Running head: STRUCTURE/BEHAVIOR RELATIONS IN SELF-DERIVATION

Relations between neural structures and children's self-derivation of new knowledge through memory integration

Patricia J. Bauer[1], Jessica A. Dugan[1], Nicole L. Varga[2], and Tracy Riggins[3]

[1]Department of Psychology, Emory University; [2]Center for Learning and Memory, University of Texas at Austin; [3]Department of Psychology, University of Maryland.

Support for this research was provided by HD067359 and NSF1748293 to Patricia J. Bauer; an Institutional Predoctoral Fellowship T32 HD071845 for Jessica A. Dugan; and HD079518 to Tracy Riggins. The authors thank Morgan Botdorf, Kelsey Canada, Lisa Cox, Jake Hansen and the Neurocognitive Development Lab at the University of Maryland for assistance with this research, and extend a special note of appreciation to the children and families who took part.

Correspondence may be addressed to Patricia J. Bauer, 36 Eagle Row, Department of Psychology, Emory University, Atlanta, Georgia 30322, USA; patricia.bauer@emory.edu.

Abstract

Accumulation of semantic or factual knowledge is a major task during development. Knowledge builds through direct experience and explicit instruction as well as through productive processes that permit derivation of new understandings. In the present research, we tested the neural bases of the specific productive process of self-derivation of new factual knowledge through integration of separate yet related episodes of new learning. The process serves as an ecologically valid model of semantic knowledge accumulation. We tested structure/behavior relations in 5- to 8-year-old children, a period characterized by both age-related differences and individual variability in self-derivation, as well as in the neural regions implicated in memory integration, namely the hippocampus and prefrontal cortex. After controlling for the variance in task performance explained by age, sex, verbal IQ, and gray-matter volume (medial prefrontal cortex, mPFC, only), we observed relations between right mPFC thickness and memory for information explicitly taught to the children as well as the new information they self-derived; relations with the volume of the right hippocampus approached significance. This research provides the first evidence of the neural substrate that subserves children's accumulation of knowledge via self-derivation through memory integration, an empirically demonstrated, functionally significant learning mechanism.

Key words: hippocampus, learning, memory integration, prefrontal cortex, self-derivation

1. Introduction

A major component of intelligence is the amount of semantic or factual knowledge that has been accumulated (so-called *crystallized* intelligence: Cattell, 1963; Horn & Cattell, 1966). Entries in the semantic knowledge base are made through direct experiences, such as explicit tuition and observation. Semantic knowledge is further expanded through productive processes, such as analogy and inference (Gentner, 1983; Goswami, 2002), as well as the productive process examined in the present research, namely, self-derivation of new factual knowledge through integration of separate yet related episodes (Bauer, 2012).

The neural structures involved in integration across episodes (e.g., Zeithamova & Preston, 2010) and the temporal unfolding of the neural processes they support (Bauer & Jackson, 2015; Varga & Bauer, 2017a) have been subjects of research in adults. Because there has been little research in children, the brain bases of self-derivation in development are unknown. In the present research, we examined self-derivation through integration in 5- to 8-year-olds, and associated age-related and individual variability in children's performance to measures of structural development in the medial temporal and frontal lobes. We selected these regions because they are implicated in memory integration in adults (see below).

1.1 Self-derivation through Integration: Behavior

The productive process of self-derivation of new facts through integration of separate episodes is of particular interest because it serves as an ecologically valid model for accumulation of knowledge (e.g., Esposito & Bauer, 2017). In this paradigm, participants learn a true but novel fact (e.g., *dolphins talk by clicking and squeaking*) in one episode of instruction. After a filled delay, in a second episode, they learn another true but novel fact that is related to the first (*dolphins live in groups called* pods). Following another filled delay, they are asked first open-ended and then forced-choice questions probing for self-derivation of another new fact, based on integration of the episodes (*how does a pod talk*?). Critically, as is frequently the case in the world outside the laboratory, learning is based on a single trial (see Bauer & San Souci, 2010, Experiment 2, for findings when learning is ensured through repeated trials). Moreover, as is also the case outside the laboratory, participants are not informed that any of the material is related (see Bauer, Varga, King, Nolen & White, 2015, for findings under "hint" conditions); they are not given practice with integration of the episodes or self-derivation from them. These testing conditions ensure the ecological validity of the paradigm. Control conditions entailing learning of only one member of a pair of related facts make clear that across-episode integration is required for self-derivation (e.g., Bauer & Larkina, 2017).

There are age-related and individual differences in self-derivation throughout childhood. Whereas 4-year-olds self-derive new factual knowledge on 13% of trials, 6- and 8-year-olds do so on roughly 50% and 83% of trials, respectively (Bauer & Larkina, 2017; see also Bauer et al., 2015; Bauer & San Souci, 2010). At all ages, performance ranges from near floor to near ceiling. Newly self-derived information is retained over time (Varga & Bauer, 2013; Varga, Stewart, & Bauer, 2016), providing support for this paradigm as a model for knowledge accumulation. As well, self-derivation relates to reading and math achievement (Esposito & Bauer, 2017; see Varga, Esposito, & Bauer, 2018, for comparable findings with adults).

1.2 Neural Substrate of Memory Integration

The neural bases of integration of separate episodes have been examined in nonhuman animals (Preston & Eichenbaum, 2013) and human adults (e.g., Kumaran, Summerfield, Hassabis, & Maguire, 2009; Sweegers, Takashima, Fernández, & Talamini, 2014; Zeithamova & Preston, 2010); the majority of human work has used the associative inference paradigm. In this paradigm, successful encoding of arbitrarily related pairs of stimuli (stimulus pairs A-B, B-C), and novel associative inferences between them (A-C; e.g., Schlichting, Zeithamova, & Preston, 2014; Zeithamova, Dominick, & Preston, 2012) relate to activations in both medial temporal, including hippocampal, and medial prefrontal cortex (e.g., ventromedial PFC or VMPFC). These regions are thought to subserve (a) reactivation of the first pair of stimuli (A-B) as the related pair (B-C) is encoded; and (b) binding of previously and newly encoded stimuli into an integrated memory representation (A-B-C) that supports novel associations (A-C) (Schlichting & Preston, 2015). These regions also are implicated in encoding of new information related to prior knowledge (e.g., Hebscher & Gilboa, 2016; van Kesteren, Rijpkema, Ruiter, Morris, & Fernández, 2014).

Though there have been no studies of functional activations as children engage in memory integration, Schlichting, Guarino, Schapiro, Turk-Browne, and Preston (2016) examined relations between hippocampal volume and memory integration in the associate inference task in 6-30-year-olds. Smaller volume in the hippocampal head correlated to higher levels of associative inference. The relation may reflect differences in the efficiency of encoding of overlapping representations and/or of integration or binding (Ghetti & Bunge, 2012; Olson & Newcombe, 2014). Although based on adult data, we may expect integration also would relate to structural differences in prefrontal cortex (e.g., Benes, 2001; Giedd et al., 1999; Gogtay et al., 2004; Huttenlocher & Dabholkar, 1997; Sowell et al., 2004), this possibility has not been tested.

1.3 Neural Substrate of Self-derivation through Memory Integration

The present research is the first developmental study of the neural bases of self-derivation of new factual knowledge through integration of separate yet related episodes of new learning. We selected the task of self-derivation through integration because it is an ecologically valid model for accumulation of semantic knowledge and thus stands to inform the neural bases of an empirically demonstrated, functionally significant learning mechanism. It also can inform the neural structures involved in multiple aspects of learning, including episodic encoding of explicitly learned facts and productive extension beyond them (self-derivation).

