ELSEVIER

Neuropsychologia 42 (2004) 1394–1413

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Review

Relationship between hippocampal volume and memory ability in healthy individuals across the lifespan: review and meta-analysis

Cyma Van Petten*

Department of Psychology, University of Arizona, Tucson, AZ 85721, USA

Received 29 September 2003; received in revised form 7 April 2004; accepted 12 April 2004

Abstract

Poor memory ability and small hippocampal volume measurements in magnetic resonance images co-occur in neurological patients. Numerous studies have examined the relationship between memory performance and hippocampal volumes in participants without neurological or psychiatric disorders, with widely varying results. Three hypotheses about volume–memory relationships in the normal human brain are discussed: "bigger is always better", a neuropsychological view that volume decreases due to normal aging are accompanied by memory decline, and a developmental perspective that regressive events in development may result in negative correlations between hippocampal volume and memory ability. Meta-analysis of results from 33 studies led to little support for the bigger-is-better hypothesis. A negative relationship between hippocampal volume and memory (smaller is better) was significant for studies with children, adolescents, and young adults. For studies with older adults, the most striking observation was extreme variability: the evidence for a positive relationship between hippocampal size and episodic memory ability in older adults was surprisingly weak. Some of the variability in results from older adults was associated with statistical methods of normalizing for age and head size, which are discussed.

Keywords: Hippocampus; MR morphometry; Aging

1. Introduction

Most, if not all, of the cerebral cortex is likely to subserve the learning and retrieval of facts and events, with the engagement of specific regions dependent on the nature of the material learned (see Damasio, 1989; McClelland, McNaughton, & O'Reilly, 1995 for theoretical accounts, and Köhler, Moscovitch, Wincocur, Houle, & McIntosh, 1998; Nyberg, Habib, McIntosh, & Tulving, 2000; Senkfor, Van Petten, & Kutas, 2002; Wheeler, Peterson, & Buckner, 2000 for empirical reports). Despite general acceptance of the idea that memory is not localized to one neural structure, there is overwhelming evidence that the medial temporal lobe (MTL) plays a central role in episodic memory for diverse material, via its interactions with distributed cortical regions (Scoville & Milner, 1957). The MTL consists of multiple cortical areas, including the hippocampal formation, surrounding entorhinal, perirhinal, and posterior parahippocampal regions within the parahippocampal gyrus in humans, and part of the temporal pole (Insausti, Insausti, Sobreviela, Salina, & Martínez-Peñuela, 1998; Insausti, Juottonen, et al., 1998). Although many cases of human domain-general amnesia have damage to multiple regions of the medial temporal lobe, or to the diencephalon (Aggleton & Brown, 1999; Brown & Aggleton, 2001; Mayes, 2000; Smith & Bigel, 2000), lesions restricted to the hippocampus alone result in memory deficits in both humans and nonhuman primates (Rempel-Clower, Zola, Squire, & Amaral, 1996; Zola & Squire, 2000). The latter conclusion was drawn from postmortem examination of the affected brains, but the availability of high-resolution magnetic resonance images has led to a more widespread research effort to tie hippocampal pathology to memory impairment over the last decade.

MR scans are used to identify and/or confirm the location of damage after a frank neural insult such as stroke, hypoxia, or closed head injury. But an additional advantage of a noninvasive method is the possibility of identifying subtle neural damage when the insult is not as obvious. Recent studies have examined hippocampal volumes in individuals with childhood seizures, posttraumatic stress disorder, borderline personality disorder, depression, high risk of schizophrenia due to affected relatives, an ApoE-4 allele, and high estrogen levels (Bremner et al., 1995; Cohen, Small, Lalonde, Friz, & Sunderland, 2001; den Heijer et al., 2003; Driessen et al., 2000; Fennema-Notestine, Stein, Kennedy, Archibald,

* Fax: +1-520-621-9306.
E-mail address: vanpettc@u.arizona.edu (C. Van Petten).

0028-3932/$ – see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2004.04.006

& Jernigan, 2002; Lawson et al., 2000; O'Driscoll et al., 2001; Plassman et al., 1997; Seidman et al., 2002; Sheline, Sanghavi, Mintun, & Gado, 1999; Simpson, Baldwin, Burns, & Jackson, 2001; Stein, Koverola, Hanna, Rochia, & McClarty, 1997; VanLandingham, Heinz, Cavazoa, & Lewis, 1998). Particularly strong attention has focused on the possibility of detecting Alzheimer's disease before its clinical onset, when interventions might be more effective (see Chetelat & Baron, 2003; Kantarci et al., 2002; Wolf et al., 2003 for recent reviews).

Some clinical investigations have included measures of memory performance in additional to MR measures. When pathology is clearly present, these studies have largely succeeded in demonstrating quantitative relationships between hippocampal volumes and memory. Positive correlations between hippocampal volume and memory performance have been frequently reported in Alzheimer's patients, for instance, as well as in other varieties of dementia, and in amnesic patients of mixed etiology (Barber, McKeith, Ballard, Gholkar, & O'Brien, 2001; Cahn et al., 1998; de Toledo-Morrell et al., 2000; Deweer et al., 1995; Jernigan, Ostergaard, & Fennema-Notestine, 2001; Köhler, Black, et al., 1998; Kopelman et al., 2001; Mizuno, Wakai, Takeda, & Sobue, 2000; Mungas et al., 2002; Petersen et al., 2000; Wilson et al., 1996). Such correlations have also been observed in samples for which memory performance is at the borderline of clinical impairment, and a pathological process is suspected, as in older adults with mild cognitive impairment (Jack et al., 2000; Soininen et al., 1994).[1]

These clinical studies, and the expanding application of MR volumetry to diverse populations, raise a fundamental question: what is the relationship between hippocampal volume and memory ability in the normal brain? Given that the measured size of hippocampi varies a good deal across healthy individuals, and memory abilities also vary, is there a detectable structure–function relationship? When this question is posed bluntly, the response of many cognitive neuroscientists is likely to be "no", in part because macroscopic size is the crudest of neurobiological metrics, and in part because the question elicits memories of the failures of phrenology a century ago. In response, it might be argued that the cognitive psychology of the 21st century is more sophisticated than what the phrenologists had to work with, so that it may be more plausible to imagine size/function relationships for "episodic memory" than for the various "faculties" proposed by Gall and Spurzheim (but see Uttal, 2001 for a thoughtful and critical discussion of the history and current status of cognitive taxonomies and the limits of localization). More to the point, a survey of the literature indicates that a modern exercise in searching for size/function relationships is already underway. Numerous recent studies report significant correlations between hippocampal volumes and memory across individuals without a neurological or psychiatric diagnosis. However, numerous studies have also failed to find such relationships; these null results are less frequently cited, perhaps because they often occur in brief descriptions of results from a control group and are overshadowed by results from a patient group.

The discrepancy among published reports may result from random variability around a mean of zero correlation. Alternatively, both significant and null results across studies could arise from a subtle but genuine relationship. By analogy, the impact on cognitive function of white matter hyperintensities (WMH) observed in MR images from healthy older adults on cognitive function was subject to debate for some time, as the first several studies variously reported a negative relationship with cognitive performance (Rao, Mittenberg, Bernardin, Haughton, & Leo, 1989; Schmidt et al., 1991), or no detectable relationship (Hunt et al., 1989; Mirsen et al., 1991; Tupler, Coffey, Logue, Djang, & Fagan, 1992). In that case, pooling results from multiple studies to obtain a larger combined *N* was successful in demonstrating a moderate negative impact of WMHs on cognitive performance (Gunning-Dixon & Raz, 2000). The present paper evaluates whether pooling results from multiple studies in a meta-analysis will similarly show a relationship between hippocampal volumes and memory ability. Below, I first describe alternative hypotheses about this relationship, which guide the subsequent analyses.

