCAS) and free or reduced-price lunch data were obtained from the Massachusetts Department of Elementary and Secondary Education. Scaled scores from the 2012 MCAS mathematics test were used; these values represented student performance relative to grade-level expectations, and allowed for comparison across students enrolled in 7th and 8th grade.

Behavioral and mediation analyses

Performance was compared across groups using mixed model ANOVAs and independent samples *t*-tests. For mediation analyses, statistical significance was tested using the Sobel test (Sobel, 1982) after checking for statistically significant relationships between (1) the independent variable (IV; income group) and the dependent variable (DV; MCAS performance), (2) the IV and the hypothesized mediators (WM activations and WM performance), and (3) the hypothesized mediators and the DV. Neural activations that correlated significantly with MCAS math scores (across all four loads) were extracted for each participant; average discrimination sensitivity across all four loads was calculated. Mediation analyses controlled for IQ (standardized Test of Nonverbal Intelligence scores). All analyses relating behavior to brain data controlled for motion (average framewise displacement, root mean square values of rotation and translation (fdrms; Jenkinson, 2003).

fMRI analysis

Functional MRI data processing and analysis were completed using Nipype and the Brain Imaging Pipeline (BIPS) (Ghosh, Keshavan, Salvatore & Klein, 2012; Gorgolewski, Burns, Madison, Clark, Halchenko *et al.*, 2011). The pre-processing of data is detailed in the Supplemental Materials. Analysis of movement showed that the higher- and lower-income groups did not differ significantly (average framewise displacement, root mean square values of rotation and translation (fdrms; Jenkinson, 2003); higher versus lower income: $U = 422$, $p = .087$, $r = .209$). Fdrms values were included as a covariate in all mixed effects analyses.

A statistical parametric map was calculated for each participant based on linear combinations of the covariates modeling each task period (0-, 1-, 2- and 3-back; Jenkinson, Beckmann, Behrens, Woolrich & Smith, 2012). Regressors for motion and motion outliers were included in this first-level model. Data were then coregistered to individual structural volumes (using bbregister in Freesurfer; Greve & Fischl, 2009) and normalized using Advanced Normalization Tools (ANTs Software; Avants, Tustison, Song, Cook, Klein *et al.*, 2011). Data were then analyzed using mixed effects higher-level modeling (Woolrich, Behrens, Beckmann, Jenkinson & Smith, 2004). To identify regions related to MCAS performance, a behavioral regressor was included in a second-level full brain regression (including fdrms as a covariate). Data were corrected for multiple comparisons by using threshold free cluster enhancement (TFCE) followed by permutation testing to adjust for family wise error (Smith & Nichols, 2009).

Regions of interest (ROIs) were created from functional clusters within anatomically defined regions. Regions were identified from the omnibus contrast (3 > 2>1 > 0-back) from the combined group data ($n = 67$) so as to be unbiased by either group. Clusters from this contrast were identified within two anatomical regions using anatomical masks (the middle frontal gyrus and lateral parietal cortex (including the intraparietal sulcus, based on the Harvard-Oxford Atlas; Kennedy & Haselgrove, 2011) on the left and right sides at a cluster forming threshold of $p < .05$ and a cluster probability threshold of $p < .05$.

Results

Test of Nonverbal Intelligence (IQ)

The higher-income group ($M = 100.35$, $SD = 10.2$) scored higher than the lower-income group ($M = 95.97$, $SD = 6.71$) on the Test of Nonverbal Intelligence, $t(61) = 2.02$, $p = .048$, $d = .518$. Given this difference, IQ-matched groups were created from a subset of participants and additional analyses were performed with these groups. To equate groups, we excluded the three highest scoring higher-income participants (TONI scores = 127, 121, 121) and the lowest scoring lower-income participant (TONI score = 79). In this matched subset, the higher-income group ($M = 97.93$, $SD = 7.22$) did not differ from the lower-income group ($M = 96.55$, $SD = 5.95$) on the Test of Nonverbal Intelligence, $t(57) = .80$, $p = .425$, $d = .210$.

N-back performance

Across all participants, a mixed model analysis of variance (ANOVA) indicated that accuracy (discrimination sensitivity, d') differed as a function of both WM load ($F(3, 195) = 57.04$, $p < .001$, $\eta_p^2 = .47$) and group ($F(1, 65) = 9.98$, $p = .002$, $\eta_p^2 = .13$), and that the effect of load differed by group ($F(3, 195) = 3.31$, $p = .021$, $\eta_p^2 = .05$; Figure 1b). Reaction times likewise differed by load ($F(3, 195) = 27.64$, $p < .001$, $\eta_p^2 = .298$) and group ($F(1, 65) = 5.14$, $p = .027$, $\eta_p^2 = .\ 073$), but there was no load-by-group interaction ($F(3, 195) = .223$, $p = .638$, $\eta_p^2 = .003$; Figure 1b, Table 2).

Across groups, performance was worst for the highest load (3-back vs. 2-back: d': $t(66) = 6.09$, $p < .001$, $d = .79$, Bonferroni corrected $\alpha = .017$), second worst for 2-back (2-back vs. 1-back: d': $t(66) = 5.70$, $p < .001$, $d = .72$, Bonferroni corrected $\alpha = .017$), and best on 1- and 0-back (for which discrimination sensitivity did not differ: d': $t(66) = .74$, $p = .46$, $d = .09$, Bonferroni corrected $\alpha = .017$). This pattern was similar within each group (Supplementary Table 1).

The higher-income group was more accurate than the lower-income group at all WM loads except for 1-back (3-back: d': $t(65) = 2.71$, $p = .009$, $d = .67$; proportion correct: $t(65) = 3.07$, $p = .003$, $d = .760$; 2-back: d': $t(65) = 2.30$, $p = .004$, $d = .74$; proportion correct: $t(65) = 3.2$, $p = .002$, $d = .788$; 1-back: d': $t(65) = .46$, $p = .65$, $d = .12$; proportion correct: $t(65) = .863$, $p = .391$, $d = .216$; 0-back: d': $t(65) = 4.01$, $p < .001$, $d = 1.06$; proportion correct: $t(65) = 3.83$, $p < .001$, $d = 1.061$; Bonferroni corrected $\alpha = .013$). This pattern of data was replicated in the IQ-matched sample (Supplementary Materials). Performance on 3-back in the higher-income group (mean $d' = 2.12$) was statistically indistinguishable from performance on 2-back in the lower-income group (lower-income mean $d' = 2.11$; difference: $t(65) = .159$, $p = .87$, $d = .04$; Figure 1b).