We focused on 5- to 8-year-olds because this is a period of both age-related and individual variability in self-derivation as well as of substantial structural change in the hippocampus (e.g., Riggins, Geng, Botdorf, Canada, Cox, & Hancock, 2018; Schlichting et al., 2016) and prefrontal cortex (e.g., Giedd et al., 1999; Gogtay et al., 2004; Sowell et al., 2004). Overall, studies of hippocampal development suggest that volumes vary as a function of age and that change is not uniform across the structure. For example, in childhood, the anterior portion of the hippocampus (i.e., head) increases in volume, whereas the posterior hippocampus (i.e., body and/or tail) shows no difference or decreases in volume with age (Riggins et al., 2018; Schlichting et al., 2016). Across development, there are hemispheric asymmetries such that the right hippocampus is larger than the left (e.g., Pfluger et al., 1999; Utsunomiya et at., 1999); the degree of asymmetry tends to decrease with development (Gogtay et al., 2006; Szabó et al., 1999). Studies of cortical development also suggest gray matter changes with age, with most studies reporting a reduction in volume or cortical thinning across childhood; there is regional specificity with regard to timing (see Brown & Jernigan, 2012, for review). Regardless of these age-related differences, in both episodic (Riggins et al.) and associative (Schlichting et al.) tasks, structural variability relates to performance.

In summary, because the paradigm testing self-derivation through memory integration operates over real-world factual knowledge, the study stands to inform the neural structures involved in learning of new semantic information, integration of the new information into the network of prior knowledge, and productive knowledge extension. We predicted that, consistent with prior research, performance on the task would relate to age and verbal IQ (e.g., Bauer, Blue, Xu, & Esposito, 2016; Esposito & Bauer, 2017). We also predicted that measures of brain structure (i.e., hippocampus and mPFC) would contribute to explanation of variance in task performance. Yet because of the paucity of research on structure/behavior relations in self-derivation of new knowledge through integration—in either children or adults—we did not make specific hypotheses regarding the direction of relations (i.e., whether larger volume or greater thickness would be positively or negatively related with behavioral performance).

2. Method

2.1 Participants

Sixty-eight 5- to 8-year-old children (33 female; *M* age=7.34 years, *SD*=1.16) participated in the present study, which is part of an ongoing longitudinal investigation examining brain and memory development in early childhood (Riggins et al., 2018). Sixty-six children provided useable MRI data (2 6-year-olds [1 female] were not scanned due to contraindications, e.g., dental work). The behavioral task measures and the relations between neural and task measures are presented for the first time in this report.

2.2 Stimuli

Stimuli were six pairs of facts (i.e., stem facts), all used in previous research (Bauer & Larkina, 2017; Bauer & San Souci, 2010; Esposito & Bauer, 2017; Varga & Bauer, 2013). There was one pair of related facts about each of plants, dolphins, muscles, the solar system, deserts, and athletes. Members of these “stem-fact” pairs could be integrated to derive a novel fact (i.e., “integration” fact). For example, the stem fact *palm tree leaves are called fronds* could be integrated with the related stem fact *palm tree leaves are used to make baskets*, to self-derive the new knowledge that *fronds are used to make baskets* (integration fact). Previous research using these stimuli (e.g., Bauer & Larkina, 2017) included a control condition, wherein participants were presented with only one fact from a given pair, to ensure that the target facts were novel and that memory integration was necessary to derive them.

Each stem fact was displayed individually on a single PowerPoint slide, with 6-10 words per slide. Importantly, the novel integration facts were never presented to the children.

2.3 Procedures

Data were collected across two sessions approximately 8 days apart (*M*=8.11 days, *SD*=4.67). All behavioral testing took place within a single laboratory session and MRI testing took place at a separate session. For most participants, MRI and behavioral data were collected at the first and second sessions, respectively. In a few cases, MRI data from the first session were discovered to be unusable due to motion artifact. These children returned to the lab after behavioral data collection for a second attempt to collect MRI data. This yielded useable MRI data for all participants who were eligible to be scanned (i.e., *n*=66).

2.3.1 Self-derivation through integration task

Children were tested individually in a quiet laboratory room. In the stem-fact presentation phase, children were read six pairs of facts. As depicted in Figure 1, they were first read one of the stem facts from each pair. After an approximately 10 min delay in which they completed unrelated buffer activities that were part of the larger longitudinal study, children were read the second stem fact from each pair. Presentation was counterbalanced such that each member of a stem-fact pair was presented among the first six facts for half of the children and among the second six facts for the other half. To ensure that children attended to the stem-fact presentation, after each fact was read, children were asked “What was this fact about?”, and given two alternatives: the correct answer (e.g., *fronds*) and a distracter (e.g., *oaks*). All children remained on task.

Following presentation of all of the stem-fact pairs, children completed approximately 10 min of unrelated buffer activities, after which the test phase was administered. Children first were asked open-ended questions that probed for self-derivation of the integration facts (e.g., *What are fronds used to make*?). The answers to these questions had not been explicitly presented but could be derived through integration of the related stem facts. Children then were asked open-ended questions that probed recall of the stem facts from which the integration facts were derived (e.g., *What are palm tree leaves called*?). After open-ended testing, for any integration fact questions not answered correctly in open-ended format, children were asked three-alternative forced-choice questions. Finally, children were asked forced-choice questions for any stem facts not answered correctly in open-ended format (see Figure 1 for example items). The question types were administered in this standard order across participants, to avoid cueing of integration facts by recall or recognition of the stem facts. In forced-choice testing, familiarity of the distracters was controlled by including alternatives that had been presented as part of other fact pairs.

Experimenters recorded children's responses on-line. For each fact produced in open-ended testing, children received one point (integration facts: max=6; stem facts: max=12). Children also received one point for each forced-choice question answered correctly. Correct forced-choice responses were added to correct open-ended responses for a total score for integration facts and a total score for stem facts. Because forced-choice questions were only asked when open-ended performance was incorrect, the total score also had a range of 0-6 for integration facts and 0-12 for stem facts. Thus the paradigm yielded 4 dependent measures:

open-ended self-derivation, total self-derivation, open-ended stem fact recall, and total stem fact recall.

2.3.2 MRI

Children underwent training in a mock scanner before MR data acquisition. Images were obtained from a Siemens 3.0-T scanner (MAGNETOM Trio Tim System, Siemens Medical Solutions, Erlangen, Germany) using a 32-channel coil and a T1 magnetization-prepared rapid gradient-echo (MPRAGE) sequence consisting of 176 contiguous sagittal slices (voxel size: .9mm isotropic, TR=1900ms, TE=2.32ms, 900ms inversion time, 9° flip angle, pixel matrix=256×256).

Hippocampal volumes and measures of cortical thickness for left and right hemispheres were obtained using FreeSurfer v5.1 (surfer.nmr.mgh.harvard.edu; Fischl, 2012). Use of FreeSurfer has been validated in children as young as 4 years (Ghosh et al., 2010). Boundary lines separating gray/white and pial surfaces were visually inspected to ensure accuracy. Reviewers inspected the data for specific errors, including slices where the pial boundary included portions of the skull and slices where the gray or white matter extended into or beyond the skull. Manual edits were made if these errors persisted for more than seven slices. Edits were made on approximately 39% of the sample and typically involved fewer than 20 slices per subject. An experienced reviewer completed a final quality check. Cortical thickness was calculated by measuring the distance from the gray/white matter boundary to the pial boundary (Fischl & Dale, 2000). The Desikan-Killiany Atlas, which includes 34 gyral-based cortical regions, was used for cortical parcellation (Desikan et al., 2006). Given our interest in medial prefrontal cortex (mPFC), thickness of rostral anterior cingulate and medial orbitofrontal cortex were combined to create a global measure of mPFC thickness for both the right and left hemispheres. Total gray matter volume was also extracted using FreeSurfer and was used as a control variable in analyses of thickness (Fischl et al., 2002).

Hippocampal volumes obtained from FreeSurfer were refined using the Automatic Segmentation Adapter Tool (ASAT, nitrc.org/projects/segadapter), which corrects systematic errors in automatic segmentations (Wang et al., 2011). To train ASAT, hippocampi for ten subjects were manually traced using boundaries set forth by the “EADC-ADNI Harmonized Protocol for Manual Hippocampal Segmentation" (Frisoni et al., 2015). Ten subjects were randomly selected for manual tracing. It was required that these scans were representative of the age range tested and that they had clear visibility of the hippocampus in both hemispheres to be used as a training case. Following recommended methods (Lee et al., 2015), the parameters used were: 4x4x4 voxel sampling radius, 50% sampling rate, 500 training iteration and dilation radius of 2 voxels. To correct minor over or under-inclusions, manual edits were performed on 13 participants (right hemisphere *n*=9, left hemisphere *n*=5) using the "EADC_ADNI Harmonized Protocol for Manual Hippocampal Segmentation" as a reference (Frisoni et al., 2015).