2. Three perspectives on hippocampal size and memory ability in healthy subjects

2.1. Bigger is better

Three general hypotheses about the possible relationship between hippocampal volume and memory ability can be imagined. The first and simplest is the *bigger is better* (BIB) hypothesis, that regardless of the causal factors underlying the size of a structure, a larger structure should result in stronger function. This hypothesis seems to be implicit in some reports, but the investigators who have explicitly considered the BIB hypothesis have rejected it, because they observed significant negative correlations (smaller is better) between hippocampal volumes and memory performance in healthy young adults (Chantôme et al., 1999; Foster et al., 1999).

2.2. A neuropsychological perspective that includes aging

The second hypothesis can be called the *neuropsychological perspective*: that any normal size structure will support normal function, but that loss of tissue will lead to a decline

[1] It is important to note that in at least a majority of these studies, the positive correlations between hippocampal volume and memory performance were observed in analyses of the patient group alone, so that the results reflect more than a simple co-occurrence of smaller volumes and memory impairments in the patients as compared to controls (e.g., Barber et al., 2001; Cahn et al., 1998; de Toledo-Morrell et al., 2000; Köhler, Black, et al., 1998a; Kopelman et al., 2001; Laakso et al., 2000; Soininen et al., 1994).

in function. This perspective is strongly supported by positive correlations between hippocampal volume and memory performance in patients for whom the tissue loss is almost certainly due to pathology, and clinical memory impairment is evident, as noted above.

The challenge for the neuropsychological perspective lies in its treatment of normal aging. Some investigators have argued that even normal aging consists of a progressive loss of neural tissue, and an accompanying relentless decline in cognitive abilities (Terry & Katzman, 2001). In the absence of known pathology, there is general agreement that whole-brain weight (from postmortem examination) and volume (from MR studies) are lower in older adults than younger adults (Courchesne et al., 2000; Good et al., 2001; Jernigan et al., 2001; Mueller et al., 1998; Pfefferbaum et al., 1994; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003; see Coffey, 2000; Kemper, 1994; Powers, 2000; Raz, 2000 for reviews). The hippocampus is not spared from this general shrinkage. Estimates of age-related volume loss in this structure vary widely across studies, but nearly all report negative correlations between age and volume. In a fairly recent review, Raz (2000) estimated a median effect size of $r = -0.30$ across fifteen studies examining hippocampal volumes across the adult age range (ranging from -0.03 to -0.63). Studies not included in that review report correlations between age and hippocampal volume as nonsignificant (Gur et al., 2000), -0.08 (Hackert et al., 2002; Sheline et al., 1999), -0.24 (Van Petten et al., 2004), -0.35 (Tisserand, Visser, van Boxtel, & Jolles, 2000), -0.52 (Raz, Rodrigue, Head, Kennedy, & Acker, 2004), and -0.80 (Woodruff-Pak, Goldenberg, Downey-Lamb, Boyko, & Lemieux, 2000). (Discussion of the variability of these results is postponed for Section 5.)

If age-related shrinkage of the hippocampus is included within the neuropsychological perspective, we might expect to find positive correlations between hippocampal volume and memory performance in older participants, but perhaps not in younger participants, as in one report (Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998). If the neuropsychological perspective is extended to include normal aging, the second challenge will be to define when "aging" begins. In cross-sectional studies with a broad range of adult ages (twenties through eighties), a linear function is usually fitted to the age–volume relationship (Raz et al., 1997; Tisserand et al., 2000). In at least one recent report, decline in hippocampal size was precipitious across a fairly young adult age range of 18–42 years, at least in men (1.5% per year, $r = -0.46$, Pruessner, Collins, Pruessner, & Evans, 2001). In a different study of normal participants in their twenties, a repeated scan after a 2-year interval yielded measurements of the hippocampus plus amygdala that were 3.5% lower than the initial scan, although the statistical significance of this reduction was not noted (Lawrie et al., 2002). A second longitudinal study reports a significant volume loss of some 13% over an interval of about 2.5 years, for both men and women in their late twenties and early thirties (Lieberman et al., 2001). Finally, a third longitudinal study with repeated scans after a 5-year interval estimates 0.51% annual loss in participants under age 50 years, and 1.18% annual decline in hippocampal volume for those over age 50 years (Raz, Rodrigue, et al., 2004). Overall, existing data suggest that shrinking hippocampi may be a fact of adult life rather than senescence alone.

Although the apparently early onset of hippocampal shrinkage is startling, the time frame is not incompatible with recent estimates of the decline of episodic memory ability across the adult lifespan. From a cross-sectional sample of 345 participants aged 20–92 years, Park et al. (2002) report that both verbal and visuospatial long-term memory performance decline in a perfectly linear fashion from decade to decade. From a meta-analysis of studies including some 2700 participants, Verhaeghen and Salthouse (1997) instead conclude that age-related decline in episodic memory accelerates after age 50 years, yet still report a substantial decline in the 20–50-year age range.

The apparent parallel between declining hippocampal volumes and declining memory abilities across the adult lifespan argues for the inclusion of normal aging into the neuropsychological perspective on hippocampal volume loss and episodic memory. If we accept this conclusion, the prediction is that MR studies that include subjects older than some age at which both hippocampal volume loss and memory decline can be detected should also report significant positive correlations between volume and memory. The appropriate definition of "some age" will turn on practical factors that limit the power of any given study, such as sample size, the precision of the volume measurements, and the sensitivity of the memory tests administered. One advantage of a meta-analysis is that we can examine a prediction that is less vulnerable to limitations of sample size (at least): that correlations between hippocampal volume and memory should become increasingly positive as the age of the sample increases.

2.3. A developmental perspective

The third possible hypothesis about volume–memory relationships arises from a *developmental perspective*, and has received less attention in the literature than has aging. This perspective emphasizes the parallel between loss of cortical gray matter and improvement in cognitive abilities during childhood and adolescence. Although the whole brain may not reach its maximum size before the age of 15 years or so, cortical gray matter volumes (both absolute and as a proportion of whole brain) begin to decline in childhood or early adolescence, while white matter volumes increase. From repeated scans, Giedd, Blumenthal, et al. (1999) estimate that the gray matter decline begins by age 11 years in the frontal and parietal lobes, but not until about age 17 years in the temporal lobe. Other investigators estimate a somewhat earlier onset for gray matter decline, around age 9 years (Courchesne et al., 2000; Pfefferbaum et al.,

1994; Sowell, Trauner, Gamst, & Jernigan, 2002). The developmental decrease in gray matter measured in MR images has been attributed to regressive events observed in microanatomical investigations of both human and animal brains, which include loss of neurons, axonal branches, and synapses that are not adequately supported by functional use (i.e., pruning, as described in Cowan, Fawcett, O'Leary, & Stanfield, 1984; Huttenlocher, 1993; Rakic, Bourgeois, Eckenhoff, Zecevic, & Goldman-Rakic, 1986).

In contrast to the growth curves for whole-cortex or lobar gray matter, developmental studies using both postmortem and MR measures suggest that the hippocampus may reach maximum size by the age of 2 or 3 years, and show little absolute change (or a small increase) in the age range of 4–18 years (Giedd et al., 1996; Giedd, Jeffries, et al. 1999; Klekamp, Riedel, Harper, & Kretschmann, 1991; Kretschmann, Kammradt, Krauthausen, & Wingert, 1986; Pfluger et al., 1999; Utsunomiya, Takano, Okazaki, & Mistudome, 1999; Yurgelun-Todd, Killgore, & Cintron, 2003).[2] But because whole brain volume continues to increase until 12–15 years of age, the hippocampus and medial temporal lobe occupy a declining percentage of the brain across childhood development (Sowell et al., 2002; Utsunomiya et al., 1999). The prediction of the developmental perspective is thus that improvement in memory performance after age 3 years will be accompanied by decreasing hippocampal volumes, at least when those are expressed as a proportion of brain volume. The predicted correlations between volume and memory performance will be negative in direction.