Count Span performance

The higher-income group ($M = 4.43$, $SD = 1.44$) performed better than the lower-income group ($M = 3.37$, $SD = 1.48$; $t(64) = 2.94$, $p = .005$, $d = .73$), a pattern that was the same in the IQ-matched sample (Supplementary Materials).

Mathematics achievement scores

The higher-income group ($M = 260.5$, $SD = 15.2$) had higher Mathematics MCAS scores than the lower-income group ($M = 240.9$, $SD = 13.99$) ($t(65) = 5.5$ $p < .001$, $d = 1.35$; Figure 1d; the pattern was the same in IQ-matched sample (Supplementary Materials)). MCAS scores can range from 200 to 280, with scores above 240 classified as proficient. In all, 90.3% of the higher-income group (all 8th graders) scored as proficient, while 53% of the lower-income group (57.1% of 7th graders ($n = 7$) and 51.7% of 8th graders ($n = 31$)) scored as proficient. In comparison, 37% of 8th grade students across the state who received free or reduced-price lunch scored proficient or above, compared to 61% of students who did not receive free lunch. As with many studies that require students to come into a laboratory setting, our participants were higher-performing than would be expected from a random sample. Nevertheless, the percentages of students reaching proficiency differed substantially between the higher- and lower-income groups (37%) as it did across the state (24%).

Correlations among WM and achievement scores

The two measures of working memory capacity (Count Span and N-back) correlated significantly within both groups. Performance on the mathematics achievement test (MCAS mathematics) also correlated significantly and positively with performance on Count Span and on N-back tasks in the higher-income group and correlated significantly positively with N-back performance in the lower-income group (Figure 1c).

Brain activation during N-back

Across all participants, there was significantly greater activation as a function of greater WM load (3-back > 2-back > 1-back > 0-back) in brain regions associated with WM (Figure 2a), including bilateral middle and inferior frontal gyri (MFG and IFG), bilateral intraparietal sulcus (IPS), bilateral caudate and putamen, and bilateral cerebellum (Table 3). The higher-income group recruited all of these regions bilaterally (MFG, IPS,

Table 2 *Performance on N-back*

	0-back			1-back			2-back			3-back		
Condition	Mean	*SD*	Range	Mean	*SD*	Range	Mean	*SD*	Range	Mean	*SD*	Range
d prime												
lower-income	2.6	.51	1.2–3.3	2.8	.73	1.5–4.2	2.1	.82	.69–3.6	1.7	.63	.42–2.7
higher-income	3.0	.28	2.4–3.4	2.9	.54	.96–3.9	2.7	.73	.91–3.9	2.1	.53	.36–3.0
proportion correct												
lower-income	.90	.048	.76–.95	.91	.060	.76–1.0	.85	.076	.89–.96	.81	.062	.68–.91
higher-income	.94	.018	.90–.97	.92	.047	.73–.98	.90	.068	.68–.98	.86	.054	.65–.93
reaction time (sec)												
lower-income	.589	.14	.405–.985	.714	.199	.458–1.18	.731	.215	.410–1.38	.776	.269	.428–1.39
higher-income	.518	.09	.394–.754	.597	.118	.429–.871	.658	.183	.415–1.14	.716	.152	.453–1.11

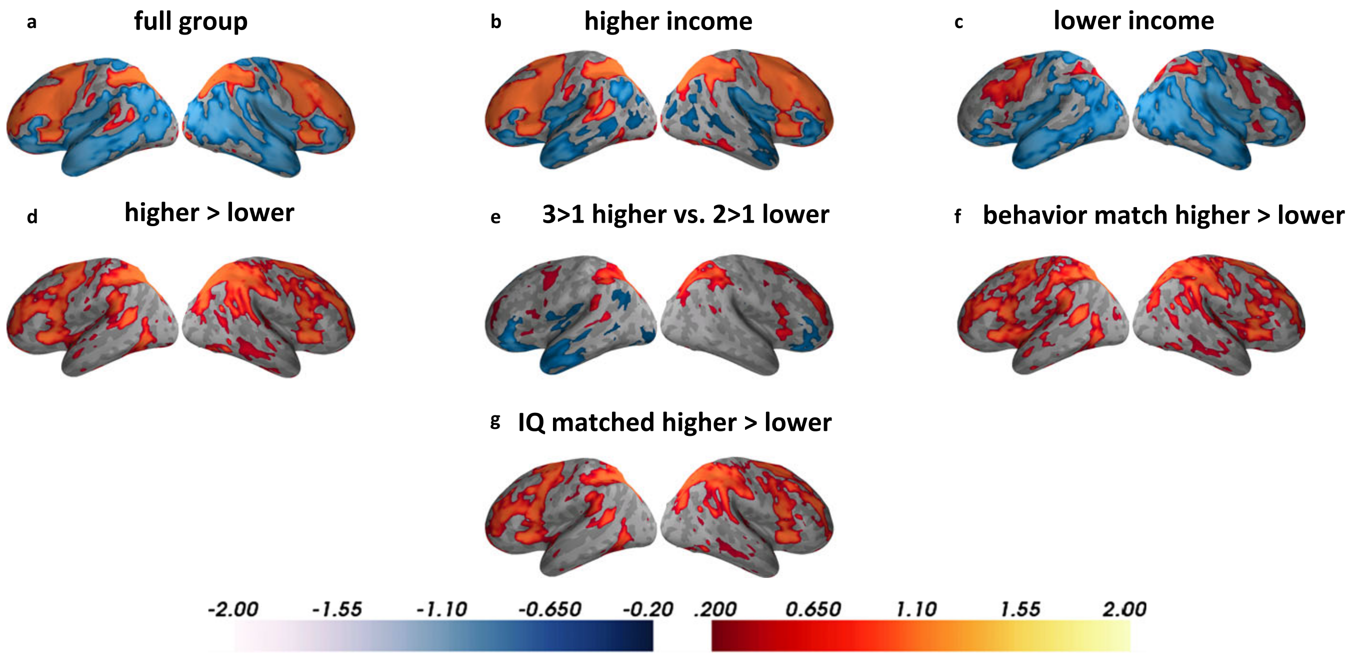

Figure 2 *WM load-related activation during N-back. Parametric activation (for the contrast 3 > 2 > 1 > 0-back) displayed for (a) the full group; (b) higher-income group; (c) lower-income group; (d) the difference between the higher- and lower-income groups (higher greater = warm colors, lower greater = cool colors); (e) for the behaviorally matched 3 > 1-back contrast in the higher-income group and 2 > 1-back contrast in the lower-income group; for (f) subsets of the two groups matched for overall performance; and (g) subsets of the two groups matched for IQ. Color bars in this and all other figures display z-scores ranging from .2 to 2. Data were corrected for multiple comparisons by using threshold free cluster enhancement (TFCE) followed by permutation testing to adjust for family wise error (FWE).*

caudate, putamen and cerebellum; Figure 2b and Table 3) with increasing load. The lower-income group, however, only showed significant load-related recruitment of left MFG and bilateral IPS (Figure 2c). When compared directly, the higher-income group recruited several regions significantly more as a function of load than the lower-income group, including bilateral MFG, IFG, IPS, caudate, putamen and cerebellum (Figure 2d; Table 3). The lower-income group did not recruit any regions significantly more than the higher-income group as a function of load.