The hippocampus was then divided into head, body, and tail subregions using manual identification of standard anatomical landmarks. The uncal apex served as the border between the head and body (Weiss et al., 2005). The boundary between the body and tail was identified as the slice at which the fornix separates from the hippocampus and becomes clearly visible (Watson et al., 1992). Raters were unaware of participant age and sex. Reliability for identification of these landmarks indicated 90% agreement within 1 slice and 98% agreement within 2 slices. Intra-class correlation coefficients (ICCs) were high (range=.84-.96).

Finally, hippocampal volumes were adjusted to control for differences in intracranial volume (ICV) using an analysis of covariance approach (Raz et al., 2005; Van Petten, 2004).

Brain extraction was conducted separately in 6 toolboxes: ANTs, AFNI, FSL, BSE, ROBEX, and SPM8. The voxels extracted by at least four toolboxes were included in the brain mask. This approach was adopted because any single toolbox is prone to error and may result in multiple outlier values. The requirement that voxels be extracted from multiple toolboxes increases the validity of the measure (see Tillman et al., 2017, for a similar approach). Previous research in children within this age range (5-8 years) has shown significant independent influences of age and sex on total brain size as well as heterogeneity in these relations as a function of age (Riggins et al., 2018). Therefore, we divided our sample into two groups: younger (i.e., 5-6 years) and older (i.e., 7-8 years) and corrections were carried out for each age group separately. Both age and sex were used to estimate ICV (adjusted volume=raw volume–*b* * (ICV–predicted ICV), in which *b* is the slope of the regression of the ROI volume (Vol(rawi)) on ICV, see Keresztes et al., 2017; Riggins et al., 2018).

2.3.3 IQ

Indices of intelligence were obtained using the vocabulary subtests from age-appropriate standardized intelligence tests (i.e., Wechsler Intelligence Scale for Children-Fourth Edition, or WISC; Wechsler, 2003; and the Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition, or WPPSI; Wechsler, 2012) one to two years prior to collection of the imaging and behavioral data that are the focus of the present report. Although the IQ measures were not obtained contemporaneous with the measures that are the subject of the present report, the vocabulary subtest provides a relatively stable assessment of verbal intelligence (Watkins & Smith, 2013). The measures were used as potential predictors of behavioral task performance.

2.4 Statistical Analysis

Relations between performance on the self-derivation through integration task and brain structure were examined using linear regression. As predictors of performance, we entered Age, Sex, and verbal IQ, and volumes for either hippocampus (right, left), or cortical thickness (right, left). To control for potential differences arising from brain size, gray matter was also included for analyses of cortical thickness.

3. Results

Descriptive statistics for performance on the self-derivation through integration task are shown in Table 1, Panel a. Bivariate correlations between demographic variables (age, sex), verbal IQ, and the self-derivation task are presented in Table 2. There were significant positive correlations with age for all four measures of performance on the self-derivation task. Girls tended to have higher total integration fact performance (open-ended plus forced-choice). As in prior research (Esposito & Bauer, 2017; Varga et al., 2018), verbal IQ also correlated with measures from the self-derivation task.

Descriptive statistics for the MRI measures are shown in Table 1, Panel b. As reflected in Table 2, age was significantly correlated to ICV-adjusted right hippocampal volume; the correlation with ICV-adjusted left hippocampal volume approached significance. Thus the size of the hippocampus tended to increase with age (consistent with Brown & Jernigan, 2012; Hu, Pruessner, Coupé, & Collins, 2013; Uematsu et al., 2012). In contrast, mPFC thickness did not show reliable or even trend-level associations with age. The difference between this finding and that reported in Brown and Jernigan (2012) may be due to the restricted age range of the present sample (i.e., most children were 6-8 years of age; few were younger than 6). Boys tended to have larger ICV-adjusted right hippocampal volumes and girls tended to have thicker left mPFC.[1]

The bivariate correlations (Table 2) revealed only one significant relation between task performance and the neural measures: left mPFC thickness was negatively correlated with the total number of stem facts recalled or recognized (open-ended plus forced-choice). The relative paucity of relations must be understood in light of the robust pattern of correlation between task performance and both age and verbal IQ, and the relatively weak relations between age and the neural measures. To determine whether, with the variance in age and verbal IQ controlled, individual variability in hippocampal volume or mPFC thickness would emerge as a significant predictor, we conducted regression analyses. The results of the analyses with hippocampal volumes are summarized in Table 3. With the variability associated with age and verbal IQ controlled, right hippocampal volume emerged as a marginally significant predictor (p=.07) of open-ended stem fact recall. The model assessing open-ended self-derivation of integration facts showed a trend in the same direction, but a weaker relation (p=.11). Thus after accounting for the substantial variance explained by age and verbal IQ (effect sizes of .40-.53 and .28, respectively), right hippocampal volume approached statistical significance as a predictor of open-ended production of integration facts and open-ended recall of the stem facts from which they were derived (effect size .25). The similar pattern of association for these two measures of task performance is not surprising given that open-ended self-derivation and stem-fact recall are strongly correlated (r=.87, p<.001; see Table 2). As illustrated in Figure 2, the relations were such that higher levels of open-ended self-derivation and recall of stem facts were associated with smaller right hippocampal volume. Left hippocampal volume did not emerge as a predictor. The asymmetrical pattern of relation was statistically significant. That is, the beta weights for self-derivation and stem fact recall for the right hippocampus were significantly different from the beta weights for the left hippocampus (ps<.05; Cummings, 2009). We conducted separate follow-up analyses for the head, body, and tail of the right hippocampus and open-ended self-derivation of integration facts and stem fact recall. None of the subregions emerged as significant independent predictors.

The results of regression analyses of performance on the self-derivation through integration task and mPFC thickness are summarized in Table 4. After controlling for the significant (or marginal) variance in task performance explained by age and sex (total integration fact performance only), right mPFC thickness emerged as a significant predictor of open-ended stem fact recall and a marginally significant predictor of total integration fact performance (open-ended plus forced-choice). As illustrated in Figure 3, Panels A and B, better integration fact performance and stem fact recall was associated with decreased right mPFC thickness. The left and right beta weights were not significantly different (*p*s>.05).

Separate follow-up analyses were conducted to examine the specificity of the relation between right mPFC thickness and both total integration fact performance and open-ended stem fact recall. For total integration fact performance, right rostral anterior cingulate thickness was a significant predictor (β=-.331, p<.05) after controlling for age, sex, verbal IQ, and total gray matter volume (Adjusted R^2=.197, $F(6,59)$=3.650, p<.01). For open-ended stem fact recall, right rostral anterior cingulate thickness was a marginally significant predictor (β=-.246, p=.055) after controlling for age, sex, verbal IQ, and total gray matter volume (Adjusted R^2=.362, $F(6,59)$=7.144, p<.001). Right medial orbital frontal cortex thickness was not related to either task-based measure (*p*s>45). As illustrated in Figure 2, Panels C and D, better integration fact performance and stem fact recall was associated with decreased thickness.

4. Discussion

The present research represents the first test of the neural bases of self-derivation of new factual knowledge through integration of separate yet related episodes of new learning in 5- to 8-year-old children. Consistent with Schlichting et al. (2016), the data suggest a relation between knowledge derived through memory integration and the volume of the hippocampus. In Schlichting et al., better associative inference performance was related to smaller volume in the hippocampal head. In the present research, smaller right hippocampal volume was associated with higher task performance. Relations were specific to the open-ended phase of testing, and extended to recall of the stem facts from which the integration facts were derived. Notably, for both variables, the effects only approached statistical significance, and then only after the variance associated with age and verbal IQ was controlled. Yet the suggestion of structure/behavior relations was strengthened by statistically significant hemispheric asymmetry in the relation (i.e., stronger relations with right than left hippocampus). Moreover, it is noteworthy that hippocampal volume was related to open-ended but not to total task performance, either for integration or for stem facts. The hippocampus is widely interpreted to be more involved in mnemonic tasks that impose greater demands, such as recall and recollection, relative to those that are less mnemonically demanding, such as recognition (Ghetti & Lee, 2014, for a review). Thus the pattern of relations is further evidence of a role for the hippocampus in the task of self-derivation through memory integration. Importantly, although hippocampal volume was positively related with age, the trend was for children with smaller hippocampi to have higher levels of open-ended self-derivation and stem-fact recall. This implies that after controlling for the effects of age, children with smaller volumes were more proficient at the task. This is an important reminder that factors other than age must be considered as determinants of performance. In the present case, performance also was impacted by individual differences in the developmental status of the hippocampus.