Only two studies have explicitly addressed the developmental prediction above. For the 7–16 years age range, Sowell, Delis, Stiles, and Jernigan (2001) found the predicted negative relationship between memory performance and a medial temporal volume that included the hippocampus and parahippocampal gyrus. For a measure restricted to the hippocampal formation in 12–17 year olds, Yurgelun-Todd et al. (2003) reported a nonsignificant negative correlation.

2.4. Relationships among the three hypotheses

Because of its simplicity, the "bigger is better" hypothesis is incompatible with both the neuropsychological and developmental perspectives as outlined above. BIB theory predicts a linear relationship between hippocampal volume and memory ability regardless of whether tissue has been lost (the central tenet of the neuropsychological perspective), and independent of the changing relationship between hippocampal and brain volume across development (a core part of the developmental hypothesis). In contrast, the neuropsychological and developmental perspectives need not be orthogonal, in that different relationships between hippocampal volume and memory could obtain during different parts of the lifespan when the hippocampus is growing, stable, or declining in size. Combining the aging neuropsychological and developmental views is, however, nontrivial. Just as including aging in the neuropsychological perspective requires a definition of when aging begins, the developmental perspective will require some definition of when development (or the consequences of development) ends.

Below, the available evidence for the three hypotheses is evaluated, by quantitative examination of studies reporting relationships between hippocampal volumes and memory performance in participants with no known neuropathology or diagnosed memory disorder.

3. Methods

3.1. Search and inclusion criteria

Papers reporting relationships between hippocampal volumes and memory measures were located by a search of the Medline and PsycLit databases using the keyword combination (hippocampus OR hippocampal OR medial temporal) AND (memory OR recognition OR recall OR neuropsychological) AND (volume OR volumetric OR atrophy), and by examining the references of those papers. Results were included when they met all of the following criteria: (1) publication in English; (2) participants without a neurological or psychiatric diagnosis; (3) quantitative measures of hippocampal volumes (atrophy scores from visual ratings were excluded); (4) quantitative measures of recall or recognition from long-term memory (working memory measures were excluded); and (5) report of statistical relationships between volume and memory measures for the healthy participants separate from those of any patient group.[3] Thirty-three published papers met these criteria, listed in Table 1.

Across studies, the total number of subjects was 1880. For 11 of the studies, the healthy participants were controls for a patient group. Table 1 shows that five studies included

[2] Utsunomiya et al. (1999) reports a steep increase in hippocampal volume up to age 2 years, followed by a small gradual increase from age 2 to 13 years. Pfluger et al. (1999) show the same steep increase to about 3 years, then a flat function to age 15. In a sample of 4–18 year olds, Giedd et al. (1996) reports significant growth for the left hippocampus in girls, but no significant trend for the right hippocampus in girls, or either hippocampi in boys. In the 13–18-year age range, with repeated scans, Giedd, Blumenthal, et al. (1999a) report no change, as does Yurgelun-Todd et al. (2003) in a cross-sectional sample of 12–17 year olds. For a medial temporal region including the hippocampus, parahippocampal gyrus, uncal cortex, and amygdala, Sowell et al. (2002) report a nonsignificant decrease across the 7–16-year age range. Two postmortem studies estimate that maximum adult volume is reached at 2–3 years of age with no change thereafter (Klekamp et al., 1991; Kretschmann et al., 1986).

[3] One recent paper included demented participants, a cognitively normal group, and a group meeting criteria for mild cognitive impairment (MCI). Scatterplot figures in this paper suggest little relationship between hippocampal volume and memory performance within the cognitively normal group (Wolf et al., in press). However, correlations were not reported for the cognitively normal subjects separate from those with MCI (these were collapsed into a "nondemented" group), so that the results were not included in the meta-analysis.

Table 1
Papers included in the meta-analysis of hippocampal volumes and long-term memory tests in healthy subjects

Study (notes)	*N*	Mean age (range)	Normalization	Memory tests (number in composite)	Correlation
Golomb et al. (1994)	54	69.0 (55–87)	Residual after cranial vault and age	Immediate recall composite (6) Delayed recall composite (6)	0.10 0.46*
Sullivan et al. (1995)	72	43.9 (21–70)	Residual after cranial vault	Immediate verbal recall Delayed verbal recall Immediate nonverbal recall Delayed nonverbal recall	0.255* 0.18 −0.19 −0.255*
Torres et al. (1997)[a]	19	29.2 (S.D. 8.5)	Residual after height	Delayed recall composite (3)	ns
Cahn et al. (1998)	20	68.4 (S.D. 5.4)	Residual after cranial vault and age	Warrington recognition memory for words Warrington recognition memory for faces	ns ns
Köhler, Black, et al. (1998)	26	70.8 (S.D. 6.3)	None	CVLT, immediate recall CVLT, delayed recall WMS-R, Visual Reproduction 1 WMS-R, Visual Reproduction 2	−0.39* −0.55* −0.13 −0.11
Lupien et al. (1998)[b]	9	73.9 (S.D. 5.6)	Divide by cranial vault	Delayed recall of line drawings	0.55*
Raz et al. (1998)[c]	95	44.0 (18–77)	Residual after height	Verbal memory composite (5) Nonverbal memory composite (2)	0.13 0.11
MacKay et al. (1998)[a]	11	21.7 (college)	Divide by hemisphere	Face recognition	−0.07
MacKay et al. (1998)[a]	18	21.7 (college)	Divide by hemisphere	Face recognition	0.015
Reiman et al. (1998)	33	nr (50–62)	Divide by cranial vault	Rey AVLT, delayed recall	0.38*
Chantôme et al. (1999)	70	24.5 (18–31)	Divide by brain volume	Stem-cued recall	−0.26*
Foster et al. (1999)[a]	18	19.8 (18–30)	Divide by hemisphere	Delayed story recall	−0.45*
Visser et al. (1999)	18	76.8 (S.D. 4.0)	Residual after cranial vault	CAMCOG memory score	ns
de Toledo-Morrell et al. (2000)	30	72.4 (64–84)	Divide by cranial vault	Verbal recall Delayed verbal recall Spatial recall Delayed spatial recall	ns ns ns ns
Gur et al. (2000)	110	26.1 (18–45)	Residual after cranial vault	Verbal memory composite (3) Nonverbal memory composite (2)	0.15 0.36*
Laakso et al. (2000)	34	72.0 (64–79)	Residual after cranial vault	Russell Visual Reproduction Test	ns
Petersen et al. (2000)	126	79.1 (nr)	Residual after cranial vault	WMS-R, Logical Memory (2) WMS-R, Visual Reproduction (2) Rey AVLT Buschke Selective Reminding (3)	ns ns ns ns
Tisserand et al. (2000)	61	55.7 (21–81)	Residual after cranial vault, sex, education	Word list immediate recall Word list delayed recall	0.00 0.00
Woodruff-Pak et al. (2000)	7	82.5 (S.D. 5.8)	Divide by cranial vault	Buschke Selective Reminding, delayed recall	0.53
Cohen et al. (2001)	25	57.0 (S.D. 8)	None	WMS-R indexes (5) Buschke Selective Reminding (3) Rey figure recall	−0.02 mean of all tests, range −0.25 to 0.20
O'Driscoll et al. (2001)[d]	14	35.4 (S.D. 8.8)	Residual after sex. Memory residual after sex, SES, IQ	WMS-R, Logical memory 1 WMS-R, Logical memory 2	ns 0.56 (Spearman)
Sowell et al. (2001)[b,c]	35	11.9 (7–16)	Residual after brain volume	CVLT, delayed free recall, Rey–Osterrieth, delayed recall	−0.22* −0.53*