To determine whether these activation differences might be driven by group differences in WM performance during scanning or IQ, we performed four additional analyses. First, we compared activation for performance-matched contrasts across all participants: 3-back > 1-back for the higher-income group versus 2-back > 1-back for the lower-income group. Despite the equated performance differences in the two contrasts, the higher-income group still showed significantly greater activation in bilateral MFG, IPS, cerebellum, and right caudate (Figure 2e; Table 3). Second, we compared activation in a subset of students (29 in each group) for whom performance did not differ significantly (d': $F(1, 56) = 1.7$, $p = .20$, $\eta_p^2 = 029$; reaction time: $F(1, 56) = 2.66$, $p = .109$, $\eta_p^2 = .045$). Even in this smaller sample, the higher-income group still showed significantly greater activation with increasing difficulty (3-back > 2-back > 1-back > 0-back) in bilateral MFG, IPS, cerebellum, and right caudate (Figure 2f). Third, we compared the IQ-matched samples (Test of Nonverbal Intelligence: $t(57) = .80$, $p = .425$, $d = .210$), and the higher-income group showed significantly greater activation with increasing difficulty (3-back > 2-back > 1-back > 0-back) in bilateral MFG, IPS, cerebellum, and right caudate (Figure 2g). Fourth, we controlled for IQ in the full sample (except for four participants for whom we lacked IQ scores) by adding IQ (in addition to motion) as a covariate to the primary analyses. The findings that controlled for IQ were nearly identical to those from the other analyses (Supplementary Materials; Figure S6). Neither the performance-matched nor the IQ-matched lower-income group exhibited greater activation than the higher-income groups as a function of load.

To identify which WM load conditions contributed to group differences in activation, we isolated the MGF and IPS (bilaterally) from the above-reported contrast

Table 3 *Brain activation during working memory*

Region	Hemisphere	Max Coordinates x	y	z	*Z*	# voxels
All students (*n* = 67), 3-back > 2-back > 1-back > 0-back						
Middle & Inferior Frontal Gyri	Left	−49	12	28	4.70	22525
Middle & Inferior Frontal Gyri	Right	46	38	28	5.69	3684
Inferior Parietal Lobule	Left	−40	−45	45	6.53	8279
Inferior Parietal Lobule	Right	44	−45	52	6.94	14657
Caudate, Putamen, Pallidum & Globus Pallidus	Left	−9	7	5	6.88	4123
Caudate, Putamen, Pallidum & Globus Pallidus	Right	13	11	6	6.74	4194
Insula	Left	−30	21	6	5.80	1553
Insula	Right	35	17	2	6.94	579
Occipital Lobe (Calcarine Fissure)	Bilateral	−9	−75	7	5.31	6640
Midbrain & Thalamus		1	−23	−12	5.37	1541
Cerebellum (Culmen & Vermis)		1	−53	−7	3.30	927
Higher income students (*n* = 31), 3-back > 2-back > 1-back > 0-back						
Middle & Inferior Frontal Gyri	Left	−47	−1	44	4.74	3964
Middle & Inferior Frontal Gyri	Right	37	32	29	4.97	6378
Inferior Parietal Lobule	Left	−35	−53	51	4.85	2171
Inferior Parietal Lobule	Right	47	−36	50	5.42	4009
Caudate, Putamen, Pallidum & Globus Pallidus	Left	−14	11	7	5.48	1922
Caudate, Putamen, Pallidum & Globus Pallidus	Right	14	11	7	5.87	2603
Insula	Left	−30	21	8	5.30	905
Insula	Right	31	19	8	6.43	795
Occipital Lobe (Cuneus)	Bilateral	−5	−78	9	4.91	6989
Midbrain & Thalamus		−3	−29	−16	5.82	1231
Cerebellum	Left	−27	−61	−27	5.25	1536
Cerebellum	Right	31	−58	−26	5.29	1628
Lower income students (*n* = 36), 3-back > 2-back > 1-back > 0-back						
Middle, Precentral & Superior Gyri	Left	−26	−1	58	3.41	3485
Middle, Precentral & Superior Gyri	Right	30	2	60	4.91	1955
Medial Frontal & Superior Frontal Gyri	Medial	−1	4	62	5.08	3138
Inferior Parietal Lobule	Left	−44	−47	60	3.85	153
Inferior Parietal Lobule	Left	−40	−45	46	4.02	909
Inferior Parietal Lobule	Right	43	−45	52	4.36	2346
Superior Parietal Lobule	Medial	−1	−68	50	4.19	2170
Higher income > Lower income, 3-back > 2-back > 1-back > 0-back						
Middle Frontal Gyrus	Left	−51	2	44	3.26	1587
Inferior Frontal Gyrus	Left	−50	13	12	2.62	795
Middle and Superior Frontal Gyri	Right	21	5	58	3.31	11037
Inferior Parietal Lobule	Left	−33	−43	39	3.24	1589
Inferior Parietal Lobule	Right	42	−40	39		
Caudate	Left	−16	0	19	3.38	2434
Caudate	Right	14	5	21	3.09	1465
Cerebellum	Left	−25	−50	−38	2.73	8019
Midbrain		3	−30	−16	3.34	5155
Higher income, 3-back > 1-back versus Lower income, 2-back > 1-back						
Middle Frontal Gyrus	Left	−48	1	54	3.38	3708
Middle Frontal Gyrus	Right	27	6	56	3.34	3620
Medial and Superior Frontal Gyri		−3	12	51	3.45	
Inferior Parietal Lobule	Left	−43	−50	43	3.15	6183
Inferior Parietal Lobule	Right	35	−58	38	3.48	1596
Caudate and Putamen	Left	−22	5	9	2.27	2249
Caudate and Putamen	Right	22	16	10	3.56	2315
Cerebellum	Left	−35	−64	−24	3.4	3461
Cerebellum	Right	32	−63	−25	4.64	4303

(3 > 2>1 > 0-back; Figure 3) and extracted parameter estimates for the 0-, 1-, 2- and 3-back conditions as compared to implicit baseline (rest) separately for participants from both higher- and lower-income backgrounds. During all conditions, students from the lower-income group recruited bilateral MFG and IPS (Figure 3; Table 3), whereas students from the higher-income group recruited bilateral MFG and IPS in only the three more demanding conditions. At the behaviorally similar 1-back load, activation in these groups was similar in these brain regions. At the two most demanding loads (2- and 3-back), the higher-income