Based on studies of activation patterns in adults (van Kesteren et al., 2012; Zeithamova et al., 2012), we also expected to find relations between mPFC thickness and self-derivation through integration. Consistent with this expectation, after accounting for the variance explained by age, sex, verbal IQ, and total gray matter, right mPFC was a marginal predictor of total integration fact performance (open-ended plus forced-choice). The specific subregion of the right rostral anterior cingulate was a significant predictor. As well, mPFC thickness was a significant predictor of open-ended stem-fact recall; the specific region of the right rostral anterior cingulate was a marginal predictor. We note that the pattern of relations between mPFC thickness and stem fact memory as revealed through the regression analyses (right mPFC predicted open-ended stem fact recall) was different from that indexed by bivariate correlations (left mPFC thickness was correlated with total stem fact performance). Although we did not predict this pattern, we suggest that it can be understood in light of Sowell et al.'s (2004) finding of associations between left inferior frontal regions and verbal IQ in children 5-11 years of age. We offer the possibility that our control of the variance associated with domain-general verbal ability suggested by this finding permitted the relation between right mPFC and stem fact recall to emerge. Though we offer this interpretation, we note that in the present research, we did not analyze inferior frontal regions and we did not observe a statistically significant correlation between verbal IQ and left mPFC. As such, the interpretation remains speculative. Future research will be necessary to establish the reliability of the finding and shed further light on hemispheric effects.

In summary, individual and developmental variability of both the hippocampus and mPFC related to task performance. In the case of the hippocampus, the relations were nominally weaker and only approached statistical significance; relations with mPFC, specifically the right rostral anterior cingulate, were significant and nominally stronger.

4.1 Age-related and Individual Variability in Self-derivation through Integration

A prerequisite to a valid test of relations between hippocampal and mPFC volume in self-derivation through integration is that there is adequate age-related and/or individual variability to explain. This criterion was met. In terms of self-derivation, the means for both open-ended and total performance (39% and 77%, respectively) are in line with prior related research (e.g., across the age range 4-8 years, *M* open-ended and total performance is roughly 45% and 70%, respectively: Bauer & Larkina, 2017). This level of correspondence was observed even though before the present study, the individual-sentence paradigm used in the present research had only been used with children 7 years of age and older (Bauer, et al., 2016; Esposito & Bauer, 2017). In previous research, children younger than 7 years were tested using a story-passage paradigm in which true but previously unknown facts were conveyed in the context of stories (e.g., Bauer & Larkina 2017). In contrast, in the present study, facts were conveyed in individual sentences, with no surrounding context. In research conducted in the classroom, differential patterns of performance have been attributed to these paradigm differences (Esposito & Bauer, in press). In the present research, the approach resulted in a desirable spread in scores: (a) age-related variability—performance was significantly correlated with age; and (b) individual variability—in open-ended testing, performance ranged from 0-83% correct in self-derivation of the integration facts, and from 0-100% correct in recall of the stem facts.

4.2 The Role of the Hippocampus and mPFC in Self-derivation through Integration

The critical role of the hippocampus in supporting encoding of relational information has long been accepted, particularly with respect to binding individual elements into episodic memory (Cohen & Eichenbaum, 1993; Mishkin, Vargha-Khadem, & Gadian, 1998; Tulving, 1983). Evidence of this role is apparent in the present research, in the relation between right hippocampal volume and recall of the stem facts upon which self-derivation depended ($p<.07$). More controversial is the role of the hippocampus in supporting *semantic* memory, such as the new factual knowledge self-derived from the stem facts. In particular, individuals with developmental amnesia who exhibit disrupted hippocampal function and episodic memory have been shown to acquire new semantic knowledge (Vargha-Khadem et al., 1997). This has led some to conclude that acquisition of semantic relations, such as factual statements, does not depend on the hippocampus (e.g., Olson & Newcombe, 2014). However, in this regard, it is noteworthy that individuals with medial-temporal lobe amnesia typically require repeated exposures to new information before showing evidence of learning (e.g., O'Kane, Kensinger, & Corkin, 2004). This pattern extends beyond arbitrary stimuli to factual knowledge relevant outside the laboratory, such as events in the news (Manns, Hopkins, & Squire, 2003). Thus although new semantic knowledge can be acquired even in cases of hippocampal damage, patterns of acquisition deviate from those shown by intact adults. One possibility consistent with the overall pattern of findings is that the hippocampus plays a crucial role in rapid acquisition of new knowledge acquired across temporally-extended episodes, as is assessed in the paradigm used in the present research (see Kumaran & McClelland, 2012, for consistent arguments). Relative to other paradigms in common use in the literature (e.g., associative inference), the self-derivation paradigm provides a strong test of this possibility, given that the information over which memory integration must be performed is new, and presented in a single trial. Under these conditions, right hippocampal volume contributed variance to prediction of performance, though with other strong predictors already in the model, the variance explained only approached statistical significance (*p*=.11).

In the present research, relations with hippocampal volume were specific to the right hippocampus; beta weights for the right and left hippocampus were significantly different. Although we did not predict this pattern, it is not without precedent in the literature. Hippocampal hemispheric asymmetry, with greater volume in the right than left, is well established; the effect is observed across development (e.g., Pfluger et al., 1999; Utsunomiya et at., 1999), though it tends to decrease with age (Gogtay et al., 2006; Szabo et al., 1999; perhaps contributing to the observation that among adults, the asymmetry may be limited to the hippocampal head: Woolard & Hecklers, 2012). Among adults, right hippocampal volume is correlated with performance on tests of general cognitive function, including immediate and delayed tests of verbal learning. The effect is specific to the right hippocampus, and the anterior region in particular; the magnitude of the relation is comparable to that observed in the present research (*r*=.22; Woolard & Hecklers, 2012). In the case of the present research, evaluation of the psychological significance of the pattern awaits tests for replication.

In research with adults, PFC has been shown to play a role in derivation of new relations based on memory integration (Zeithamova & Preston, 2010), as well as in encoding of new information related to prior knowledge (van Kesteren et al., 2014). In particular, functional hippocampal-VMPFC coupling has been shown to support extraction of semantic commonalities across separate episodes (Kumaran et al., 2009). In the associative inference paradigm, learning-related increases in VMPFC activation and corresponding decreases in hippocampal activation relate to successful associative inference (Zeithamova et al., 2012; see also van Kesteren et al., 2014, for a similar pattern when new stimuli overlap with prior knowledge). In the present research, we observed a marginally significant contribution of right mPFC to self-derivation when measured in terms of total performance (i.e., open-ended plus forced-choice); the relation with the specific region of the right rostral anterior cingulate was statistically significant. It is possible that the differential pattern of subregion specificity observed in the present research with children compared to that observed in adults (VMPFC) reflects the relative lack of connectivity between prefrontal and medial temporal structures in development (Benes, 2001; Giedd et al., 1999; Gogtay et al., 2004; Huttenlocher & Dabholkar, 1997; Sowell et al., 2004). Lower levels of structural connectivity could correspond to differential functional connectivity and thus, logically, to a somewhat different role in behavior, relative to that observed in adults. Interestingly, the relation was observed in the context of marginally significant variance explained by age, and in terms of total integration fact performance, including that in response to forced-choice options. In adults, memory integration most often is measured via forced-choice, as opposed to open-ended report (e.g., Zeithamova et al. 2012). Thus the relation with mPFC and the specific relation with right rostral anterior cingulate observed in the present research may be an early sign of the emergence of more adult-like function. Open-ended recall of the stem facts showed similar patterns of relation with mPFC and with right rostral anterior cingulate. As noted earlier, in light of the strong relation between self-derivation and stem-fact recall, similarity in patterns is not unexpected.