Table 1 (*Continued*)

Study (notes)	*N*	Mean age (range)	Normalization	Memory tests (number in composite)	Correlation
Hackert et al. (2002)	511	73.0 (60–90)	Residual after cranial vault and age	Word list immediate recall	0.12*
				Word list delayed recall	0.14*
MacLullich et al. (2002)[e]	95	nr (65–70)	Residual after cranial vault	Benton Visual Retention	ns
				WMS-R, Visual Reproduction (2)	ns
				Rey AVLT	ns
				WMS-R, Logical Memory (2)	ns
Marquis et al. (2002)	60	83.2 (S.D. 7.9)	None	WMS-R, Logical Memory 2	ns
Sanfilipo et al. (2002)	27	35.7 (S.D. 8.7)	Residual after cranial vault and age	Buschke-Selective Reminding (4)	−0.30
				WMS-R, Visual Reproduction (2)	ns
Seidman et al. (2002)	47	40.1 (20–68)	None	WMS-R, Logical Memory 1	0.27
				WMS-R, Logical Memory 2	0.26
Convit et al. (2003)	30	68.6 (53–89)	Divide by cranial vault. Memory residual after age and MMSE	WMS-R, Logical Memory 1	nr
				WMS-R, Logical Memory 2	0.49*
Maguire et al. (2003)	26	25.8 (18–56)	Residual after age	Environmental scene recognition	ns
				AMIPB, delayed story recall	ns
Yurgelun-Todd et al. (2003)	37	14.6 (12–17)	Divide by brain volume	WAIS, Digit-Symbol delayed recall	−0.125
Driscoll et al. (2003)[f]	16	26.1 (20–39)	None	Virtual Morris water maze	0.44*
	16	77.6 (60–85)		Transverse patterning discrimination	0.44*
Rosen et al. (2003)[a]	14	69.5 (all elderly)	Divide by cranial vault	Paragraph recall, immediate	0.515*
				Paragraph recall, delayed	0.485
				Word list recall, immediate	0.315
				Word list recall, delayed	0.195
Rodrigue & Raz (2004)	48	57.6 (26–82)	Residual after cranial vault	Associative memory, immediate (2)	0.33*
				Associative memory, delayed (2)	ns
				Free recall, immediate (2)	ns
				WMS-R, Logical Memory 2	ns
Van Petten et al. (2004)[a,g]	48	73.2 (65–85)	Residual after cranial vault	CVLT, long delay cued recall	−0.18
				WMS-3, Logical Memory 1	−0.14
				WMS-3, Logical Memory 2	−0.11
				WMS-3, Verbal Paired Associates 1	−0.09
				WMS-3, Verbal Paired Associates 2	−0.16
				WMS-R, Visual Paired Associates 1	−0.21
				WMS-R, Visual Paired Associates 2	−0.11
				WMS-3, Face Recognition 1	0.21

Note: Correlations significant in original reports marked with *; “ns” means that relationships were reported only as nonsignificant; “nr” means “not reported”. Correlations are Pearson unless noted. Age ranges of the sample are noted when available, otherwise the standard deviation (S.D.) of the mean age. Tests are combined in the manner reported by the original authors, with number of tests within a composite measure in parentheses. For tests with both an immediate and a delayed version, “(2)” means that both immediate and delayed versions were administered. Note that only tests of long term memory are listed here; original reports may include additional tests. AMIPB: Adult Memory and Information Processing Battery; CAMCOG: cognitive portion of the Cambridge Examination of Mental Disorders of the Elderly; CVLT: California Verbal Learning Test; Rey AVLT: Rey Auditory–Verbal Learning Test; MMSE: Mini-Mental State Exam; WAIS: Wechsler Adult Intelligence Scale; WMS-R: Wechsler Memory Scale, Revised; WMS-3: Wechsler Memory Scale 3. Many studies use experimenter-designed memory tests and/or composite measures based on multiple neuropsychological tests, see original papers for details.

[a] Participants selected from a larger sample as having high or low memory performance.

[b] Delayed recall measure corrected for immediate recall by subtraction or regression.

[c] Combined measure of hippocampus and parahippocampal gyrus.

[d] Combined measure of anterior hippocampus and amygdala.

[e] Significant positive correlations before normalization for intracranial vault size.

[f] Significant positive correlations also reported for unspecifed method of normalizing for cranial vault size.

[g] Original report included correlations between raw hippocampal volumes and memory performance, and partial correlations after regressing on cranial vault size and age. Correlations here are normalized only for cranial vault size.

only children, adolescents, or young adults, with $N = 189$ (Chantôme et al., 1999; Foster et al., 1999; MacKay et al., 1998; Sowell et al., 2001; Yurgelun-Todd et al., 2003); 17 included only older adults, with $N = 1140$ (Cahn et al., 1998; Cohen et al., 2001; Convit, Wolf, Tarshish, & de Leon, 2003; de Toledo-Morrell et al., 2000; Golomb et al., 1994; Hackert et al., 2002; Köhler, Black, et al., 1998; Laakso, Hallikainen, Hänninen, Partanen, & Soininen, 2000; Lupien et al., 1998; MacLullich et al., 2002; Marquis et al., 2002; Petersen et al., 2000; Reiman et al., 1998; Rosen et al., 2003; Van Petten et al., 2004; Visser et al., 1999; Woodruff-Pak et al., 2000); four included primarily subjects in their thirties or forties, with $N = 170$ (Gur et al., 2000; O'Driscoll et al., 2001; Sanfilipo et al., 2002; Torres, Flashman, O'Leary, Swayze, & Andreason, 1997); and seven included a broad range of adult ages, with $N = 381$ (Driscoll et al., 2003; Maguire et al., 2003; Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998; Rodrigue & Raz, 2004; Seidman et al., 2002; Sullivan, Marsh, Mathalon, Lim, & Pfefferbaum, 1995; Tisserand et al., 2000).

3.2. Statistical methods

Most of the included papers reported volume–memory relationships as Pearson correlations (r); these were converted to z-scores by the Fischer transform (Z_r). The single Spearman's *rho* was treated in the same way (as recommended by Hunter & Schmidt, 1990). For two papers (Driscoll et al., 2003; Hackert et al., 2002), probability levels and sample size were estimate Z_r. Relationships reported as not significant without numerical correlations were coded as correlations of zero. When results were reported separately for men and women, or for right and left hippocampi, arithmetic averages of the correlations were computed before transformation to Z_r. (Too few studies report separate correlations by gender or side for this information to be used.) When correlations were reported for more than one memory test, the arithmetic average of the correlations was computed (immediate and delayed memory tests are considered separately in an analysis of possible moderator variables). One paper reported two studies from the same population, with the same memory test (MacKay et al., 1998); these results were combined (weighted by sample size) prior to analysis.

