Developmental changes in category-specific brain responses to numbers and letters in a working memory task

Melissa E. Libertus a,b,⁎, Elizabeth M. Brannon a,b, Kevin A. Pelphrey c

a Department of Psychology and Neuroscience, Duke University, USA
b Center for Cognitive Neuroscience, Duke University, USA
c Yale Child Study Center, Yale University, USA

A R T I C L E   I N F O
Article history:
Received 29 June 2008
Revised 11 September 2008
Accepted 16 October 2008
Available online 5 November 2008

A large body of research on the neural substrate supporting working memory (WM) in adults implicates a network of parietal and prefrontal brain regions (Cabeza and Nyberg, 2000). An important question is whether the type of information to be retained affects which regions within this network are recruited. In adults there is evidence that object and spatial WM tasks recruit distinct neural networks. Whereas spatial WM recruits the parietal, dorsal visual stream and the dorsal prefrontal cortex, object WM – specifically for faces – relies on the temporal, ventral visual stream and ventral prefrontal cortical regions (Belger et al., 1998; Courtney et al., 1996; Haxby et al., 2000; Ungerleider et al., 1998). For example, activity in the fusiform face area and prefrontal cortex was found to be modulated by the memory load in a face WM task (Druzgal and D’Esposito, 2001, 2003). Verbal and non-verbal WM have also been found to elicit differential activation in the left and the right hemisphere, respectively (Fiez et al., 1996). Furthermore, Knops et al. (2006) found a further subdivision within verbal WM. Using an n-back paradigm with numbers and letters in adults, they found enhanced activation of the horizontal segment of the intraparietal sulcus (a region that is commonly engaged by numerical processing tasks) in a numerical n-back task as compared to a letter n-back condition. The authors suggested that numerical information is accessed automatically and that a numerical n-back task might not assess purely verbal WM components.

In general, structural neuroimaging studies show that regions of parietal and especially frontal cortex undergo changes well into adolescence and early adulthood. For example, cortical thickness increases in these brain areas between around 5 and 10 years of age and then decreases during adolescence until it stabilizes in adulthood (Shaw et al., 2008; Sowell et al., 1999). Furthermore, gray matter in frontal and parietal cortex declines – most likely due to increases in white matter and myelination – between childhood and adolescence (Sowell et al., 2004) and continues to decline even between adolescence and early adulthood (Sowell et al., 1999). These findings therefore suggest that the brain systems underlying WM and thus WM performance undergo changes well into adolescence and even early adulthood.

⁎ Corresponding author. Center for Cognitive Neuroscience Levine Science Research Center, Rm B203, Box 90999, Duke University, Durham, NC 27708, USA. Fax: +1 919 681 0815.
E-mail address: melissa.libertus@duke.edu (M.E. Libertus).

Available online 5 November 2008

© 2008 Elsevier Inc. All rights reserved.

A B S T R A C T

Neuroimaging studies have identified a common network of brain regions involving the prefrontal and parietal cortices across a variety of working memory (WM) tasks. However, previous studies have also reported category-specific dissociations of activation within this network. In this study, we investigated the development of category-specific activation in a WM task with digits, letters, and faces. Eight-year-old children and adults performed a 2-back WM task while their brain activity was measured using functional magnetic resonance imaging (fMRI). Overall, children were significantly slower and less accurate than adults on all three WM conditions (digits, letters, and faces); however, within each age group, behavioral performance across the three conditions was very similar. FMRI results revealed category-specific activation in adults but not children in the intraparietal sulcus for the digit condition. Likewise, during the letter condition, category-specific activation was observed in adults but not children in the left occipital–temporal cortex. In contrast, children and adults showed highly similar brain-activity patterns in the lateral fusiform gyri when solving the 2-back WM task with face stimuli. Our results suggest that 8-year-old children do not yet engage the typical brain regions that have been associated with abstract or semantic processing of numerical symbols and letters when these processes are task-irrelevant and the primary task is demanding. Nevertheless, brain activity in letter-responsive areas predicted children’s spelling performance underscoring the relationship between abstract processing of letters and linguistic abilities. Lastly, behavioral performance on the WM task was predictive of math and language abilities highlighting the connection between WM and other cognitive abilities in development.
Previous research on WM in children has shown that similar brain regions are activated when children and adults retain information in WM (e.g. Casey et al., 1995; Kwon et al., 2002). However, in comparison to adults, children show greater reliance on posterior brain regions that mature earlier than prefrontal regions in childhood (Bunge et al., 2002; Ciesielski et al., 2004, 2006) and their brain activity appears to be more diffuse during WM tasks, becoming more focal with age (Casey et al., 1995, 2000; Klingberg et al., 2002; Kwon et al., 2002; Nelson et al., 2000). Furthermore, activation in the fronto-parietal WM network increases with age and improvement in WM (Crone et al., 2006; Kwon et al., 2002). It is not yet known whether children, like adults, recruit specialized brain regions when remembering different types of information in WM. Here we sought to address this issue.