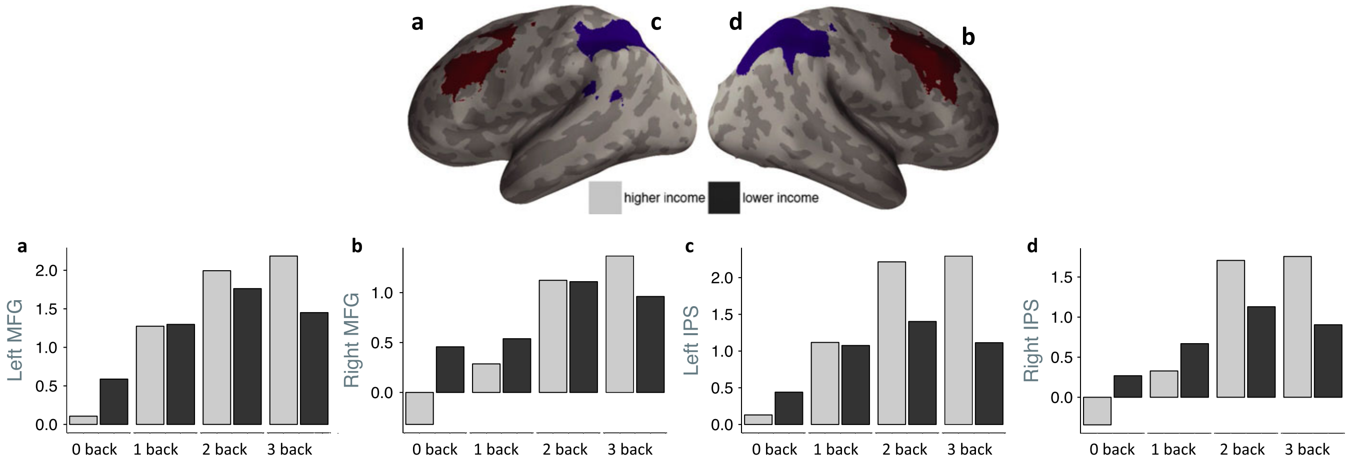

Figure 3 *WM activation in regions of interest during each condition. Mean parameter estimates (z-scores) plotted separately for the higher-income (light gray bars) and lower-income (dark gray bars) groups in (a) left middle frontal gyrus; (b) right middle frontal gyrus; (c) left intraparietal sulcus; and (d) right intraparietal sulcus.*

group exhibited greater activation than the worse performing lower-income group.

Thus, across the least to most demanding WM loads, there was a reversal of relative activation, such that the lower-income group exhibited greater activation at the 0-back load and lesser activation at the 2- and 3-back loads (Figure 4a and b); this difference was significant in bilateral MFG and right IPS (Figure 4c). Subsidiary analyses indicated that these findings held independent of age, gender, race, and ethnicity (Supplemental Material).

Relation of WM activations to Mathematics MCAS scores

Both groups exhibited correlations between activation across the four loads and academic achievement on the Mathematics MCAS (Figure 5). More specifically, both groups exhibited positive correlations between bilateral parietal recruitment and mathematics scores (Figure 5). The lower-income group exhibited negative temporal and inferior frontal correlations, whereas the higher-income group exhibited more extensive positive correlations in bilateral prefrontal cortices and left temporal regions. This pattern was similar in the IQ-matched subset (Supplementary Materials).

Mediation analyses

Given the relationships between income group and both working memory performance and brain activation during working memory (reported above), and the relationship between income group and performance on the Math MCAS test (also reported above), we asked first whether WM ability mediated the relationship between income status and achievement test performance and, second, whether brain activation during WM performance partially mediated the relationship between income status and achievement test performance. A Sobel test revealed that WM performance (average *d* prime; $z = 2.74$, $p = .006$) significantly mediated the relationship between income status and achievement even when controlling for IQ. Thus, including WM performance – in addition to income status (and controlling for IQ) – to predict MCAS performance reduced the income–achievement gap by 26% from 19.63 to 14.49 points. A separate Sobel test additionally revealed that brain activation during working memory (controlling for motion (fdrms) and IQ; $z = 2.933$, $p = .003$) significantly mediated the relationship between income status and achievement. Including WM brain activations, in addition to income status (and controlling for motion and IQ), to predict MCAS mathematics performance therefore reduced the income–achievement gap by 28%, from 19.63 to 14.1 points. As with all mediation analyses, these reductions in the achievement gap could reflect either a direct influence of WM performance and brain activation during WM on achievement, or the influence of unmeasured differences between higher-income and lower-income students that correlated with both achievement and each of these measures.

Discussion

Higher-income students, relative to lower-income students, had higher scores on the statewide mathematics

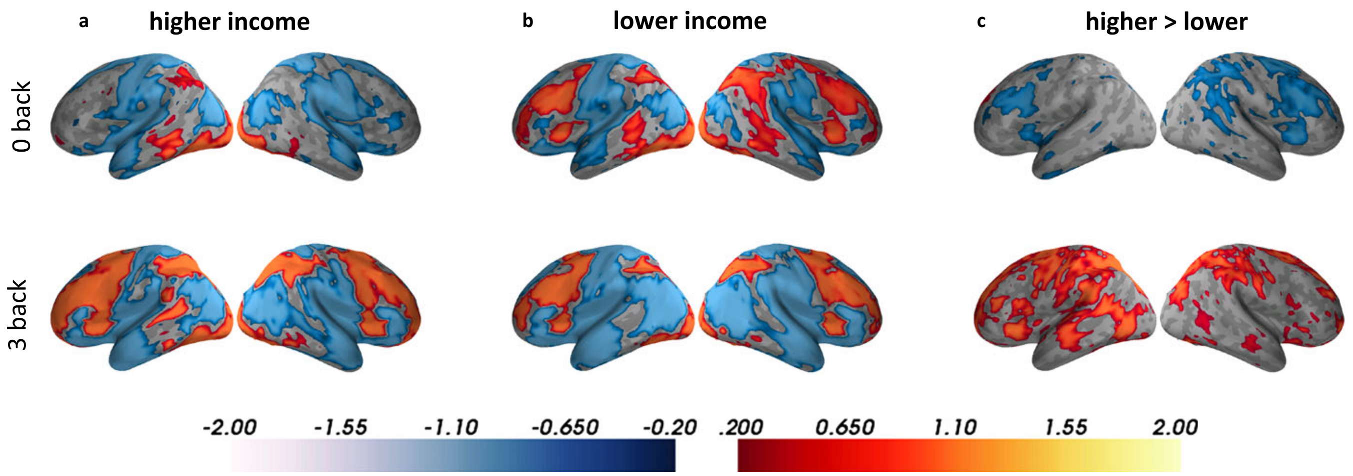

Figure 4 *WM load-related differences in activation during the least and most demanding conditions. Activations for the contrasts 0-back greater than implicit baseline (rest) and 3-back greater than implicit baseline (rest) are displayed for (a) higher-income (b) lower–income groups; (c) differences between the higher- and lower-income groups (higher greater = warm colors, lower greater = cool colors) are displayed for each contrast.*