Task differences are another possible contributor to the differential pattern of subregion specificity observed in the present research with children compared to that observed in adults. The focus of the present research was self-derivation of new factual knowledge through integration of separate yet related episodes in which true but previously unknown facts were learned. Learning of the explicitly taught stem facts was based on a single trial; children were not informed that any of the facts were related to one another; children were not given practice at self-derivation prior to the test; both open-ended and forced-choice tests were administered. These conditions of testing differ markedly from those typical in associative inference tasks, in which what is learned are arbitrary paired associates. By design, arbitrary associations make little contact with prior knowledge, which may alter the contribution of the PFC, as well as the interaction of the PFC and hippocampus (see Shing & Brod, 2016, for discussion). What is more, in associative inference tasks, learning often is brought to a high criterion (e.g., 85% in Preston, Shrager, Dudukovic, & Gabrieli, 2004), participants are given practice with the task of recognizing novel associations in advance of testing (e.g., Schlichting et al., 2016; Schlichting & Preston, 2015), and testing is forced-choice only. Separately—or in combination—these significant task differences could contribute to different patterns of relations, whether for children or adults.

4.3 Implications, Limitations, and Directions for Future Research

The means by which separate episodes of experience are integrated with one another and by which the resulting novel representations are used productively to guide behavior are major questions in contemporary cognitive science and neuroscience (e.g., Bauer & Varga, 2017; Kumaran & McClelland, 2012; Preston & Eichenbaum, 2013; van Kesteren et al., 2012, 2014). One challenge in understanding these processes is that they span the seemingly great divide between the nominally separate mnemonic systems of episodic and semantic memory. Unique experiences presumably are encoded in episodic memory, a system specialized for preservation of patterns that differentiate one experience from another. Conversely, semantic memory is responsible for extraction of general patterns that give rise to representations that are timeless and placeless. Traditionally, rapid formation of episodic memories has been thought to be the purview of the hippocampus whereas the more gradual accumulation of general representations that make up semantic memory has been thought to be subserved by neocortex.

Results from memory integration paradigms in general, and the self-derivation through integration paradigm in particular, call this division of labor into question. They make it increasingly apparent that the hippocampus and prefrontal cortex are involved in interactive fashion to encode new experiences and seemingly simultaneously, extract generalizations from them (Schlichting, Mumford, & Preston, 2015; although see Varga & Bauer, 2017a, for evidence of temporal staging of the processes of encoding and extraction of relational meaning, even over rapid time scales). This process is readily apparent in the case of self-derivation of new factual knowledge through integration. Even young children encode new information on the basis of a single experience of it. They preserve the information such that they are able to recall it or recognize it later in the session, after a delay. Moreover, they use the information productively, to create novel fact representations that themselves are retained over time (e.g., Varga & Bauer, 2013; Varga et al., 2016; see Varga & Bauer, 2017b, for evidence of long-term retention of self-derived information in adults). Though the hippocampus seemingly plays a role in the self-derivation of the new knowledge—as evidenced by the structural relations observed in the present research—it does not seem reasonable to consider the new information an “episodic” memory, given that it features no spatial or temporal contextual tags (see Bauer & Jackson, 2015, for a similar argument). Similarly, though prefrontal cortex is implicated in gradual extraction of semantic knowledge—as evidenced by the structural relations observed in the present research—it also plays a role in rapid encoding of new factual knowledge on the basis of a single trial, a process typically ascribed to episodic memory. Based on these dual relations, it seems reasonable to argue either that the hippocampus and prefrontal cortex both subserve both episodic and semantic memory, or that the division itself is artificial.

The present research is not without limitations. One limitation is that one of the control variables, the estimate of verbal IQ, was obtained roughly 2 years prior to the assessment of self-derivation and the structural measures reported in this research. Although in future research, it will be desirable to have contemporaneous measures, we do not view this as a major impediment to interpretation of the present findings. This is because the verbal IQ measures used were the scaled scores from vocabulary subtests of the Wechsler scales (Wechsler, 2003, 2012), which provide stable assessments of verbal intelligence, even over delay intervals exceeding that in the present research (Watkins & Smith, 2013). Thus we can have confidence that the measure was valid, even if dated. A second limitation of the present research is that we tested children over the age range of 5 to 8 years and did not have sufficient power to test possible interactions with age. Interaction effects are not a significant concern in terms of the behavior of self-derivation through integration: prior research has revealed main effects of age, but no interactions. In contrast, measures of neural structure change in nonlinear as well as linear fashion, motivating future tests for interaction effects. In future research it will be especially important to include older children, in whom reductions in hippocampal volume and cortical thickness are to be expected, thus permitting more definitive tests of the pattern of relations between changes in structural volume and thickness and self-derivation of new factual knowledge through memory integration.

We also note that in light of evidence from the adult literature of the importance of correlated activity for successful memory integration and associative inference, the absence of measures of structural connectivity from the present research is salient. In future research, it will be necessary to assess structural connectivity as well as volume and thickness. Critically, the present research sets the stage for future efforts that should assess structural connectivity, as well as function and functional connectivity and patterns of activation more broadly. For example, it will be important to examine lateral aspects of the PFC, which have been implicated in memory processes such as semantic elaboration (e.g., Blumenfeld & Ranganath, 2007; Staresina, Gray, & Davachi, 2009). To date, there have been no neuroimaging studies that have measured memory integration of factual information, or the derivation of new factual knowledge from integrated representations. There is thus a great deal of work to be done before we understanding the neural bases of integration of separate episodes of learning of factual knowledge and subsequent derivation from it.

The present research also sets the stage for future work that bridges the scanner and laboratory and the classroom. There are many paradigms that assess memory integration. In the present research, a major motivation for focus on the process of self-derivation of new factual knowledge through integration of separate yet related episodes of new learning is that it serves as an ecologically valid model for accumulation of semantic knowledge. As such, the research stands to inform the neural bases of an empirically demonstrated, functionally significant learning mechanism. Consistent with this contention, self-derivation performance relates to academic achievement in reading and math (Esposito & Bauer, 2017; see Varga et al., 2018, for comparable findings with adults). A next step in the research process will be to determine whether the patterns of relation observed in the present research, between neural and laboratory measures, replicate when self-derivation of new factual knowledge through integration of separate yet related episodes of new learning is tested in classroom or other educational settings. Strictly behavioral classroom research largely replicates the patterns observed in the laboratory (e.g., Esposito & Bauer, 2017; in press). It will be left to future research to determine whether the structure/behavioral relations observed in the present research replicate as well.

4.4 Conclusions

The present research represents the first test of the neural bases of self-derivation of new factual knowledge through integration of separate yet related episodes of new learning in young children. We observed relations between the behavior and volume of the hippocampus as well as between behavior and mPFC thickness. Although some of the relations did not reach conventional levels of statistical significance, the overall pattern of findings implies that the hippocampus and prefrontal cortex both play crucial roles in rapid acquisition of information explicitly taught or learned, as well as in supporting extraction of semantic commonalities across separate episodes of experiences, thus permitting self-derivation of new knowledge acquired across temporally-extended episodes. Relative to other paradigms in common use in the literature (e.g., associative inference), the self-derivation paradigm provides especially strong evidence for this argument, given that the information over which memory integration and self-derivation must be performed is presented in a single trial, and is novel yet related to prior real-world factual knowledge.

5. Footnote

[1] When tested together in regression analyses along with verbal IQ, age and sex together accounted for 16% of the variance in ICV-adjusted right hippocampal volume, $F(3,62)=5.160$, $p<.01$. Only age explained significant variance (10%) in ICV-adjusted left hippocampal volume, $F(3,62)=3.634$, $p< .05$. Only sex and total gray matter volume explained significant variance (11%) in left mPFC, $F(4,61)=3.009$, $p< .05$. Only sex explained significant variance in predicting right mPFC (7%), however the total model was not statistically significant, $F(4,61)=2.265$, $p>.10$. Verbal IQ did not contribute significant variance in any of the models.