Behaviorally, WM undergoes substantial growth over the course of development. For example, counting and listening span increase from 7–13 years of age (Siegel and Ryan, 1989), and digit span increases from 4–5 items at 4–5 years of age to 5–7 items at 14–15 years of age (Conklin et al., 2007; Pickering, 2006). On verbal and spatial self-ordered search tasks that require strategy generation and constant monitoring of responses, the number of errors that children make decreases with age even well into adolescence (Conklin et al., 2007). At least two hypotheses have been put forward to explain the dramatic increase in WM span during childhood and adolescence: 1) greater efficiency of rehearsal strategies that can be viewed as domain-general changes in executive control (Case et al., 1982; Swanson, 1999; Towsie and Hitch, 2007), and 2) greater efficiency of basic processes necessary for WM, i.e. domain-specific changes (Hulme and Mackenzie, 1992). The latter proposes mechanisms including increases in the speeds of object identification and articulation over the course of development. These increases in turn lead to increases in the amount of information that can be stored and rehearsed in each of the subsystems of WM (Hitch et al., 1989).

Another important question concerns how individual differences in working memory during childhood relate to differences in other cognitive abilities. For example, both reading and math involve WM. Reading involves rapid orthographic, phonological, contextual, and semantic processing in which different pieces of information have to be held in memory for later processing steps (Adams, 1990; de Jong, 2006). While phonological awareness (the sensitivity to the particular sound structure of a language) is an important factor in learning to read (see Ziegler and Goswami, 2005, for a review), verbal WM also plays an important role both during early reading acquisition as well as reading comprehension in adulthood (Daneman and Carpenter, 1980; Hitch et al., 2001; Just and Carpenter, 1992; Rohl and Pratt, 1995). WM deficits have been linked to poor reading skills in both children and adults (Chiatte et al., 2000; Gathercole et al., 2006), and WM is also related to spelling performance, especially learning to spell (Ormrod and Cochran, 1988).

Furthermore, WM is essential for many forms of numerical processing but especially for mental arithmetic in which different arithmetic operations and intermediate results have to be held on-line to arrive at the correct final solution (e.g. Furst and Hitch, 2000; Lee and Kang, 2002; Lemmaitre et al., 1996; Logie et al., 1994; Siegler and Shrager, 1984). A number of studies have found correlations between math disabilities and poor WM skills suggesting that math disabilities might be due to poor WM performance (e.g. Bull et al., 1999; Geary et al., 1991; Hitch and McAuley, 1991; Passolunghi and Siegel, 2001). However, there is controversy over whether children with math disabilities suffer from general WM impairments or whether these WM deficits are specific to numerical information (e.g. Wilson and Swanson, 2001). In past behavioral studies, this question has been addressed using various WM span tasks (digit, letter, reading, listening span). We hypothesized that individual differences in domain-specific activation patterns in the WM system and behavioral performance on the n-back WM task would be related to individual differences in math and linguistic abilities.

In the present fMRI study, we tested 8-year-old children and adults on a 2-back WM task with three different stimulus conditions (numbers, letters, and face images) while their brain activity was measured using functional magnetic resonance imaging (fMRI). We chose to test 8-year-olds because previous behavioral studies investigating the relationship between WM and other cognitive functions have focused on children at this age (e.g. Swanson and Beebe-Frankenberger, 2004) and initial pilot testing showed that children at this age are capable of performing our WM task successfully. In addition, we tested children’s math and linguistic abilities as well as their performance on a standardized measure of intelligence. This design allowed us to compare the neural systems that children and adults engage to accomplish WM tasks involving different stimulus categories and to evaluate a hypothesized relationship between activation in components of the WM system and other cognitive abilities during childhood.

Methods

Participants

Fifteen normally developing 8-year-old children participated (7 girls, mean age: 8 years 8 months; range: 8 years 11 days–9 years 1 month at the time of the fMRI scan). The adult group consisted of fifteen healthy adults (8 female, mean age: 25 years; range 20–35 years). All participants were recruited from the local community near Duke University. Adult participants and parents of the child participants gave informed consent based on a protocol approved by the Duke University Institutional Review Board.

Behavioral tasks

Dyscalculia and dyslexia screeners for children

Children completed the nferNelson Dyscalculia screener (Butterworth, 2003). For this age group, the screener assesses general response time speed, counting abilities, number comparison abilities, and performance on simple addition tasks in the following four subsections of the test: (1) to assess the children’s general response speed for any task, the Simple Reaction Time subtest required the children to press a left or right keyboard button as fast as possible when a dot appeared on the screen. (2) For Dot Enumeration, children saw a set of between one and nine dots on the left side of the screen and an Arabic numeral between one and nine on the right. They clicked on a left key (“no”) to indicate that the number of dots and the number did not match or a right key (“yes”) to indicate that they did match. (3) Numerical Stroop showed children two Arabic numerals (of varying physical size) and had children click on either a right or left key to indicate which number was greater in numerical value. On some trials, physical size and numerical size were congruent; on others, they were incongruent. (4) The Addition subtest showed children a completed addition problem (like “2+3=5”) and had them respond either by pressing a left key that the solution to the addition problem was not correct (“no”) or a right key that the problem was correct (“yes”).

Children also completed the nferNelson Dyslexia screener (Turner and Smith, 2004), which screens for potential literacy problems. At this age, the dyslexia screener includes six subtests, two that assess general problem solving abilities (Missing Pieces and Vocabulary), two tests to assess information processing efficiency (Visual Search and Word Sounds), and two attainment tests (Reading and Spelling): (1) in Missing Pieces, the children determined which option shown at the bottom of the screen best completed the empty space in a design presented above. This subtest measured children’s understanding of patterns and relationships between shapes presented in a design. (2) In the Vocabulary subtest, children indicated which of four photographs best illustrated the meaning of a spoken word. (3) Visual
was a speeded response task in which the children saw a design at the
top of the screen and they clicked the matching shape from the six
designs shown at the bottom as fast as possible. (4) Word Sounds had
four parts. First, children heard a letter and clicked on which of four
drawings depicted a word starting with that letter; then, they chose the
corresponding word to a word that rhymed with the word they heard.
Next, they chose the picture of the word formed when part of a longer
word was omitted. And, lastly, they indicated the number of syllables in
spoken real and pseudo-words. (5) In the Reading subtest, children first
had to select which of three written words corresponded to the word
read aloud. They then chose which of three words best completed a
written sentence. (6) Lastly, in the Spelling subtest, children chose which
of three written words gave the correct spelling of the word they heard.