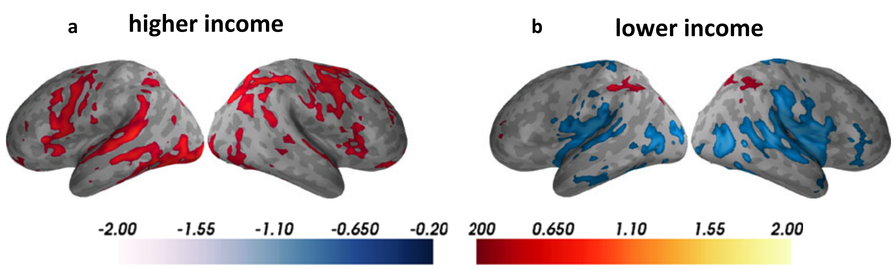

Figure 5 *WM recruitment and MCAS performance. Parametric activation (for the contrast 3 > 2 > 1 > 0-back) correlated positively with MCAS mathematics scores (a) in frontal, parietal, and temporal regions in the higher-income group, and (b) only in parietal regions in the lower-income group.*

achievement test (MCAS), greater working memory capacity by two measures, greater functional brain responses to increasing demands on WM capacity, and differential brain activation correlations with scores on mathematics achievement. These findings indicate that family income is related to WM capacity and its functional neural architecture in adolescents, as well as to performance on a statewide test of academic achievement in mathematics.

Behaviorally, higher income was associated with both greater WM capacity (on both the N-Back and Count Span measures) and higher mathematics achievement test scores. WM capacity is conceptualized as reflecting executive function ability (Miyake, Friedman, Emerson, Witzki, Howerter *et al.*, 2000; Miyake & Shah, 1999) and is related to separate measures of general cognitive ability (indexed by measures like IQ tests (Engle, Tuholski, Laughlin & Conway, 1999; Kane, Hambrick, Tuholski, Wilhelm, Payne *et al.*, 2004; Kyllonen & Christal, 1990; Miyake *et al.*, 2000)). The finding that lower income was associated with lesser WM capacity is consistent with prior findings that lower-SES children and adolescents perform worse on measures of executive function, including spatial working memory and verbal tasks (Farah, Shera, Savage, Betancourt, Giannetta *et al.*, 2006; Hackman *et al.*, 2015; Heckman, 2007; Noble *et al.*, 2005). These findings also contradict the claim that SES does not influence WM capacity (Engel, Santos & Gathercole, 2008). The correlations between both WM capacity measures collected with individualized testing in the laboratory with mathematics MCAS scores is consistent with prior evidence of such a correlation when WM capacity measures were collected in group classroom testing (Finn *et al.*, 2014). Further, the similar SES differences on the out-of-scanner Count Span and in-scanner N-back measures of WM capacity suggest that the measurement of WM during MRI scanning did not interact with group WM capacity differences.

Greater WM demands were, on average, associated with greater activation in brain regions known to underlie WM, but income background varied with the extent to which these regions were recruited with increasing demands. Overall, greater WM loads were associated with greater activation in bilateral DLPFC, parietal cortex, basal ganglia (caudate and putamen), and cerebellar regions. This is consistent with prior studies of WM in children and adolescents (Casey, Cohen, Jezzard, Turner, Noll *et al*., 1995; Finn, Sheridan, Hudson Kam, Hinshaw & D'Esposito, 2010; O'Hare, Lu, Houston, Bookheimer & Sowell, 2008; Thomas, King, Franzen, Welsh, Berkowitz *et al*., 1999; Thomason, Race, Burrows, Whitfield-Gabrieli, Glover *et al*., 2009). The higher-income group exhibited increasing magnitudes of activation from 0-back through 3-back conditions. In contrast, the lower-income group exhibited much less growth in activation across loads. The lower-income group, relative to the higher-income group, exhibited greater activation at the 0-back load, and then lesser activation at the most demanding 2-back and 3-back loads. This was true even in a subset of students who were matched across income groups for behavioral performance or for nonverbal (fluid) IQ and also when IQ was statistically controlled for across all participants.

One explanation for the group differences in brain activation is that there are limited neural resources in the less well performing lower-income group. At the lowest load, WM circuits may have been taxed to a greater extent in the lower-income group, and additional resources were no longer available to support performance at the highest loads. The pattern of similar or even greater activation at a low WM load and lesser activation at a high WM load has been observed previously in other groups with reduced WM capacity and executive functions, such as healthy older relative to healthy younger adults (Mattay, Fera, Tessitore, Hariri, Berman *et al*., 2006), or patients with schizophrenia (Jansma *et al*., 2004). It is possible that chronic exposure to stress or less enriching environments could produce these differences as such exposure (or lack of) has been known to have an impact across the brain and on frontal areas in particular (Sheridan & McLaughlin, 2014).

It is difficult to interpret the meaning of differences in brain activation when there are also related differences in performance, as was the case in this study. The brain activation differences could reveal causes of the WM performance differences, or could instead be a consequence of differences in WM capacity. There are three suggestions that the activation differences were not strictly secondary to performance differences. First, at the 0-back load, the lower-income group exhibited greater MFG and IPS activation than the higher-income group, even though their performance was less accurate. (This also shows that lesser activation was not an inevitable consequence of lower income.) Second, even when performance differences were matched across different WM loads (1-back to 3-back in the higher-income group versus 1-back to 2-back in the lower-income group), there were significant differences in activation in PFC and parietal regions. Third, when subgroups of both groups were selected so that overall performance was not significantly different between groups, there were still significant differences in activation. In addition, the same patterns of findings occurred when the higher- and lower-income groups were equated for nonverbal IQ or when IQ was statistically controlled. Thus, the income-group differences in brain activation were not a direct consequence of better or worse WM performance during scanning.