Most of the papers report volume–memory relationships only after some form of normalization for head or body size: division by whole-brain volume, ipsilateral hemisphere, or intracranial vault size, or residual volume after regressing on brain volume, intracranial vault, or height (method listed in Table 1). When papers included correlations for both raw and normalized volumes, normalized correlations were selected in order to conform to the larger number of papers that did not offer a choice of measures. Six papers report partial correlations between volume and memory only after additionally regressing on the age of the subjects (Cahn et al., 1998; Convit et al., 2003; Hackert et al., 2002; Maguire et al., 2003; Sanfilipo et al., 2002). A priori, it seems that this procedure might reduce volume–memory correlations if both hippocampal volume and memory ability decline across the age range included in the sample. Analyses were conducted both with and without these studies.

The common Z_r across studies were computed as $\sum Z_r(n-3)/\sum(n-3)$, where n is the sample size for each study. The common Z_r was then transformed to estimate the common r across studies (Hedges & Olkin, 1985). The Z_r transform produces a small bias toward larger (positive or negative) correlation estimates, so that some authors recommend using a correction factor, or indeed using the raw r's instead (Edwards, 1963; Hunter & Schmidt, 1990). These different procedures yielded estimates of the common r that were identical to the second decimal place, so that the formula above was used. The significance of the common r was then evaluated by

$$Z_{\text{test}} = (\text{common } Z_r) \times (\text{square root of}(N - 3k)),$$

where N is the total number of subjects across studies, and k is the number of studies (Hedges & Olkin, 1985).

For the entire set of studies, and for selected subsets, tests of heterogeneity were conducted to evaluate the null hypothesis that the set of correlations were drawn from the same underlying population, using the X^2 statistic as recommended by various authors (Edwards, 1963; Hedges & Olkin, 1985). To compare correlations across different subsets of studies, the statistic

$$Z_{\text{comparison}} = \frac{Z_{r1} - Z_{r2}}{\text{square root of}((1/N_1 - 3) + (1/N_2 - 3))}$$

was used (Wolf, 1986).

4. Results

4.1. Common correlation across studies, and heterogeneity of the results

The common correlation between hippocampal volume and memory performance across all 33 studies (weighted by sample size) was 0.071 ($Z_{\text{test}} = 3.00$, $P < 0.005$). Although the estimated correlation across studies was significantly positive, Table 1 shows a great deal of variability across studies: reported correlations range from positive 0.55 to negative 0.55. Fig. 1 shows the distribution of reported correlations, which appears to deviate from normal. The heterogeneity test confirmed the visual impression that the reported correlations are unlikely to come from the same underlying distribution ($X^2 = 63.5$, d.f. $= 32$, $P < 0.001$).

The initial reasoning above was that partialling out the effect of age might reduce the magnitude of the volume–memory correlation, because both hippocampal volume and memory performance are typically lower in elderly than young subjects. However, three of the six studies that used this procedure reported substantial positive correlations, so that excluding all six resulted in a nonsignificant

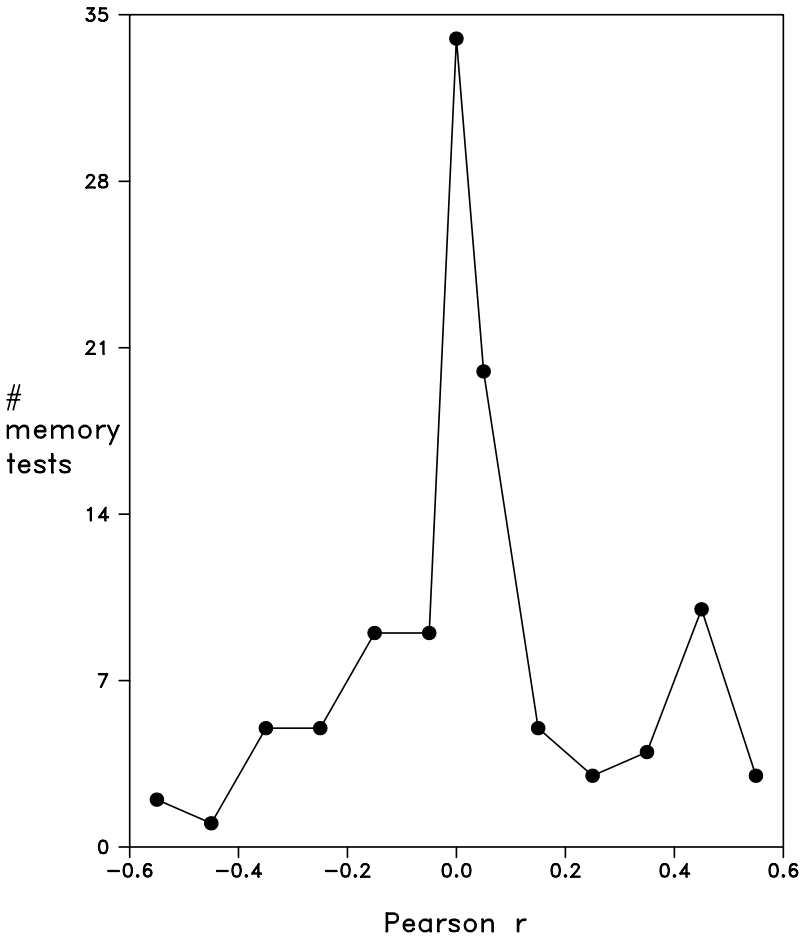

Fig. 1. Frequency distribution of reported correlations between hippocampal volume and the 107 memory tests used across the studies. Correlations between 0.60 and 0.50 were collapsed, and plotted as 0.55; correlations between 0.50 and 0.40 collapsed and plotted as 0.45, etc. When correlations for a composite of several memory tests were reported as a single *r*, the reported *r* was counted as occurring several times (equal to the number of tests in the composite). Correlations reported only as nonsignificant were counted as $r = 0$.

estimate of the common *r* as 0.035 ($Z_{\text{test}} = 1.17$). The heterogeneity test remained significant for the remaining 27 studies ($X^2 = 47.0$, d.f. $= 26$, $P < 0.01$).

The heterogeneity tests (as well as even a cursory inspection of Table 1), suggest that some factor may moderate the relationship between hippocampal volume and memory performance. Below I examine three candidates: the nature of the memory measures, the age of the subjects, and the method used to normalize hippocampal volume for head or body size.

4.2. Is nature of the memory test a moderator variable?

Table 1 indicates that a variety of memory tests were used across the different studies, although tests from the Wechsler Memory Scale Revised (WMS-R) are strongly represented. Because of this variety, and the frequent combination of individual tests into composite measures, it is not possible to examine whether some particular memory tests are more strongly associated with hippocampal volumes than others. However, the WMS-R and some other standard neuropsychological batteries have both immediate memory tests, and delayed versions in which the subject attempts to recall the same material some twenty minutes later. Some investigators have suggested that delayed tests show a stronger relationship to hippocampal volume than those in which recall follows immediately after the study phase (see Köhler, Black, et al., 1998 for an example in Alzheimer's patients). In the first published study of volume–memory relationships in a nondemented sample of older adults, Golomb et al. (1994) observed a significant correlation for delayed but not immediate recall. Following this lead, other investigators have used a retention measure of delayed minus immediate recall (Lupien et al., 1998; Sowell et al., 2001).

Despite the potential importance of this issue, it is not simple to compare the results of immediate to delayed memory tasks across the published studies, for several reasons: (1) independent scores for immediate and delayed versions cannot be estimated in those studies which report only a retention measure; (2) some of the test composites used by

various authors include both immediate and delayed tests, with results reported only for the composite; and (3) some studies which included both immediate and delayed tests report both sorts of correlations only as "nonsignificant". Paring down the reported results to only those studies in which correlations are reported for individual memory tests, or for test composites consisting of only one sort of test yielded eleven studies with immediate tests, and twenty with delayed tests. Weighted by sample size, the estimated common r is 0.086 ($P < 0.02$) for the immediate tests, and 0.090 ($P < 0.005$) for the delayed tests (not significantly different). Although based on only a subset of the published results, this comparison provides little support for the idea that delayed tests have a stronger association to hippocampal volume than immediate tests in healthy individuals.