Letter and digit span task

Participants were presented with streams of digits or letters (1 per
second) played from an mp3-player and projected by speakers in a
quiet testing room. The task was divided into four conditions. First,
they completed the “forward digits” portion of the task in which they
verbally repeated the digits back to the experimenter following each
stream in the same order in which they were heard. Then, in the
“backward digits” portion, the participants repeated the digits in the
reverse order in which they were heard. Next, the participant
completed the “forward letters” and “backward letters” subtasks in
the same manner as with digits. The first two streams were of length
two and the streams increased in increments of one digit or one letter
with two streams played consecutively of each length. Each condition
ended the first time a participant made errors on both streams of the
same length.

FMRI task

2-back WM task

During the scanning session, adult and child participants were
tested on the same 2-back WM task including three different stimu-
lus types: letters, digits, and images of faces. We used nine different
letters (A, E, K, M, O, R, T, X, and Z), nine digits (1–9), and nine different
emotionally neutral faces from the NimStim face set (Tottenham et al.,
in press). Digits and letters varied in color (red, blue, or green) and
type face (Times New Roman or Arial) to provide some perceptual
variation and help children to stay more engaged. Five of the face
images were female faces, four were male, and ethnicity varied
(5 European-American faces, 2 African-American, 1 Latino-American,
and 1 Asian-American).

Participants were presented with a stream of stimuli, each being
shown for 1500 ms at the center of the screen and separated by an
average inter-stimulus interval (ISI) of 2000 ms (range: 1750–
2250 ms). Starting with the third stimulus, participants were asked
to respond by pressing one of two buttons indicating whether the
target stimulus matched or did not match the stimulus presented two
images before the target. They pushed one button to indicate “same”
and another button for “different”. As illustrated in Fig. 1, stimuli were
presented in blocks of 14 images for each stimulus type requiring
twelve responses equally balanced between “same” and “different”
responses. Following each block, a reward screen was shown for 3 s to
help keep participants engaged. The reward screen consisted of a
ClipArt image and the words “Good work! You earned X points!!”
whereby X was a predetermined number consistent across partici-
pants that was not linked to their actual performance. The reward
screen was followed by a 12-second fixation cross. There were a total
of three runs, each contained six blocks, two for each stimulus type.
Thus, each run lasted about 6 min.

Procedure

Child participants

Children participated in three separate sessions. During the first
visit, children completed the nferNelson Dyscalculia screener (Butter-
worth, 2003) and the nferNelson Dyslexia screener (Turner and Smith,
2004). During the second visit, they participated in the Wechsler
Abbreviated Scale of Intelligence (WASI; 1999) and the letter and digit
span task to assess their WM. Finally, they were familiarized with the
scanning environment on a mock MRI scanner and the 2-back WM
task for the scanning session. At the final visit, they performed the 2-
back WM task in the actual MRI system.

Adult participants

Adults participated in only one visit during which they completed
the forward and backward letter and digit span task outside the
scanner and the 2-back WM task inside the scanner.

Image acquisition

MRI data were acquired on a 3-Tesla General Electric Signa Excite
scanner at the Brain Imaging and Analysis Center at Duke University.
An echo-planar imaging pulse sequence was employed to detect BOLD
T2* contrast (TR=2000 ms, TE=27 ms, flip angle=60°, FOV=25.6 cm,
matrix=64×64, slice thickness=4 mm). There were three functional
runs per session and 200 volumes were collected in each run. High-
resolution structural T1 contrast images were acquired at the
beginning of each session (TR=7.3 ms, TE=2.9 ms, flip angle=12°,
FOV=25.6 cm, matrix=256×256, slice thickness=2 mm). During
scanning sessions with children, one experimenter remained with
the child in the scanner room.

Data analysis

Analysis of the behavioral data

For the letter and digit span task, four scores were computed for
each participant, one for each condition (forward or backward, letter
or digit). Scores were the length of the longest stream repeated

Fig. 1. Experimental design of 2-back WM task. Participants were presented with blocks of 14 digits, letters or face images (stimulus duration: 1500 ms, inter-stimulus interval: 1750–
2250 ms). Starting with the third stimulus, participants indicated whether the image was the same as two items before (Yes/No). At the end of each block, a reward screen was
presented that was independent of the actual task performance.
correctly (at least once) in each condition. Reaction time and accuracy data from the 2-back WM task were compared across age groups and stimulus conditions. Additionally, we examined the correlation between WM span and accuracy on the 2-back WM task. Lastly, accuracies on the three stimulus conditions of the 2-back WM task were used as predictors in stepwise multiple regression models to examine the best predictors of IQ and performance on the various subtests of the Dyslexia and Dyscalculia screeners.