An apparent paradox is that performance-matched lower-income students had reduced activation in PFC and parietal regions as a function of load. This finding leads one to question what brain regions were supporting strong performance at higher WM loads in those lower-income students. An alternative neural circuitry supporting strong performance was not identified because in no analysis did performance-matched or IQ-matched lower-income students exhibit significantly greater activation than the higher-income students. One possibility is that the lower-income students were heterogeneous as to the brain regions supporting strong performance, and the lack of homogeneity would not support a consistent and statistically significant group difference in any single brain region.

An unexpected finding was that the lower-SES group performed less accurately than the higher-SES group on the 0-back condition, in which there was no WM load and participants simply responded to the target letter 'W'. This raises the concern that the lower-SES group might have been less engaged in task performance. This concern is mitigated by two other observations. First, there was no group difference in the more demanding 1-back condition, which indicates that at least in that condition, engagement was equal in the two groups (and the greater activation in the lower-SES group suggests greater engagement in that group). Second, group differences persisted when groups were equated by performance at different loads or subgroups were selected to equate overall performance (both discrimination sensitivity and reaction time).

The strong correlations between WM capacity, mathematics achievement scores, and the brain organization of WM are consistent with evidence that WM capacity supports and constrains the learning of mathematics. Similar correlations between WM capacity and performance on mathematics achievement tests have been documented in other studies (Finn *et al.*, 2014; Gathercole *et al.*, 2004). Mathematics achievement in particular is related to WM capacity (Dumontheil & Klingberg, 2012; Raghubar *et al.*, 2010), suggesting that WM capacity may be a critical resource for learning mathematics. Although it is difficult to establish causal relations between highly correlated measures of WM and mathematics achievement, the finding that WM performance predicts mathematics performance 2 years later (Dumontheil & Klingberg, 2012) suggests that WM may be an important resource for learning in mathematics. It is difficult to know, however, whether WM capacity per se is constraining achievement in mathematics, because WM capacity is highly correlated with other broad measures of cognitive ability (Engle *et al.*, 1999; Kane *et al.*, 2004; Kyllonen & Christal, 1990; Miyake *et al.*, 2000).

Correlations between WM brain activations and mathematics achievement scores were similar in the parietal lobes for the two groups, but dissimilar in PFC. Both income groups exhibited a positive correlation between WM activation in the bilateral IPS, a brain region associated with numerical operations (Grabner, Ansari, Reishofer, Stern, Ebner *et al.*, 2007; Hubbard, Piazza, Pinel & Dehaene, 2005), and the amount of information being stored in working memory (McNab & Klingberg, 2008; Vogel, McCollough & Machizawa, 2005). This suggests that greater activation in parietal cortices on the WM task indexed neural processes useful for performing mathematical operations.

The higher-income group uniquely exhibited significant positive correlations between WM activation and mathematics scores in PFC. The PFC has been associated generally with complex reasoning and rule use (Badre & D'Esposito, 2007; Crone, Wendelken, van Leijenhorst, Honomichl, Christoff *et al.*, 2009; Ferrer, O'Hare & Bunge, 2009) and specifically with complex mathematics problems that require multiple arithmetic operations (Prabhakaran, Rypma & Gabrieli, 2001). For students who performed less well as a group on the more difficult problems, there might be less variation to relate between PFC functions and mathematics performance. For students who performed better as a group on the more difficult problems, which demanded more reasoning and arithmetic operations, there may have been greater variation to relate between PFC functions and mathematics performance.

The mediation analysis indicated that the relations between income status and achievement in math were at least partially mediated by both WM capacity and brain activation during WM even when controlling for IQ. These findings are consistent with a longitudinal behavioral study of a much larger cohort of students, which reported that executive functions, mainly measured by WM capacity, partially mediated the relation between SES and longitudinal change in math achievement (Lawson & Farah, 2015). The present study replicates the relations among SES, WM, and math achievement, and extends such findings to brain functions.

The present study has a few limitations. First, the measurement of family income, whether students qualify for free or reduced lunch, was somewhat imprecise. While this measure is accurate (we were able to determine whether students in this study qualified for 3 years leading up to participation), it does not stipulate the actual income of families and somewhat artificially binarizes students into lower- and higher-income groups. Future studies should measure parental income and education and use this information as a continuous variable. Further, although controlling for race and ethnicity did not alter the pattern of reported findings (Supplement), the groups of students in the present study were not matched in terms of racial and ethnic demographics (as occurs generally in relation to SES in the United States). Similar relations between SES and executive functions, including WM, have been found when income groups were equated for racial membership (Farah *et al.*, 2006; Noble *et al.*, 2005).

In addition, there were differences in the years of schooling for students from higher- and lower-income backgrounds in this study. Although the groups did not differ in terms of age, four of the lower-income students were in 7th grade (and not 8th grade as the rest were) at the time of scanning (and seven of the lower-income students were in 7th grade at the time of taking the MCAS test). Some of the lower-income students reported repeating at least one grade, and none of the higher-income students reported such retention. Although controlling for grade at the time of testing did not alter the results (Supplementary Figure S1), these variables need careful measurement in future studies, especially if students have repeated a grade, in order to determine whether and how number of years in school might influence this pattern of data.

Although this study documented functional brain differences related to family income at around age 14, there is considerable evidence that substantial brain plasticity occurs at this and older ages. Training on WM and other executive functions, such as selective attention, alters performance, evoked brain activation, the connectivity of intrinsic brain networks and white matter pathways, and the timing of brain responses as measured with event-related potentials across as little as 4 to 5 weeks of training (Kundu, Sutterer, Emrich & Postle, 2013; Olesen, Westerberg & Klingberg, 2004; Takeuchi, Sekiguchi, Taki, Yokoyama, Yomogida *et al.*, 2010;

Takeuchi, Taki, Nouchi, Hashizume, Sekiguchi *et al*., 2013; Westerberg & Klingberg, 2007; Zhao, Zhou & Fu, 2013) and even across two sessions with direct neural feedback in the scanner (Zhang, Yao, Zhang, Long & Zhao, 2013).

The achievement gap associated with family income is of concern to educators who want students to thrive academically regardless of the environment into which they were born. The present study revealed one functional brain basis for that gap, expressed as an alteration in the functional neural architecture supporting a greater WM capacity that was also associated with achievement in mathematics. Future research can examine whether interventions focused on WM and its underlying neural circuitry could boost the achievement of lower-income students and reduce that gap.

Acknowledgements

This research was funded by the Bill & Melinda Gates Foundation (to J.D.E. and C.F.O. Gabrieli) and the National Institutes of Health (NRSA to ASF). We thank Brett Alessi, C.J. El-Dahr, Margaret Sheridan, Jon Fullerton, Jared Silver, Brian Chan, and Micah Nishigaki for their help. We thank the students and their parents for their participation, and the MATCH Charter Public Middle School for helping recruit participants. We also thank the staff of the Athinoula A. Martinos Imaging Center at the McGovern Institute for Brain Research, MIT.