6. References

Bauer, P. J. (2012). Facilitating learning and memory in infants and young children: Mechanisms and methods. In S. L. Odom, E. P. Pungello, & N. Gardner-Neblett (Eds.), *Infants, toddlers, and families in poverty: Research implications for early child care* (pp. 49-74). New York, NY: The Guilford Press.

Bauer, P. J., Blue, S. N., Xu, A., & Esposito, A. G. (2016). Productive extension of semantic memory in school-aged children: Relations with reading comprehension and deployment of cognitive resources. *Developmental Psychology*, *52*(7), 1024-1037.

Bauer, P. J., & Jackson, F. L. (2015). Semantic elaboration: ERPs reveal rapid transition from novel to known. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(1), 271-282.

Bauer, P. J., & Larkina, M. (2017). Realizing relevance: The influence of domain-specific information on generation of new knowledge through integration in 4-to 8-year-old children. *Child Development*, *88*(1), 247-262.

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, *107*(4), 452-465.

Bauer, P. J., & Varga, N. L. (2017). Similarity and deviation in event segmentation and memory integration. [Peer commentary on the paper “Event perception: Translations and applications” by L.L. Richmond, D. A. Gold, & J. M. Zacks]. *Journal of Applied Research in Memory and Cognition*, *6*(2), 124-128.

Bauer, P. J., Varga, N. L., King, J. E., Nolen, A. M., & White, E. A. (2015). Semantic elaboration through integration: Hints both facilitate and inform the process. *Journal of Cognition and Development*, *16*(2), 351-369.

Benes, F. M. (2001). The development of prefrontal cortex: The maturation of neurotransmitter systems and their interactions. *Handbook of Developmental Cognitive Neuroscience*, *2*, 79-92.

Blumenfeld, R. S., & Ranganath, C. (2007). Prefrontal cortex and long-term memory encoding: An integrative review of findings from neuropsychology and neuroimaging. *Neuroscientist*, *13(3*), 280-291.

Brown, T. T., & Jernigan, T. L. (2012). Brain development during the preschool years. *Neuropsychology Review,* 22(4), 313-333. https://doi.org/10.1007/s11065-012-9214-1

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, *54*(1), 1-22.

Cohen, N. J., & Eichenbaum, H. (1993). Memory, amnesia, and the hippocampal system. *Cambridge, MA: MIT Press*, *3,* 378-389.

Cummings, G. (2009). Inference by eye: Reading the overlap of independent confidence intervals. *Statistics in Medicine, 28(2),* 205-220.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... & Albert, M. S. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31*(3), 968-980.

Esposito, A. G., & Bauer, P. J. (2017). Going beyond the lesson: Self-generating new factual knowledge in the classroom. *Journal of Experimental Child Psychology*, *153,* 110-125.

Esposito, A. G., & Bauer, P. J. (in press). Building a knowledge base: Predicting self-derivation through integration in 6- to 10-year-olds. *Journal of Experimental Child Psychology*.

Fischl, B. (2012). FreeSurfer. *Neuroimage, 62*(2), 774-781.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050-11055.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., ... & Montillo, A. (2002). Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, *33*(3), 341-355.

Frisoni, G. B., Jack, C. R., Bocchetta, M., Bauer, C., Frederiksen, K. S., Liu, Y., ... & Grothe, M. J. (2015). The EADC-ADNI Harmonized Protocol for manual hippocampal segmentation on magnetic resonance: evidence of validity. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *11*(2), 111-125.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170.

Ghetti, S., & Bunge, S. A. (2012). Neural changes underlying the development of episodic memory during middle childhood. *Developmental Cognitive Neuroscience*, *2*(4), 381-395.

Ghetti, S., & Lee, J. K. (2014). The development of recollection and familiarity during childhood: Insight from studies of behavior and brain. In P. J. Bauer & R. Fivush (Eds.), *The Wiley-Blackwell Handbook on the Development of Children's Memory* (pp. 309-335). West Sussex, UK: Wiley-Blackwell.

Ghosh, S. S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., ... & Fischl, B. (2010). Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4 to 11years of age. *NeuroImage, 53,* 85-93.

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., ... & Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. *Nature Neuroscience*, *2*(10), 861-863.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *PNAS, 101*, 8174-8179.

Gogtay, N., Nugent, T. F., Herman, D. H., Ordonez, A., Greenstein, D. K., Hayashi, K. M., et al. (2006). Dynamic mapping of normal hippocampal development. *Hippocampus, 16*, 664-672.

Goswami, U. (2002). Inductive and deductive reasoning. In U. Goswami (Ed.), *Blackwell handbooks of developmental psychology. Blackwell handbook of childhood cognitive development* (pp. 282-302). Malden: Blackwell Publishing.

Hebscher, M. & Gilboa, A. (2016). A boost of confidence: The role of the ventromedial prefrontal cortex in memory, decision-making, and schemas. *Neuropsychologia*, *9*0, 46-58.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*, *57*(5), 253-270.

Hu, S., Pruessner, J. C., Coupé, P., & Collins, D. L. (2013). Volumetric analysis of medial temporal lobe structures in brain development from childhood to adolescence. NeuroImage, 74, 276-287.

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*(2), 167-178.

Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proceedings of the National Academy of Sciences*, 201710654.

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological Review*, *119*(3), 573-616.

Kumaran, D., Summerfield, J. J., Hassabis, D., & Maguire, E. A. (2009). Tracking the emergence of conceptual knowledge during human decision making. *Neuron*, *63*(6), 889-901.

Lee, J. K., Nordahl, C. W., Amaral, D. G., Lee, A., Solomon, M., & Ghetti, S. (2015). Assessing hippocampal development and language in early childhood: Evidence from a new application of the automatic segmentation adapter tool. *Human Brain Mapping*. doi: 10.1002/hbm.22931

Manns, J. R., Hopkins, R. O., & Squire, L. R. (2003). Semantic memory and the human hippocampus. *Neuron*, *38*(1), 127-133.

Mishkin, M., Vargha-Khadem, F., & Gadian, D. G. (1998). Amnesia and the organization of the hippocampal system. *Hippocampus*, *8*(3), 212-216.

O'Kane, G., Kensinger, E. A., & Corkin, S. (2004). Evidence for semantic learning in profound amnesia: an investigation with patient HM. *Hippocampus*, *14*(4), 417-425.

Olson, I. R., & Newcombe, N. S. (2014). Binding together the elements of episodes: Relational memory and the developmental trajectory of the hippocampus. In P. J. Bauer & R. Fivush (Eds.), *The Wiley Handbook on the Development of Children's*

Memory, Volume 1/II, (pp. 285-308). Wiley-Blackwell.

Pfluger, T., Weil, S., Weis, S., Vollmar, C., Heiss, D., Egger, J., Scheck, R., & Hahn, K. (1999). Normative volumetric data of the developing hippocampus in children based on magnetic resonance imaging. *Epilepsia, 40*, 414-423.

Preston, A. R., & Eichenbaum, H. (2013). Interplay of hippocampus and prefrontal cortex in memory. *Current Biology, 23*(17), R764-R773.

Preston, A. R., Shrager, Y., Dudukovic, N., & Gabrieli, J. D. E. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus, 14,* 148-152.

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., … Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*(11), 1676–1689. https://doi.org/10.1093/cercor/bhi044

Riggins, T., Geng, F., Botdorf, M., Cox, L., & Hancock, G. R. (2018). Protracted hippocampal development is associated with age-related improvements in memory during early childhood. *NeuroImage, 174,* 127-137.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2016). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience, 29*(1), 37-51.

Schlichting, M. L., Mumford, J. A., & Preston, A. R. (2015). Learning-related representational changes reveal dissociable integration and separation signatures in the hippocampus and prefrontal cortex. *Nature Communications, 6*, 8151.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: Neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1-8.

Schlichting, M. L., Zeithamova, D., & Preston, A. R. (2014). CA1 subfield contributions to memory integration and inference. *Hippocampus*, *24*(10), 1248-1260.