**Analysis of the fMRI data**

The fMRI data for one adult was excluded from the analysis due to technical difficulties during data acquisition, and the fMRI data sets from two children were excluded from the analyses due to excessive movement (>2 mm in any direction) during all three of their functional runs. Partial data from two other children was excluded due to excessive motion in one (n=1) or two (n=1) of the functional runs. Images were processed and analyzed in SPM2 (http://www.fil.ion.ucl.ac.uk/spm). All volumes were 1) corrected for slice acquisition timing, 2) spatially aligned to the Montreal Neurological Institute standard template at a resolution of 2 x 2 x 2 mm. Data for each participant were first analyzed with a fixed effects model to produce contrast images, which were then employed for group analyses using a random-effects approach.

The linear model applied to the images included the SPM2 standard hemodynamic response function (HRF) convolved with block onsets and block length, separately for each stimulus type (letter, digit or face WM), a temporal derivative parameter, and six motion parameters. Fixation periods and reward screens were not included in the model. The contrasts of greatest relevance to us were digit WM vs. letter and face WM (digit-specific WM), letter WM vs. digit and face WM (letter-specific WM), and face WM vs. digit and letter WM (face-specific WM). Group-level contrast thresholds were initially set to p < 0.01 (uncorrected) on the voxel-level and p < 0.05 on the cluster-level and a cluster size of 8 or more contiguous voxels. Interaction effects between age and WM condition were analyzed more conservatively with a threshold of p < 0.005 (uncorrected) at the voxel-level and p < 0.05 on the cluster-level and a cluster size of 8 or more contiguous voxels.

In addition, we examined whether category-specific WM activity predicts performance on any of the subtests of the Dyscalculia and Dyslexia screener, IQ, WM span, and/or performance on the 2-back WM task itself. To this end, we extracted the mean activity of the 10 most strongly activated voxels in a 15 mm sphere around the peak activation for each stimulus-specific region. We restricted our analyses to regions of interest that showed the strongest category-specificity for each WM condition. These mean activation values were then entered as predictors into a stepwise multiple regression to find the best model that predicted each of the behavioral measures of interest. For children, we extracted peak activations both from stimulus-specific regions found in children as well as those found in adults because we were interested if more mature stimulus-specific activation during the 2-back WM task might predict better performance on other cognitive tasks. For adults, we only used peak activation from stimulus-specific regions in adults as predictors for performance on the 2-back WM task. The exact coordinates of the peak voxels for these analyses are indicated in Table 3.

### Table 1

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Mean score (SD)</th>
<th>Median RT (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyscalculia screener</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speeded response</td>
<td>75 (3)</td>
<td>551.5 (60.1)</td>
</tr>
<tr>
<td>Dot enumeration</td>
<td>112 (13)</td>
<td>2560.0 (737.4)</td>
</tr>
<tr>
<td>Numerical Stroop</td>
<td>111 (14)</td>
<td>1273.3 (233.7)</td>
</tr>
<tr>
<td>Addition</td>
<td>109 (12)</td>
<td>3096.1 (1357.0)</td>
</tr>
<tr>
<td>Missing pieces</td>
<td>119 (10)</td>
<td></td>
</tr>
<tr>
<td>Word sounds</td>
<td>138 (9)</td>
<td></td>
</tr>
<tr>
<td>Spelling</td>
<td>138 (7)</td>
<td></td>
</tr>
<tr>
<td>Visual search</td>
<td>92 (10)</td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td>114 (10)</td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td>108 (8)</td>
<td></td>
</tr>
<tr>
<td>Dyslexia screener</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading subsection</td>
<td></td>
<td>751.5 (60.1)</td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td>2560.0 (737.4)</td>
</tr>
<tr>
<td>Visual scanning</td>
<td></td>
<td>1273.3 (233.7)</td>
</tr>
<tr>
<td>Stroop</td>
<td></td>
<td>3096.1 (1357.0)</td>
</tr>
<tr>
<td>Missing pieces</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocabulary</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Scores are standardized to 100.

### Results

#### Behavioral results

Table 1 provides a summary of the children’s scores on the Dyscalculia and Dyslexia screeners. All children had an average or above average WASI full-scale IQ score [mean = 121; standard deviation (SD) = 14]. None of the children in this sample had indications of dyslexia or dyscalculia. Table 2 provides results of the forward and backward letter and digit span task for children and adults. As expected, adults had a significantly longer span in all four measures than children, (t(28) = 3.6, p < 0.01).

For the 2-back WM task, we performed separate 2-way ANOVAs for accuracy and reaction time with age (adults vs. children) and stimulus type (digits, letters, or faces) as factors. Children were significantly slower than adults (Fig. 2A; F(1,28) = 11.42, p < 0.01) and less accurate (Fig. 2B; F(1,28) = 25.05, p < 0.001). For accuracy, we also found a significant interaction between age and stimulus type (F(2,56) = 4.79, p < 0.05). Separate one-way ANOVAs for children and adults revealed significant main effects of stimulus type for children (F(2,28) = 4.19, p < 0.05) but not adults (F(2,28) = 2.36, p = 0.11). Post-hoc pairwise t-tests indicated that children’s accuracy was marginally better for faces than for digits only (t(14) = 2.42, p = 0.09, Bonferroni corrected).

#### Relationships among behavioral measures

To assess relationships among our different measures of WM, we calculated correlations between the forward and backward WM span scores and the corresponding 2-back WM tasks as well as overall accuracy on the 2-back WM task. For adults, there were no significant correlations between these measures. Children’s backward digit span, however, correlated significantly with their overall accuracy (i.e. the average across all three WM conditions) on the 2-back WM task (r = 0.55, p < 0.05).