References

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A. *et al*. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, **54**, 2033–2044.

Awh, E., & Vogel, E.K. (2008). The bouncer in the brain. *Nature Neuroscience*, **11**, 5–6.

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience*, **19**, 2082–2099.

Brooks-Gunn, J., Guo, G., & Furstenberg, F.F. (1993). Who drops out of and who continues beyond high school? A 20-year follow-up of black urban youth. *Journal of Research on Adolescence*, **3**, 271–294.

Brown, L., Sherbenou, R.J., & Johnsen, S.K. (2010). *Test of Nonverbal Intelligence – Fourth edition*. San Antonio, TX: The Psychological Corp.

Casey, B.J., Cohen, J.D., Jezzard, P., Turner, R., Noll, D.C. *et al*. (1995). Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *NeuroImage*, **2**, 221–229.

Conway, A.R.A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O. *et al*. (2005). Working memory span tasks: a methodological review and user's guide. *Psychonomic Bullitin and Review*, **12**, 769–786.

Conway, A.R.A., Kane, M.J., & Engle, R.W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences*, **7**, 547–552.

Cowan, N., Elliott, E.M., Saults, S.J., Morey, C.C., Mattox, S. *et al*. (2005). On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology*, **51**, 42–100.

Crone, E.A., Wendelken, C., van Leijenhorst, L., Honomichl, R.D., Christoff, K. *et al*. (2009). Neurocognitive development of relational reasoning. *Developmental Science*, **12**, 55–66.

Cubillo, A., Smith, A.B., Barrett, N., Giampietro, V., Brammer, M. *et al*. (2014). Drug-specific laterality effects on frontal lobe activation of atomoxetine and methylphenidate in attention deficit hyperactivity disorder boys during working memory. *Psychological Medicine*, **44**, 633–646.

Curtis, C.E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences*, **7**, 415–423.

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, **19**, 450–466.

Dumontheil, I., & Klingberg, T. (2012). Brain activity during a visuospatial working memory task predicts arithmetical performance 2 years later. *Cereberal Cortex*, **22**, 1078–1085.

Duncan, G.J., Yeung, W.J., Brooks-Gunn, J., & Smith, J.R. (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review*, **63**, 406–423.

Engel, P.M.J., Santos, F.H., & Gathercole, S.E. (2008). Are working memory measures free of socioeconomic influence? *Journal of Speech, Language, and Hearing Research*, **51**, 1580–1587.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, **128**, 309–331.

Farah, M.J., Shera, D.M., Savage, J.H., Betancourt, L., Giannetta, J.M. *et al*. (2006). Childhood poverty: specific associations with neurocognitive development. *Brain Research*, **1110**, 166–174.

Ferrer, E., O'Hare, E.D., & Bunge, S.A. (2009). Fluid reasoning and the developing brain. *Frontiers in Neuroscience*, **5**, 46–51.

Finn, A.S., Kraft, M.A., West, M.R., Leonard, J.A., Bish, C.E. *et al*. (2014). Cognitive skills, student achievement tests, and schools. *Psychological Science*, **25**, 736–744.

Finn, A.S., Sheridan, M.A., Hudson Kam, C.L., Hinshaw, S., & D'Esposito, M. (2010). Longitudinal evidence for functional specialization of the neural circuit supporting working memory in the human brain. *Journal of Neuroscience*, **30**, 11062–11067.

Gathercole, S.E., Pickering, S.J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: evidence from national curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, **18**, 1–16.

Ghosh, S., Keshavan, A., Salvatore, J., & Klein, A. (2012). BIPS: a framework for curating and executing brain imaging pipelines. Paper presented at the 5th INCF Congress of Neuroinformatics.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum (Ed.), *Handbook of physiology, the nervous system, higher functions of the brain* (pp. 373–417). Bethesda, MD: American Physiological Society.

Gorgolewski, K., Burns, C.D., Madison, C., Clark, D., Halchenko, Y.O. *et al*. (2011). Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. *Frontiers in Neuroinformatics*, **5**, 13.

Grabner, R.H., Ansari, D., Reishofer, G., Stern, E., Ebner, F. *et al*. (2007). Individual differences in mathematical competence predict parietal brain activation during mental calculation. *NeuroImage*, **38**, 346–356.

Greve, D.N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *NeuroImage*, **48**, 63–72.

Hackman, D.A., Gallop, R., Evans, G.W., & Farah, M.J. (2015). Socioeconomic status and executive function: developmental trajectories and mediation. *Developmental Science*, **18**, 686–702.

Hanson, J.L., Chandra, A., Wolfe, B.L., & Pollak, S.D. (2011). Association between income and the hippocampus. *PLoS ONE*, **6**, e18712.

Hanson, J.L., Hair, N., Shen, D.G., Shi, F., Gilmore, J.H. *et al*. (2013). Family poverty affects the rate of human infant brain growth. *PLoS ONE*, **8**, e80954.

Heckman, J.J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the National Academy of Sciences, USA*, **104**, 13250–13255.

Hubbard, E.M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience*, **6**, 435–448.

Jansma, J.M., Ramsey, N.F., van der Wee, N.J., & Kahn, R.S. (2004). Working memory capacity in schizophrenia: a parametric fMRI study. *Schizophrenia Research*, **68**, 159–171.

Jednoróg, K., Altarelli, I., Monzalvo, K., Fluss, J., Dubois, J. *et al*. (2012). The influence of socioeconomic status on children's brain structure. *PLoS ONE*, **7**, e42486.

Jenkinson, M. (2003). Measuring transformation error by rms deviation from http://www.fmrib.ox.ac.uk/analysis/techrep/tr99mj1/tr99mj1/index.html

Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., & Smith, S.M. (2012). FSL. *NeuroImage*, **62**, 782–790.

Kane, M.J., Hambrick, D.Z., Tuholski, S.W., Wilhelm, O., Payne, T.W. *et al*. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, **133**, 189–217.

Kennedy, D., & Haselgrove, C. (2011). *Harvard-Oxford atlas*. From: http://neuro.imm.dtu.dk/wiki/Harvard-Oxford_Atlas

Kishiyama, M.M., Boyce, W.T., Jimenez, A.M., Perry, L.M., & Knight, R.T. (2009). Socioeconomic disparities affect prefrontal function in children. *Journal of Cognitive Neuroscience*, **21**, 1106–1115.

Kundu, B., Sutterer, D.W., Emrich, S.M., & Postle, B.R. (2013). Strengthened effective connectivity underlies transfer of working memory training to tests of short-term memory and attention. *Journal of Neuroscience*, **33**, 8705–8715.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, **14**, 389–433.