Shing, Y. L., & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education*, *10(3)*, 153-161.

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W. (2004). Longitudinal mapping of cortical thickness and brain growth in normal children. *Journal of Neuroscience*, *24*(38), 8223-8231.

Staresina, B. P., Gray, J. C., & Davachi, L. (2009). Event congruency enhances episodic memory encoding through semantic elaboration and relational binding. *Cerebral Cortex*, *19*, 1198-1207.

Sweegers, C. C., Takashima, A., Fernández, G., & Talamini, L. M. (2014). Neural mechanisms supporting the extraction of general knowledge across episodic memories. *Neuroimage*, *87*, 138-146.

Szabó, C. A., Wyllie, E., Siavalas, E. L., Najm, I., Ruggieri, P., Kotagal, P., & Lüders, H. (1999). Hippocampal volumetry in children 6 years or younger: Assessment of children with and without complex febrile seizures. Epilepsy Research, 33(1), 1-9.

Tillman, R. M., Stockbridge, M. D., Nacewicz, B. M., Torrisi, S., Fox, A. S., Smith, J. F., & Shackman, A. J. (2017). Intrinsic functional connectivity of the central extended amygdala. *bioRxiv*.

Tulving, E. (1983). Ecphoric processes in episodic memory. *Philosophical Translations of the Royal Society of London*, *302*(1110), 361-371.

Uematsu, A., Matsui, M., Tanaka, C., Takahashi, T., Noguchi, K., Suzuki, M., & Nishijo, H. (2012). Developmental trajectories of amygdala and hippocampus form infancy to early adulthood in healthy individuals. *PloS ONE, 7(10):* e46970. doi.org/10.1371/journal.pone.0046970.

Utsunomiya, H., Takano, K., Okazaki, M., & Mitsudome, A. (1999). Development of the temporal lobe in infants and children: Analysis by MR-based volumetry. *American Journal of Neuroradiology, 20,* 717-723.

van Kesteren, M. T., Ruiter, D. J., Fernández, G., & Henson, R. N. (2012). How schema and novelty augment memory formation. *Trends in Neurosciences*, *35*(4), 211-219.

van Kesteren, M. T., R., Rijpkema M., Ruiter, D. J., Morris, R. G. M., &Fernández, G. (2014). Building on prior knowledge: Schema-dependent encoding processes relate to academic performance. *Journal of Cognitive Neuroscience, 26(10), 2250-2261.*

Van Petten, C. (2004). Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis. *Neuropsychologia*, *42*(10), 1394-1413.

Varga, N. L., & Bauer, P. J. (2013). Effects of delays on 6-year-old children's self-generation and retention of knowledge through integration. *Journal of Experimental Child Psychology*, *115*(2), 326-341.

Varga, N. L., & Bauer, P. J. (2017a). Using event-related potentials to inform the neurocognitive processes underlying knowledge extension through memory integration. *Journal of Cognitive Neuroscience*, *29*(11), 1932-1949.

Varga, N. L., & Bauer, P. J. (2017b). Young adults self-derive and retain new factual knowledge through memory integration. *Memory & Cognition*, *45*(6), 1014-1027.

Varga, N. L., Esposito, A. G., & Bauer, P. J. (2018). *Cognitive correlates of memory integration across development: Explaining variability in an educationally relevant phenomenon.* Manuscript under review.

Varga, N. L., Stewart, R. A., & Bauer, P. J. (2016). Integrating across episodes: Investigating the long-term accessibility of self-derived knowledge in 4-year-old children. *Journal of Experimental Child Psychology*, *145*, 48-63.

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, *277*, 376-380.

Wang, H., Das, S. R., Suh, J. W., Altinay, M., Pluta, J., Craige, C., … Yushkevich, P. A. (2011). A learning-based wrapper method to correct systematic errors in automatic image segmentation: Consistently improved performance in hippocampus, cortex and brain segmentation. *NeuroImage*, *55*(3), 968–985.

Watkins, M. W., & Smith, L. G. (2013). Long-term stability of the Wechsler Intelligence Scale for Children—Fourth Edition. *Psychological Assessment, 25,* 477-483.

Watson, C., Andermann, F., Gloor, P., Jones-Gotman, M., Peters, T., Evans, A., ... & Leroux, G. (1992). Anatomic basis of amygdaloid and hippocampal volume measurement by magnetic resonance imaging. *Neurology*, *42*(9), 1743-1743.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children, Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2012). *Wechsler Preschool and Primary Scale of Intelligence-Fourth Edition*. San Antonio, TX: The Psychological Corporation.

Weiss, A. P., Dewitt, I., Goff, D., Ditman, T., & Heckers, S. (2005). Anterior and posterior hippocampal volumes in schizophrenia. *Schizophrenia Research*, *73*(1), 103–112. doi.org/10.1016/j.schres.2004.05.018

Woolard, A. A., & Heckers, S. (2012). Anatomical and functional correlates of human hippocampal volume asymmetry. *Psychiatry Research: Neuroimaging*, *1(30*), 48-53.

Zeithamova, D., Dominick, A. L., & Preston, A. R. (2012). Hippocampal and ventral medial prefrontal activation during retrieval-mediated learning supports novel inference. *Neuron*, *75*(1), 168-179.

Zeithamova, D., & Preston, A. R. (2010). Flexible memories: Differential roles for medial temporal lobe and prefrontal cortex in cross-episode binding. *Journal of Neuroscience*, *30*(44), 14676-14684.

Table 1. *Descriptive Statistics for Behavioral (n = 68) and Neural (n = 66) Measures*

		Measure			
Behavior/Neural Measure		**Mean**	**SD**	**Min**	**Max**
Panel a: Behavioral Measures					
Integration facts (max=6)					
	Open-ended	2.35	1.69	0	5
	Total	4.63	1.30	1	6
Stem facts (max=12)					
	Open-ended	6.00	3.15	0	12
	Total	10.03	1.98	5	12
Panel b: Neural Measures					
Hippocampal volume (mm^3)					
	Left	3366.00	319.48	2370.06	4190.14
	Right	3311.30	287.00	2521.70	3817.22
mPFC thickness (mm^2)					
	Left	3.28	0.24	2.60	3.85
	Right	3.20	0.22	2.76	3.73

Table 2. Bivariate correlations between measures of interest.

		1	2	3	4	5	6	7	8	9	10	11
1. Age	r											
	p											
2. Sex	r	-.03										
	p	.83										
3. Verbal IQ (SS)	r	.14	.13									
	p	.27	.28									
4. Integration Facts—Open-ended	r	**.42****	.13	**.30***								
	p	**.<.001**	.29	**.01**								
5. Integration Facts—Total	r	**.26***	**.34****	.13	**.62****							
	p	**.03**	**<.01**	.29	**<.001**							
6. Stem Facts—Open-ended	r	**.54****	.16	**.32****	**.87****	**.60****						
	p	**<.001**	.19	**.01**	**<.001**	**<.001**						
7. Stem Facts—Total	r	**.44****	.09	**.26***	**.67****	**.61****	**.74****					
	p	**<.001**	.47	**.03**	**<.001**	**<.001**	**<.001**					
8. ICV-Adjusted Left hippocampal volume	r	*.22*	-*.23*	-*.23*	.05	-.12	.06	.18				
	p	*.08*	*.06*	*.06*	.67	.32	.62	.14				
9. ICV-Adjusted Right hippocampal volume	r	**.32****	**-.32****	.01	-.01	-.07	.01	.15	**.63****			
	p	**.01**	**.01**	.95	.92	.56	.94	.24	**<.001**			
10. Left mPFC thickness	r	-.05	**.27***	-.10	<.01	.06	-.11	**-.26***	-.07	-.02		
	p	.66	**.03**	.44	1.00	.62	.39	**.03**	.57	.87		
11. Right mPFC thickness	r	.02	.20	-.19	-.09	-.08	-*.21*	-*.22*	-.02	.01	**.69****	
	p	.84	.11	.13	.45	.52	.09	*.07*	.86	.92	**<.001**	