Next, we were interested in the relationship between the WASI IQ score as well as the different subsections of the Dyslexia and Dyscalculia screener and performance on the 2-back WM task. To assess whether accuracy on any combination of the three conditions of the 2-back WM task predicts IQ or performance on the Dyslexia and Dyscalculia screener, we conducted stepwise multiple regressions in which accuracy on the digit, letter, and face 2-back WM task were entered as possible predictors. There was no combination of WM performance that predicted a significant amount of variability in IQ. A significant amount of the variation in performance on the reading subsection of the Dyslexia Screener was predicted by accuracy differences on the 2-back letter WM task (Fig. 3; r = 0.52, standardized β = 0.52, F(1,13) = 4.87, p < 0.05) and by a model that included accuracy on the 2-back letter and face WM tasks as statistically significant predictors of IQ.

### Table 2

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Children Mean ± SD</th>
<th>Adults Mean ± SD</th>
<th>Statistical results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forward digit span</td>
<td>5.6 (0.7)</td>
<td>7.1 (1.5)</td>
<td>(t(28)) = 3.6, p &lt; 0.01</td>
</tr>
<tr>
<td>Backward digit span</td>
<td>4.0 (0.8)</td>
<td>5.9 (1.5)</td>
<td>(t(28)) = 4.5, p &lt; 0.001</td>
</tr>
<tr>
<td>Forward letter span</td>
<td>4.4 (0.7)</td>
<td>6.4 (1.2)</td>
<td>(t(28)) = 5.4, p &lt; 0.001</td>
</tr>
<tr>
<td>Backward letter span</td>
<td>3.6 (0.7)</td>
<td>6.3 (1.9)</td>
<td>(t(28)) = 5.4, p &lt; 0.001</td>
</tr>
</tbody>
</table>
predictors ($r=0.75$, $F(2,12)=7.92$, $p<0.01$). Both letter (standardized $\beta=1.8$, $p<0.01$) and face WM performance (standardized $\beta=-1.4$, $p=0.02$) demonstrated significant effects on reading scores. However, whereas accuracy on the letter WM task showed a positive relationship, accuracy on the face WM task showed a negative relationship. No other models were significant predictors of performance on any of the other subsections of the Dyslexia Screener.

For the Dyscalculia Screener, we found that variation in accuracy on the digit WM task predicted a significant amount of the variance in performance in the Numerical Stroop subsection (Fig. 3; $r=0.58$, standardized $\beta=0.58$, $F(1,13)=6.59$, $p<0.03$) as well as in the addition subsection (Fig. 3; $r=0.64$, standardized $\beta=0.64$, $F(1,13)=9.22$, $p=0.01$). No other models provided significant predictions of performance on these or the other subsections of the Dyscalculia Screener.

**FMRI results**

**Digit WM**
To identify brain regions uniquely involved in digit WM, we contrasted activations during the digit WM with activations during the letter and the face WM tasks. For adults, a random effects analysis revealed significantly greater activation for the digit WM condition than for the letter and the face WM conditions in the inferior parietal lobules bilaterally (see Fig. 4), and extending into the superior parietal lobe and the post-central gyrus in the right hemisphere (see Table 3). This activation cluster is consistent with regions previously reported to be active during semantic numerical processing (for review see Dehaene et al., 2005). Children showed significantly more activation for the digit condition than the letter and the face WM conditions in the right pre- and post-central gyrus and in the right inferior frontal gyrus (see Fig. 4 and Table 3). A direct comparison between children and adults revealed significantly greater activation in adults than children in the right middle and superior frontal gyri.

**Face WM**
To identify regions specifically involved in face WM, we compared activations for the face WM condition to activations for the digit and letter WM conditions. As shown in Fig. 6, for both children and adults, we found significantly greater activation in the lateral fusiform gyrus (Fig. 6). This region is typically found to be active in response to faces compared to other visual stimulus categories (Kanwisher et al., 1997; Dehaene et al., 2005).

**Letter WM**
To identify brain regions specifically involved in letter WM, we contrasted activations for the letter WM condition to activations for the digit and the face WM conditions. For adults, the letter condition evoked more activity in the left middle occipital gyrus (see Fig. 5 and Table 3). This activation is consistent with regions that have been previously reported as engaged in abstract encoding of letters and words (Cohen et al., 2000; Dehaene et al., 2002; Nobre et al., 1994). For children, we found only marginally significant activation in the right cerebellum (see Fig. 5 and Table 3). A direct comparison of children and adults revealed significantly greater activation for children than adults in the right middle and superior frontal gyri.
Puce et al., 1996). In addition, we found significantly greater activation in the face condition in the right precuneus, left cingulate gyrus, and bilateral superior frontal gyri for adults (Table 3). For children, we also observed activation in the right medial frontal gyrus and the bilateral anterior cingulate as well as the right posterior cingulate (Table 3). A direct comparison of children and adults revealed only significantly greater activation for children than adults in the left inferior parietal lobule (Table 4).

Multiple regression results for brain activation and behavioral performance

We assessed whether category-specific WM activation was a significant predictor of performance on the 2-back WM task or other behavioral measures. To this end, we first extracted the mean activity of the 10 most strongly activated voxels in a 15 mm sphere around the peak activation for each stimulus-specific region in adults (Table 3). Thus, our analysis took inter-subject variability in the exact peak activation locations into account. These mean activations were then entered as possible predictors for performance on the category specific 2-back WM tasks in stepwise multiple regression analyses.