4.3. Age of the subjects as a moderator variable

The age of the participants is critically important for both the aging neuropsychological and developmental perspectives. The neuropsychological hypothesis predicts positive correlations between hippocampal volume and memory in older adults, after age-related atrophy has resulted in memory impairment, but no relationship or a smaller positive correlation in young adults. The developmental hypothesis predicts negative correlations in children and possibly in young adults, and is agnostic about older adults. The two hypotheses converge in predicting increasingly positive correlations with advancing age. The ideal method of examining this prediction would consider volume, memory performance, and age for each participant in the 33 studies. In contrast, mean age of the sample is what is available from published reports, but examination of the reported volume–memory correlations by mean age should provide at least a first look at the prediction.

Fig. 2 plots average correlations from each study (zero for those reported only as nonsignificant) against the mean age of the subjects. Some tendency for the volume–memory correlations to become more positive with increasing age can be discerned, but this is disrupted by the large degree of scatter at the elderly end of the continuum. A linear regression on the corresponding Z_r values from each study against age, weighted by sample size, yielded a marginal result (weighted $r = 0.23$, $P = 0.10$). A log transform of age produced a somewhat stronger relationship with the volume–memory correlation (weighted $r = 0.30$, $P < 0.05$).

The potential impact of age was pursued by examining subsets of studies from the two age extremes. Three studies included only adults under the age of 30 (Chantôme et al., 1999; Foster et al., 1999; MacKay et al., 1998), and two studies included only children and adolescents (Sowell et al., 2001; Yurgelun-Todd et al., 2003). Considered together, these five studies have a common r of negative 0.25 ($Z_{\text{test}} = 3.09$, $P = 0.001$, weighted by sample size). The heterogeneity test for these studies (the squares in Fig. 2) was nonsignificant ($X^2 = 3.32$, d.f. $= 4$, $P > 0.40$),

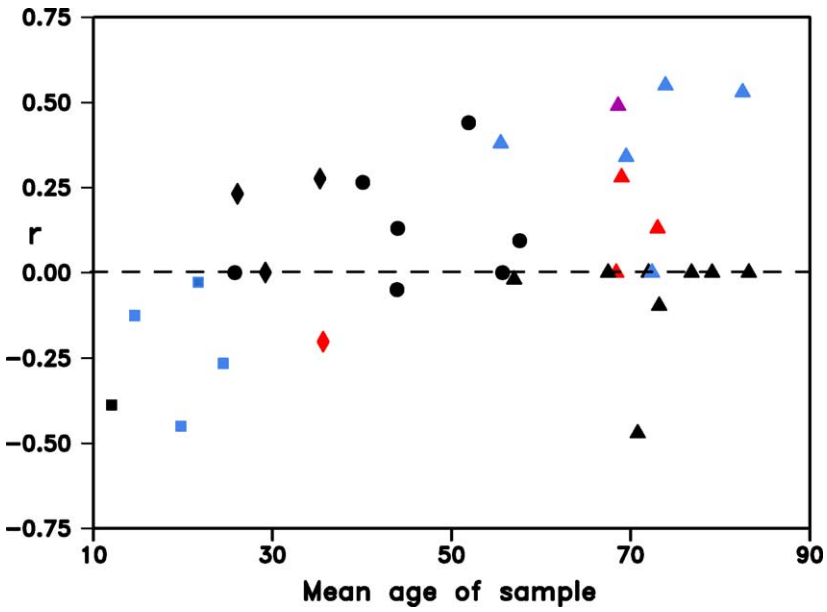

Fig. 2. Correlations between hippocampal volume and memory performance are plotted against mean age of the subject sample (one correlation for each study, averaged across all memory tests in the study). Squares denote studies including only young adults (<30 years), children, or adolescents; diamonds denote studies with adults in their thirties to forties; triangles denote studies with older adults (>50 years); circles denote studies with a broad range of adult ages. Black symbols are for studies using raw volumes or residual volumes after removing variance associated with head or body size. Blue symbols denote studies in which hippocampal volumes were expressed as a proportion of intracranial vault or brain volume. Red symbols denote studies that use age-adjusted residual scores for hippocampal volume or memory performance. Purple indicates age-adjusted residuals for memory performance, and hippocampal volume expressed as a proportion of intracranial volume. Note that the graph does not display the number of participants in each study (which varied widely), although the statistical analyses in the text are weighted by sample size.

suggesting that they form a coherent cluster drawn from the same underlying distribution.

Seventeen studies included only older adults (all over 50, triangles in Fig. 2), with an estimated common r of positive 0.097 ($Z_{test} = 3.20$, $P < 0.001$). The heterogeneity test was marginal ($X^2 = 22.8$, d.f. $= 16$, $P = 0.12$). The estimated common r's for the studies with young versus elderly participants were significantly different from one another ($Z_{comparison} = 4.24$, $P < 0.0001$).

One concern about the validity of the common correlation for the elderly end of the age range is that one study (Hackert et al., 2002) contributes nearly half of the total N (511 of the 1120 older adults), so that the meta-analysis might be largely equivalent to the single large study. If this study is excluded, the other 16 studies have a marginal common r of 0.063 ($Z_{test} = 1.50$, $P = 0.07$, X^2 for heterogeneity $= 21.2$, d.f. $= 15$, $P > 0.15$). The large-sample study was one of four with elderly participants that used a measure of hippocampal volume or of memory performance residualized on age; excluding all four studies results in a nonsignificant common r across the remaining thirteen studies ($r = 0.024$, X^2 for heterogeneity $= 10.5$, ns). Both of the estimated r's after excluding some elderly studies remain significantly different from the estimated common r of -0.25 from the young studies ($Z_{comparison} > 2.88$, $P = 0.002$).

4.4. Is the method of normalizing for head size a moderator variable?

Head size is correlated with body size, and brain size in turn correlated with head size. Because most investigators share the intuition that large people are not invariably superior to small people in cognitive abilities, morphometry studies frequently include some sort of correction factor for overall differences in head or body size prior to examining the relationship between the volume of any given brain structure and cognitive performance. Table 1 shows a wide variety of normalization methods which may contribute to the variety of reported results. These include: (1) analysis of raw hippocampal volumes; (2) volume divided by brain volume; (3) volume divided by vault size; (4) removal of variance shared with brain volume via regression; (5) removal of variance shared with vault size via regression or ANCOVA; and (6) removal of variance shared with height via regression or ANCOVA. (In Table 1, the regression and ANCOVA methods are both referred to as "residual" methods.)

For analysis, I collapse these six methods into three: (1) *raw* volumes (or residuals after removing variance due to age or sex, but not head or body size); (2) *division* by vault or brain; and (3) *residuals* after height, vault, or brain volume. This particular scheme focuses on the statistical method of correction, while ignoring the nature of the correction factor itself. The argument for combining analyses using intracranial vault size, brain size, and height as the correction factors is that these can be expected to be strongly correlated with one another. Note that vault and brain volume can become dissociated as total brain volume begins to decline in late life; however, the studies which use brain volume as the correction factor were all conducted in participants under the age of 30, in which total brain volume and vault size will be closely related. Thus, while there may be subtle differences dependent on the particular correction factor applied (Peters et al., 1998; Raz, Gunning-Dixon, et al., 2004), these are likely to be less important than the major differences in statistical method.