Lawson, G.M., Duda, J.T., Avants, B.B., Wu, J., & Farah, M.J. (2013). Associations between children's socioeconomic status and prefrontal cortical thickness. *Developmental Science*, **16**, 641–652.

Lawson, G.M., & Farah, M.J. (2015). Executive function as a mediator between SES and academic achievement throughout childhood. *International Journal of Behavioral Development*. doi: 10.1177/0165025415603489

LeBlanc, M.D., & Weber-Russell, S. (1996). Text integration and mathematical connections: a computer model of arithmetic word problem solving. *Cognitive Science*, **20**, 357–407.

Luby, J., Belden, A., Botteron, K., Marrus, N., Harms, M.P. *et al*. (2013). The effects of poverty on childhood brain development: the mediating effect of caregiving and stressful life events. *JAMA Pediatrics*, **167**, 1135–1142.

Mackey, A.P., Finn, A.S., Leonard, J.A., Jacoby-Senghor, D.S., West, M.R. *et al*. (2015). Neuroanatomical correlates of the income–achievement gap. *Psychologial Science*, **26**, 925–933.

McNab, F., & Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience*, **11**, 103–107.

Mattay, V.S., Fera, F., Tessitore, A., Hariri, A.R., Berman, K.F. *et al*. (2006). Neurophysiological correlates of age-related changes in working memory capacity. *Neuroscience Letters*, **392**, 32–37.

Miyake, A., Friedman, N.P., Emerson, M.J., Witzki, A.H., Howerter, A. *et al*. (2000). The unity and diversity of executive functions and their contributions to complex 'frontal lobe' tasks: a latent variable analysis. *Cognitive Psychology*, **41**, 49–100.

Miyake, A., & Shah, P.E. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Noble, K.G., Houston, S.M., Brito, N.H., Bartsch, H., Kan, E. *et al*. (2015). Family income, parental education and brain structure in children and adolescents. *Nature Neuroscience*, **18**, 773–778.

Noble, K.G., Houston, S.M., Kan, E., & Sowell, E.R. (2012). Neural correlates of socioeconomic status in the developing human brain. *Developmental Science*, **15**, 516–527.

Noble, K.G., Norman, M.F., & Farah, M.J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, **8**, 74–87.

O'Hare, E.D., Lu, L.H., Houston, S.M., Bookheimer, S.Y., & Sowell, E.R. (2008). Neurodevelopmental changes in verbal working memory load-dependency: an fMRI investigation. *NeuroImage*, **42**, 1678–1685.

Olesen, P.J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, **7**, 75–79.

Owen, A.M., McMillan, K.M., Laird, A.R., & Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, **25**, 46–59.

Passolunghi, M.C., & Siegel, L.S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, **80**, 44–57.

Peirce, J. (2007). PsychoPy: psychophysics software in Python. *Journal of Neuroscience Methods*, **162**, 8–13.

Perlstein, W.M., Carter, C.S., Noll, D.C., & Cohen, J.D. (2001). Relation of prefrontal cortex dysfunction to working memory and symptoms in schizophrenia. *American Journal of Psychiatry*, **158**, 1105–1113.

Prabhakaran, V., Rypma, B., & Gabrieli, J.D. (2001). Neural substrates of mathematical reasoning: a functional magnetic resonance imaging study of neocortical activation during performance of the necessary arithmetic operations test. *Neuropsychology*, **15**, 115–127.

Raghubar, K.P., Barnes, M.A., & Hecht, S.A. (2010). Working memory and mathematics: a review of developmental, individual difference, and cognitive approaches. *Learning & Individual Differences*, **20**, 110–122.

Raizada, R.D., Richards, T.L., Meltzoff, A., & Kuhl, P.K. (2008). Socioeconomic status predicts hemispheric specialisation of the left inferior frontal gyrus in young children. *NeuroImage*, **40**, 1392–1401.

Sheridan, M.A., & McLaughlin, K.A. (2014). Dimensions of early experience and neural development: deprivation and threat. *Trends in Cognitive Sciences*, **18**, 580–585.

Sheridan, M.A., Sarsour, K., Jutte, D., D'Esposito, M., & Boyce, W.T. (2012). The impact of social disparity on prefrontal function in childhood. *PLoS ONE*, **7**, e35744.

Smith, S.M., & Nichols, T.E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*, **44**, 83–98.

Sobel, M.E. (1982). Asymptotic intervals for indirect effects in structural equation models. In S. Leinhart (Ed.), *Sociological methodology* (pp. 290–312). San Francisco, CA: Jossey-Bass.

Stevens, C., Lauinger, B., & Neville, H. (2009). Differences in the neural mechanisms of selective attention in children from different socioeconomic backgrounds: an event-related brain potential study. *Developmental Science*, **12**, 634–646.

Swanson, H.L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, **96**, 471–491.

Takeuchi, H., Sekiguchi, A., Taki, Y., Yokoyama, S., Yomogida, Y. *et al*. (2010). Training of working memory impacts structural connectivity. *Journal of Neuroscience*, **30**, 3297–3303.

Takeuchi, H., Taki, Y., Nouchi, R., Hashizume, H., Sekiguchi, A. *et al*. (2013). Effects of working memory training on functional connectivity and cerebral blood flow during rest. *Cortex*, **49**, 2106–2125.

Thomas, K.M., King, S.W., Franzen, P.L., Welsh, T.F., Berkowitz, A.L. *et al*. (1999). A developmental functional MRI study of spatial working memory. *NeuroImage*, **10**, 327–338.

Thomason, M., Race, E., Burrows, B., Whitfield-Gabrieli, S., Glover, G. *et al*. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognitive Neuroscience*, **21**, 316–332.

Vogel, E.K., McCollough, A.W., & Machizawa, M.G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, **438**, 500–503.

Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory: a single-subject analysis. *Physiology and Behavior*, **92**, 186–192.

Woolrich, M.W., Behrens, T.E.J., Beckmann, C.F., Jenkinson, M., & Smith, S.M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *NeuroImage*, **21**, 1732–1747.

Zhang, G., Yao, L., Zhang, H., Long, Z., & Zhao, X. (2013). Improved working memory performance through self-regulation of dorsal lateral prefrontal cortex activation using real-time fMRI. *PLoS ONE*, **8**, e73735.

Zhao, X., Zhou, R., & Fu, L. (2013). Working memory updating function training influenced brain activity. *PLoS ONE*, **8**, e71063.

Zheng, X., Swanson, H.L., & Marcoulides, G.A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, **110**, 481–498.

Received: 10 March 2015
Accepted: 30 March 2016

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Method.