For adults, category-specific activations did not predict performance on the 2-back WM tasks. For children, activity in the letter-specific area for adults was a significant predictor of spelling performance (Fig. 7; \( r=0.63 \), standardized \( \beta=0.63 \), \( F(1,11)=7.31, p<0.03 \)). No other models provided significant predictions of behavioral performance on the Dyslexia or Dyscalculia Screener or the 2-back WM task. There were also no models with activation in category-specific brain regions for children as possible predictors that provided significant predictions of their behavioral measures.

Discussion

Consistent with prior studies (Knops et al., 2006), we found that adults activate regions in the bilateral inferior and superior parietal lobules more during a 2-back WM task with Arabic numerals as compared to the same task with letters or face stimuli. They also activate the left middle occipital gyrus more during a 2-back WM task with letters as compared to the same task with digits or face stimuli. In contrast, 8-year-old children did not exhibit category-selective activity for letters and digits in these regions. However, children and adults exhibited face-specific activity in the lateral fusiform gyri during a 2-back WM task with faces. The following three (not mutually exclusive) explanations for these developmental differences are discussed below: 1) Children do not process the stimuli as fully as adults if the primary task is very demanding. 2) Children do not yet activate category-specific regions of the brain in general. 3) Children do not yet activate category-specific regions of the brain when such processes are task-irrelevant.

Digit WM

Regions in the parietal cortex, especially the intraparietal sulcus (IPS) have been implicated in processing numerical information. For example, the horizontal segment of the IPS is active bilaterally during number comparisons (Pinel et al., 2001), numerical estimations
Table 3
Summary of random effects analyses contrasting stimulus-specific WM activity across the whole brain for children and adults

<table>
<thead>
<tr>
<th>Age</th>
<th>Contrast</th>
<th>Region</th>
<th>Coordinates</th>
<th>Z-score</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Adults</td>
<td>D=(L and F)</td>
<td>L Inf parietal lobule</td>
<td>−38</td>
<td>−45</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−56</td>
<td>−30</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Sup parietal lobule</td>
<td>−60</td>
<td>−38</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>19</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R post-central gyrus</td>
<td>38</td>
<td>−41</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Middle occipital gyrus</td>
<td>−69</td>
<td>−64</td>
<td>−7</td>
</tr>
<tr>
<td>F=(D and L)</td>
<td>R fusiform gyrus</td>
<td>45</td>
<td>−60</td>
<td>−21</td>
<td>4.84*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L fusiform gyrus</td>
<td>−30</td>
<td>−83</td>
<td>−28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R precuneus</td>
<td>4</td>
<td>64</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Cingulate gyrus</td>
<td>50</td>
<td>33</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R precuneus</td>
<td>4</td>
<td>−56</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Sup frontal gyrus</td>
<td>11</td>
<td>41</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Sup frontal gyrus</td>
<td>−8</td>
<td>49</td>
<td>42</td>
</tr>
<tr>
<td>Children</td>
<td>D=(L and F)</td>
<td>R precentral gyrus</td>
<td>50</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R post-central gyrus</td>
<td>60</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf frontal gyrus</td>
<td>56</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Cerebellum</td>
<td>11</td>
<td>−60</td>
<td>−7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Cerebellum</td>
<td>26</td>
<td>−64</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R occipital lobe, lingual gyrus</td>
<td>41</td>
<td>−33</td>
<td>−21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Inf parietal lobule</td>
<td>−23</td>
<td>−98</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Fusiform gyrus</td>
<td>−26</td>
<td>−86</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Cerebellum</td>
<td>−19</td>
<td>−90</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Occipital lobe, lingual gyrus</td>
<td>19</td>
<td>−98</td>
<td>−7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R medial frontal gyrus</td>
<td>4</td>
<td>41</td>
<td>−7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Ant cingulate</td>
<td>4</td>
<td>19</td>
<td>−7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Ant cingulate</td>
<td>−4</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Cerebellum</td>
<td>−45</td>
<td>−53</td>
<td>−28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Cerebellum</td>
<td>−41</td>
<td>−49</td>
<td>−21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Post-central gyrus</td>
<td>−8</td>
<td>56</td>
<td>21</td>
</tr>
</tbody>
</table>

Coordinates refer to Talairach coordinates; D=digit condition; L=letter condition; F=face condition; R=right hemisphere; L=left hemisphere; Inf=inf inferior; Sup=superior; Ant=anterior; Post=posterior. The threshold for significance on the voxel-level was set to p<0.01 (uncorrected). Threshold for significant clusters reported here was set at p<0.05 on the cluster-level and a cluster size of 8 voxels, except *p=0.0196 cluster-level, and **p=0.065 cluster-level.

a Peak voxel for brain-behavior analysis based on adults’ category-specific brain activation.
b Peak voxel for brain-behavior analysis based on children’s category-specific brain activation.