Table 2
Results split by method of head-size correction

Method	Number of studies	N	Common r	Heterogeneity X^2
All studies				
Raw	7	230	0.079	15.3*
Division	10	277	0.025	26.6***
Residual	16	1373	0.081***	18.0
Older adults				
Raw	3	111	−0.119	4.5
Division	6	123	0.363****	17.5****
Residual	8	926	0.087***	6.8
Older adults without age normalization				
Raw	3	111	−0.119	4.5
Division	5	93	0.290*	9.7*
Residual	5	321	−0.014	0.4

* $P < 0.05$.
*** $P < 0.005$.
**** $P < 0.001$.

The top section of Table 2 shows the results of all the studies split into three sets based on normalization method. Although only the common correlation from studies using the residual method reached statistical significance (because of the larger N), there were no statistically significant differences among the three estimated correlations ($Z_{comparison} <$ 1 for all three paired comparisons).

4.4.1. Age and head-size correction

It was not possible to examine the impact of normalization method for studies with only young participants, as most of these (four of five) used the division method. Studies restricted to older adults have, however, used all three methods, so that the middle section of Table 2 compares these. The method of removing variance in hippocampal size associated with head or body size via regression or ANCOVA ("residual method") yielded a small positive correlation between hippocampal volume and memory performance, much like the overall analysis of older participants. The method of dividing hippocampal volume by cranial vault yielded a more substantial positive correlation, but with a significant degree of heterogeneity among studies as well. Finally, given the impact of normalizing results by age noted above, the bottom section of Table 2 includes results from only those studies of older adults that did not adjust hippocampal volumes or memory scores for age. In this last case, only those studies using the division method yielded

a statistically significant common correlation between hippocampal volume and memory performance (and a significant heterogeneity test as well). The positive correlation for studies using the division method was significantly different from the estimated correlations using the other two methods (division versus residual: $Z_{comparison} = 2.39$, $P < 0.05$; division versus raw: $Z_{comparison} = 3.58$, $P < 0.001$). The topic of how division by vault size differs from adjustment based on residuals is reserved for Section 5.

5. Discussion

5.1. Limitations of meta-analyses

A potential hazard of any meta-analysis is the so-called "file drawer problem": studies that are unpublished due to null or difficult-to-interpret results. Some protections against this problem are built into the present topic. The first stems from the expense, effort, and relative novelty of relating structural MR data to cognition, so that investigators are likely to be strongly motivated to publish, and editors likely to show a similar interest. The second protection applies to a subset of the studies in which the healthy participants were controls for neurological or psychiatric patients (11 of the 33 studies): in these cases the investigators, reviewers, and editors may have had a relatively minimal theoretical stake in the direction of these results. Whatever the reasons, Table 1 indicates that results of all flavors have reached print, so that the file-drawer problem is likely to be minimal here. Table 1 does, however, show a diluted version of the problem: volume–memory correlations reported only as "nonsignificant" without numbers. These were coded as correlations of zero in the analyses because an effect size of zero is, in general, the average null result. Fig. 2 indicates that removal of all the "ns" results would not influence the young end of the age range, and would continue to show substantial variability for the elderly end of the range. To the extent that there is some bias in reporting the details of nonsignificant results, it is likely that the a priori logic of the aging-neuropsychological hypothesis would lead to under-reportage of negative volume–memory correlations for older participants. I would thus argue that if coding the "ns" correlations as zero introduced a bias in the meta-analysis, it would be in the direction of a too-positive estimate of the volume–memory correlation in older adults.

Another general feature of meta-analytic methods is that qualitative aspects of data are disregarded in favor of those that can be quantified or categorized. In the present analyses, the quantitative variables were the number of participants and their mean age; categorical variables were immediate versus delayed memory tests, and the method of handling variation in head or body size. However, it is possible to imagine that other variables resulted in measurements of hippocampal volume that were more or less accurate. These include: (1) imaging parameters such as the pulse sequence, slice thickness, gap (or no gap) between slices, slice angle relative to orientation of the structure, rotation of slices prior to measurement, method for differentiating tissue from cerebrospinal fluid, etc.; (2) the measurement protocol itself, namely the anatomical landmarks used to define the hippocampus; and (3) the skill and care of the operators in following the measurement protocol. Across studies, the Methods sections show some variation in imaging parameters, but there was no obvious way of capturing this variation via categorical variables. Similarly, the measurement protocols depended, in part, on the slice angle and thickness, although visual identification of the hippocampus is not especially difficult and the published reports generally refer to the same landmarks. Finally, #3 is routinely evaluated by including measures of the similarity or difference in repeated measures of the same structure in the same scans, and most studies include acceptable test–retest reliability measures. It should be noted however, that reliability measures can be inflated by an operator's memory of how he/she processed a particular scan the first time, in addition to the reliability gained by following the same procedure on each scan. In general, test–retest reliability is not the same as accuracy. Overall, it is not hard to imagine that the measured volumes in some studies are more accurate measures of true hippocampal volume than those in other studies, but this speculation cannot be subjected to any sort of quantitative analysis. However, methodological variation across studies is one argument for conducting a meta-analysis, as one would expect genuine and robust effects to emerge from data collected across laboratories.

5.2. Evaluation of the three hypotheses

Examining the reported relationships between hippocampal volume and memory performance across 33 studies yielded no support for the "bigger is better" hypothesis, because of the heterogeneous results across studies, which included both positive and negative correlations between hippocampal volume and memory performance. Instead, some support for both the developmental and aging hypotheses was observed, in that the volume–memory correlation tended to grow more positive as the age of the sample increased (Fig. 2). However, the trend across age appeared to be more strongly influenced by the young end of the age continuum than the elderly end. The strongest association emerged from studies of children, adolescents, and young adults, which have uniformly reported negative correlations between hippocampal volume and memory performance (Chantôme et al., 1999; Foster et al., 1999; MacKay et al., 1998; Sowell et al., 2001; Yurgelun-Todd et al., 2003). Their estimated common r was negative 0.25, with no suggestion of heterogeneity among the studies. More studies in children and young adults are desirable to confirm this finding, but the extant results support a relationship between hippocampal size and memory ability in youth. The similarity

of results from young adults and children suggests that the primary determinant of the volume–memory relationship in early adulthood is developmental.

The neurobiological explanation for the observed developmental relationship is much less clear. In one of the first reports of a negative volume–memory correlation, Foster et al. (1999) made reference to regressive events during development, specifically that "insufficient pruning of the hippocampus during childhood and adolescence (following adequate growth) may lead to reduced mnemonic efficiency". No alternative explanation has been offered, but neither is the pruning hypothesis clearly supported by MR studies of hippocampal development. In the childhood/adolescent age range when cortical gray matter begins a general decrease in volume, hippocampal volumes seem to be relatively stable (Courchesne et al., 2000; Pfefferbaum et al., 1994; Giedd et al., 1996; Giedd, Blumenthal, et al. 1999; Giedd, Jeffries, et al., 1999; Sowell, Trauner, Gamst, & Jernigan, 2002, and see footnote 1), rather than also declining as the simplest version of the pruning hypothesis would predict. But, of course, volumetric measures provide no insight into the organizational processes that occur during neural development. Because memory formation is almost certainly dependent on afferent/efferent relationships between the medial temporal lobe and the neocortex, changes in the ratios of hippocampal/neocortical neuron numbers might play some role in the developmental relationship between hippocampal volume and episodic memory performance.