* $p \leq .05$, ** $p \leq .01$, ***$p \leq .001$

Table 3. *Regression Analyses Examining Relations between Performance on the Self-derivation through Integration Task and Hippocampal Volumes*

	Measure			
Predictor variables	Integration Facts		Stem Facts	
	Open-ended	Total	Open-ended	Total
	*β*s	*β*s	*β*s	*β*s
Age	.407***	.280*	.531***	.380**
Sex	.065	.301*	.086	.056
Verbal IQ	.282*	.046	.288**	.277*
Left Hippocampus	.203	-.108	.194	.249
Right Hippocampus	-.252†	.003	-.258†	-.115
Adj. R^2	.211	.123	.345	.212
$F(5,60)=$	4.473**	2.822*	7.845***	4.503***

† $p \leq .11$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, *ns* = not significant

Table 4. *Regression Analyses Examining Relations between Performance on the Self-derivation through Integration Task and mPFC Thickness*

	Measure			
Predictor	Integration Facts		Stem Facts	
variables	Open-ended	Total	Open-ended	Total
	βs	*βs*	*βs*	*βs*
Age	.378**	.373†	.500***	.395***
Sex	.125	.231**	.191	.133
Verbal IQ	.206	.031	.181	.167
Total Gray Matter Volume	.070	.196	.087	.044
Left mPFC	.136	.173	.090	-.196
Right mPFC	-.201	-.321†	-.313*	-.103
Adj. R^2	.182	.173	.358	.232
$F(6,59)$=	3.406**	3.269**	7.044***	4.282***

† $p \leq .07$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$, *ns* = not significant

Figure 1. Schematic representation of the structure of the paradigm used to test self-derivation through memory integration, from the stem-fact presentation through the test phase. In the test phase, children were asked forced-choice questions only for items they failed to answer correctly in the open-ended phase of testing.

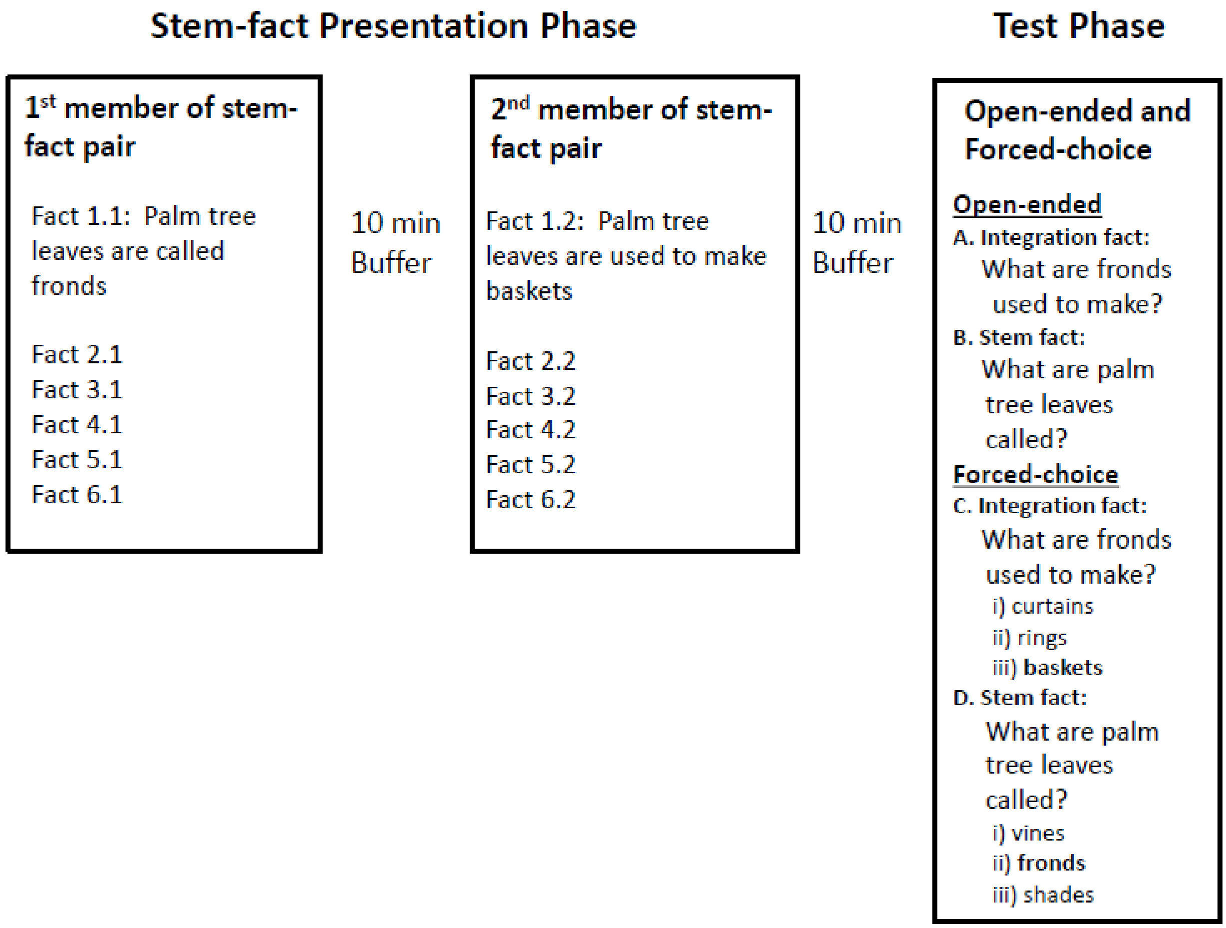
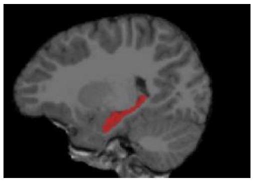


Figure 2. Partial regression plot showing association between right hippocampal volume and A) open-ended integration fact performance, and B) open-ended stem fact recall.



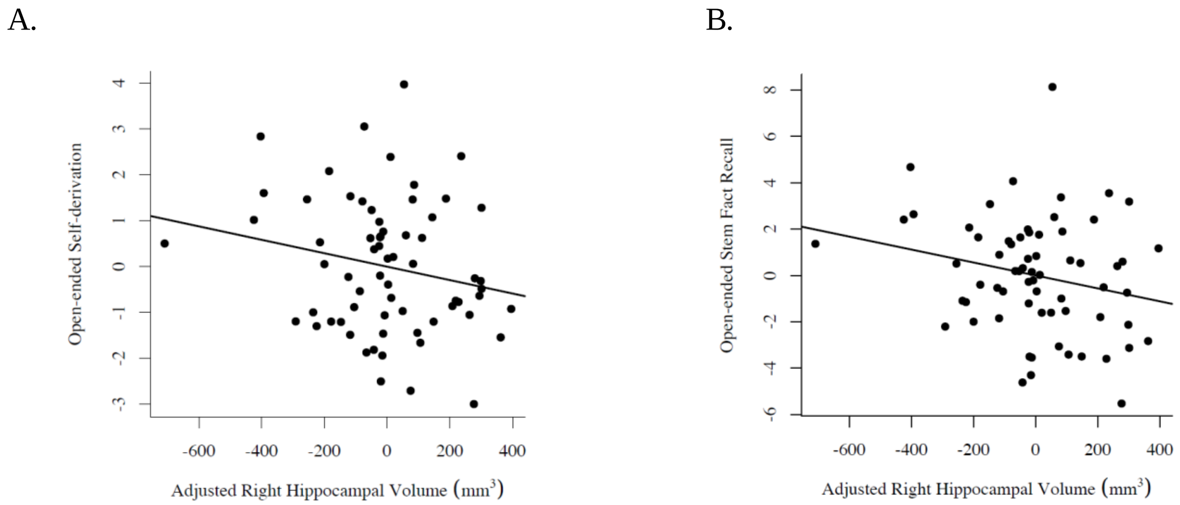


Figure 3. Partial regression plots showing associations between right mPFC thickness (green and yellow) and A) total integration fact performance (open-ended plus forced choice), B) open-ended stem fact recall; and associations between right rostral anterior cingulate thickness (yellow) and C) total integration fact performance (open-ended plus forced choice) and D) open-ended stem fact recall.