Table 4
Summary of significant interactions between age (children vs. adults) and stimulus type across the whole brain

<table>
<thead>
<tr>
<th>Age comparison</th>
<th>Contrast</th>
<th>Region</th>
<th>Coordinates</th>
<th>Z-score</th>
<th>Cluster size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults&gt;children</td>
<td>D=(L and F)</td>
<td>L Inf parietal lobule</td>
<td>−34</td>
<td>−49</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L Inf parietal lobule</td>
<td>−45</td>
<td>−34</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−53</td>
<td>−34</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Sup parietal lobule</td>
<td>19</td>
<td>68</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Precuneus</td>
<td>26</td>
<td>56</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−38</td>
<td>−45</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R middle frontal gyrus</td>
<td>34</td>
<td>34</td>
<td>42</td>
</tr>
<tr>
<td>Children&gt;adults</td>
<td>L=(D and F)</td>
<td>R Sup frontal gyrus</td>
<td>38</td>
<td>23</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Sup frontal gyrus</td>
<td>23</td>
<td>34</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−34</td>
<td>−41</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−56</td>
<td>−30</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R Inf parietal lobule</td>
<td>−45</td>
<td>−30</td>
<td>42</td>
</tr>
</tbody>
</table>

Coordinates refer to Talairach coordinates; D=digit condition; L=letter condition; F=face condition; R=right hemisphere; L=left hemisphere; Inf=inferior; Sup=superior. The threshold for significance on the voxel-level was set to p<0.005 (uncorrected). Threshold for significant clusters reported here was set at p<0.05 on the cluster-level and a cluster size of 8 voxels, except *p<0.0196 cluster-level, and **p<0.005 uncorrected voxel-level.
the IPS region during physical but not numerical judgments (Kaufmann et al., 2006). This difference might be attributable to automatic processing of numerical symbols during simple physical size judgments in adults but not children. An alternative explanation of the lack of an interference effect in the numerical Stroop task is that young children might not access numerical information fast enough to interfere with the size judgment (Mussolin and Noel, 2007; Szucs et al., 2007). Thus, in our study, numerical stimuli could have been presented too rapidly to be fully analyzed semantically. However, this explanation seems unlikely since children at this age can make semantic numerical judgments within the range of our stimulus presentation rate (Holloway and Ansari, in press; Sekuler and Mierkiewicz, 1977).

Letter WM

Neuroimaging studies of single letter and letter string processing in the visual domain have repeatedly found activation in a region of left ventral occipito-temporal cortex dubbed the visual word form area (VWFA) (Dehaene et al., 2002; Flowers et al., 2004; Nobre et al., 1994; Petersen et al., 1990; Tarkiainen et al., 1999). The VWFA is thought to process visually but not auditorily presented letter strings (Dehaene et al., 2002) and to be insensitive to perceptual variations in font, size or color of the letters (Polk and Farah, 2002). Thus, it is thought to mediate between perceptual and high-level linguistic processes (McCandliss et al., 2003). Adults also show activation in the VWFA even in the absence of consciously perceived letter strings (Dehaene et al., 2001). Adults, but not children, in our study showed larger activation for letters than digits or faces in the 2-back WM task in a region close to the expected location of the VWFA (Dehaene et al., 2002). These findings suggest that adults recognize letters and process them differently than other perceptually similar symbols such as Arabic numerals. However, the direct comparison between children and adults failed to reach significance suggesting that children may show some activation of this region albeit not a strong one.

Our finding that adults but not children show activation in the VWFA during the letter WM task is consistent with the idea that the specificity of the VWFA is a result of experience and acquired expertise rather than an innate specialization (McCandliss et al., 2003). In fact, some children did show activation in the VWFA and the level of activation predicted children’s spelling performance; however, as a group, children did not exhibit consistent activation of the VWFA. The lack of a significant group-level activation in the VWFA for the letter WM task in children may be explained in one of three ways. First, the primary task demands might have been so high that there were no resources to more fully process the letter stimuli. Thus, future studies should vary task difficulty to assess whether letter-specific activation is modulated by primary task demands.

Second, it is possible that 8-year-old children do not yet process letters in the left occipito-temporal region, or third, it is possible that they do not yet activate this region if abstract processing of the letters presented too rapidly to be fully analyzed semantically. However, this explanation seems unlikely since children at this age can make semantic numerical judgments within the range of our stimulus presentation rate (Holloway and Ansari, in press; Sekuler and Mierkiewicz, 1977).

Although there is ample evidence that processing of visual letter strings activates the VWFA, it is important to note that this region may not exclusively process words. Indeed, some studies have found activation in this region for tasks that do not require processing of visual word forms (Price and Devlin, 2003) or require processing of visual symbols foreign to the participants’ language (Xue and Poldrack, 2007).
is task-irrelevant. A previous study demonstrated that children and adults exhibit similar activity in the left occipito-temporal region for letter strings as compared to other symbols when performing a simple target detection task (Parviainen et al., 2006). In that study, the only difference between children and adults was in the timing of the activation, which was delayed in children. Thus, it seems unlikely that in our task children show different activation patterns than adults because of a lack of sensitivity to letters in the VWFA per se. Instead the primary task demands may have prevented them from processing the stimuli more fully.

In contrast to adults, children showed trends of greater category-specific activation for letters in superior and middle frontal gyrus. It is possible that this activity is due to greater task difficulty for children than adults. However, this explanation seems unlikely given the differences in category-specific activation for children between the three different WM conditions.

Face WM

Both children and adults showed greater activation for the 2-back face WM task as compared to the digit and letter WM task in the bilateral inferior temporal lobes, in a region consistent with the expected location of the fusiform face area (FFA; Kanwisher et al., 1997; Puce et al., 1996). These findings are in line with previous studies that found graded activation in the FFA as a function of memory load for faces (Druzgal and D’Esposito, 2001; 2003) supporting the idea that the FFA is part of a WM system for faces. Thus, it is conceivable that this part of the WM system is already functional by 8 years of age. However, our experimental design does not permit us to test strict face-specificity due to the lack of a perceptually comparable non-face stimulus-category such as buildings or house-hold objects.

Nevertheless, it is important to note that this condition shows that differences in brain activity between children and adults in the digit and letter WM conditions are not due to differences in performance. Children performed significantly worse on the face WM task than adults, but their brain activity exhibited a very similar pattern to that observed in adults. This finding in itself is noteworthy as previous studies have reported delayed maturation of the FFA in children and face-specificity in the FFA usually does not arise until adolescence (Golari et al., 2007; Scherf et al., 2007). However, as mentioned above further studies employing stricter perceptual controls are necessary to determine if the results observed in this study will hold up to a more conservative definition of face-specificity.

Despite the similarities in brain activity in the fusiform gyrus, children showed significantly greater face-specific activity than adults in inferior parietal regions. It is possible that this effect is due to children’s greater difficulty with the task. However, as mentioned above it seems unlikely since task difficulties existed throughout all WM conditions but did not lead to this activation pattern for digit or letter WM. It is also possible that adults showed less activation in the parietal regions for the face WM task as compared to the other two conditions, whereas children activated this region equally throughout all three WM conditions.

Brain-behavior relationships

For children, activity in the VWFA was a significant positive predictor of spelling scores in the Dyslexia Screener. Thus, it seems that more automatic abstract letter identification as measured by greater activation of the VWFA in the 2-back letter WM task is related to increases in spelling performance. Shaywitz et al. (2002) previously reported positive correlations between children’s brain activity in a similar region in the left occipito-temporal lobe and pseudo-word reading scores as measured by the Woodcock–Johnson Word Attack test. In our study, however, reading scores did not correlate with activation in the VWFA. One possible explanation may be the nature of the spelling task used in the Dyslexia Screener. Since the screener is fully computerized, the spelling task requires children to match a spoken word to one of several visually presented words on the screen. Incorrect choices are pseudo-words but are close in spelling to the correct choice. The reading subtest of the Dyslexia Screener comprises a similar matching task in addition to a task in which children have to find an appropriate completion for a visually presented sentence. Not surprisingly, the correlation between scores on the spelling and the reading subtest in the validation sample of the Dyslexia Screener are very high \( r = 0.83 \). Thus, it is possible that our spelling task is more similar to the pseudo-word reading in the Word Attack task used by Shaywitz et al. (2002) and that the actual reading task we used requires greater sentence comprehension and semantic processing of the words, i.e. higher level cognitive processes that are not directly related to processes in the VWFA. It is important to note that brain activity in category-specific regions did not predict performance on the WM task for either of the two age groups.

Relationships among behavioral measures

Accuracy on the 2-back letter WM task was a significant positive predictor of reading scores. Furthermore, adding accuracy on the face WM task to the model improved its predictive power; however, face WM performance was negatively related to reading scores. Thus, variation in reading abilities in a sample of typically developing children without any signs of dyslexia seems to be positively related to processes in the phonological loop. This result is in line with findings in previous studies in adults (Daneman and Carpenter, 1980). However, in addition to the phonological loop, reading scores seem to benefit from decreased processes in the visuo-spatial sketchpad. This result is surprising and should be investigated further in future studies.

Performance on the addition and numerical Stroop subtests of the Dyscalculia screener were best predicted by performance on the 2-back digit WM task suggesting that these numerical abilities are related to domain-specific processes in the phonological loop. Previous findings in adults have provided evidence for a relationship between arithmetic and domain-specific WM components (Heathcote, 1994; Trbovich and LeFevre, 2003) as well as domain-general WM components (see DeStefano and LeFevre, 2004, for a review; Logie et al., 1994). Studies examining WM in children with math disabilities also found correlations between arithmetic abilities and mostly domain-general WM performance (Bull et al., 1999; Geary et al., 1991; Passolunghi and Siegel, 2001; Passolunghi and Siegel, 2004; Passolunghi et al., 2007); however, some studies have also found correlations between arithmetic abilities and domain-specific WM performance (Siegel and Ryan, 1989). Our results provide further evidence that arithmetic abilities in typically developing children are related to number-specific WM capacities rather than domain-general WM.

Finally, children showed a significant correlation between overall accuracy on the 2-back WM task and the backward digit span but not for the backward letter span. Adults showed no correlation between the 2-back WM task and the backward span tasks. It is possible that this absence is due to adults’ overall high performance on the 2-back WM task. Further work will be necessary to delineate the exact differences between these tasks and whether there are additional age differences.

Conclusion

In summary, during a 2-back WM task involving digits and letters adults but not children activate stimulus-specific regions known to be active during abstract or semantic processing of these stimuli. In
contrast, children and adults showed highly similar brain activity in the fusiform face area when performing the same task with face stimuli. These findings raise the possibility that children do not yet have mature domain-specific regions for symbols (digits and letters). Alternatively, domain-specific regions in children may not be recruited due to the high processing demands of the WM task, which may have rendered children unable to abstractly or semantically process the symbols. Future research will be necessary to parse apart these competing hypotheses.

Acknowledgments

We thank Lauren Marx and Margaret Vogel for help with data collection, Jessica Cantlon for help with data analysis, and Klaus Libertus and two anonymous reviewers for comments on an earlier draft of this paper. This research was supported by John Merck Scholars Awards and a grant from the National Institutes of Health (HD057173-01) to EMB and KAP.

References