The a priori argument for an aging/neuropsychological perspective on volume–memory relationships was plausible, in that both hippocampal volumes and episodic memory abilities show a general decline across adulthood, so that one might posit a causal relationship. The empirical support for this hypothesis was, however, surprisingly weak. Studies with only healthy elderly participants have reported both positive and negative correlations between volume and memory. The common correlation was a significant 0.097 when all studies of older adults were included, but dropped to a nonsignificant 0.024 with exclusion of studies that analyzed hippocampal volumes after regressing volume against age. Results from older adults also appeared to be influenced by the method of adjusting (or not) for head size; positive correlations emerged primarily from studies that expressed hippocampal volume as a proportion of intracranial vault capacity (Table 2, Fig. 2). The fragility and variability of the results from healthy older adults are in contrast to the essentially unanimous reports of positive volume–memory correlations in Alzheimer's patients, for whom volume loss co-occurs with plaques and tangles. Perhaps the most striking aspect of the healthy elderly results reviewed here is the variability of the reported correlations between hippocampal and memory measures. The variability issue is taken up below, and then the potential impact of statistical methods for normalizing across age and head/body size.

5.3. Individual variability in older adults

The idea that a variety of cognitive performance measures show greater inter-individual variability in elderly than in young adults is frequently noted in the cognitive aging literature. Hale, Myerson, Smith, and Poon (1988) have suggested that, for RT measures, larger standard deviations in older than younger samples should be interpreted cautiously because larger means are normally accompanied by larger S.D.s, so that generally slower responses from an elderly group may masquerade as increased variability. This critique is less relevant for measures of accuracy in episodic memory tests, at least if ceiling effects do not truncate the range for young adults. Previous meta-analyses and individual studies examining dispersion around the group mean confirm greater individual variability in episodic memory performance for older adults as compared to younger adults (Morse, 1993; Nelson & Dannefer, 1992; Rabbit, 1993; Verhaeghen, 1990; see Devolder, 1991 for contrary results). Longitudinal studies of older adults also indicate large and increasing individual differences in episodic memory over time (Christensen et al., 1999; Wilson et al., 2002).

As noted in the Introduction, estimates of age-related decline in hippocampal volume have varied widely from study to study. In contrast to the cognitive aging literature, age-related changes in dispersion measures have not been a focus of attention in morphometry studies. However, because morphometry studies typically include smaller samples than behavioral studies of aging, variability across studies may also reflect individual variability between the participants recruited by different laboratories. A differential impact of age on hippocampal volume across individuals could arise from both genetic and experiental factors. From a study of monozygotic and dizygotic elderly twins, Sullivan, Pfefferbaum, Swan, and Carmelli (2001) estimate that some 40% of the variance in late-life hippocampal volume can be attributed to genetics, while the other 60% reflects experiential factors. Such experiential factors could operate to increase or decrease volumes. On the one hand, incipient pre-clinical dementia would act to reduce volumes. On the other hand, neurogenesis continues in the primate hippocampus throughout adulthood, and survival of new neurons is reported to be dependent on new learning, at least in rodents (Gould, Beylin, Tanapat, Reeves, & Shors, 1999; Gould, Reeves, et al., 1999; Kornack & Rakic, 1999). Because the new neurons have been observed only in the dentate gyrus, which comprises a small portion of the human hippocampus (2.7% of total volume, according to Harding, Halliday, & Kril, 1998), it seems unlikely that neurogenesis would have a major impact on macroscopic volume measures. However, growth or shrinkage in the dendritic arbors of existing neurons is also likely to occur in adulthood. One example of adult plasticity in human hippocampal volumes occurs in patients with pathologically high levels of glucocorticoids due to Cushing's disease. Small volumes are observed in untreated cases, but these are

reported to rebound after treatment (Bourdeaux et al., 2002; Starkman et al., 1999). Extended discussion of the causes of gains or losses in hippocampal volume are outside the scope of this review, but variability in both memory abilities and hippocampal volumes in older adults is pertinent to the neuropsychological hypothesis. If indeed age-related declines in both memory and hippocampal volume are wildly variable across individuals, chronological age will be a poor predictor of who has suffered what sort of decline.

The observation of negative correlations between hippocampal volumes and memory in young subjects, combined with increased variability in both sorts of measures in the elderly, suggests a way of salvaging the aging/neuropsychological hypothesis. With revision, the hypothesis would specify that, to the extent that a small hippocampus reflects tissue loss in old age (rather than pre-existing differences due to developmental factors), poor memory abilities will result, and a positive correlation between hippocampal volume and memory performance will be observed. The revised hypothesis makes a moderately strong prediction about cross-sectional studies of adults, namely that the memory–volume correlation will increase as the age–volume correlation offers stronger evidence for age-related tissue loss (i.e., becomes increasingly negative). Unfortunately, only eight of the adult studies reviewed here include both of the relevant correlations (without age-normalization; Gur et al., 2000; Raz et al., 1998; Reiman et al., 1998; Sullivan et al., 1995; Tisserand et al., 2000; Van Petten et al., 2004; Woodruff-Pak et al., 2000), so that the power to assess this hypothesis was weak. The relationship between the age–volume correlation and the memory–volume correlation was nonsignificant across those studies (weighted $r = 0.311$, rather than the negative correlation predicted).

Longitudinal studies offer a more transparent means of assessing the revised aging/neuropsychological hypothesis: to the extent that an individual loses hippocampal volume over time, his/her memory performance should also decrease. The data here are very sparse, but do not support the hypothesis. Ylikoski et al. (2000) assessed memory performance in older adults at two time points 5 years apart, together with a visual rating of hippocampal atrophy at Time 2. Approximately half (15 of 35) of the participants had detectable atrophy at Time 2. The groups with and without atrophy showed significant declines in four WMS-R memory tests across the 5 years, although remaining in the normal range. However, the degree of memory decline did not differ between groups, or correlate with degree of atrophy. Rodrigue and Raz (2004) similarly conducted repeated volume measures after a 5-year interval in older adults. Although hippocampal volume showed a significant decline, the magnitude of change was uncorrelated with memory performance at Time 2. In a subset of subjects with memory measures at both time points, hippocampal change significantly predicted memory change for only one of five tests administered, which the authors suggest may reflect a chance finding. Finally, Cohen et al. (2001) repeated both neuropsychological testing and hippocampal volume measures after a 2-year interval. Declines in both memory scores and volumes were noted, with greater hippocampal decrease in subjects possessing an APOE-4 allele than in those without this gene (2.3% versus 0.8% annual decrease). But the two groups did not differ in extent of memory decline, nor was there a correlation between volume loss and memory change across all the subjects. The authors of these longitudinal studies note the relatively mild hippocampal losses in their cognitively intact samples as compared to Alzheimer's patients, and allude to the possibility of a threshold effect—that deterioration of the hippocampus need be extensive before any relationship to cognitive change can be detected.

5.4. Normalizing measures for age

For the studies of hippocampal volume and memory in older adults, a significant common correlation was obtained only with the inclusion of studies that used a residual measure of hippocampal volume after regressing against age, or similarly regressing the memory scores against age (age-adjusted residuals). This procedure removes variance associated with chronological age, and redefines data points as the difference between an individual's measures and others of similar ages in the sample. From published results, it is not possible to evaluate what impact the age-adjustment procedure had on the reported volume–memory correlations. It is possible that the impact was minor. On the other hand, it is very possible that the results of these studies tell us less about the general influence of advancing age on the hippocampus and memory, and more about the significance of deviating from the norm. It is generally accepted that patients with probable Alzheimer's disease as well as those diagnosed with mild cognitive impairment (MCI) have smaller-than-expected hippocampi for their age (Chetelat & Baron, 2003; Kantarci et al., 2002; Wolf et al., 2003, in press). Age-normalization is reported to improve discriminability between normal and AD hippocampal volumes (Hampel et al., 2002). All of the studies reviewed here excluded cognitively impaired participants, but some studies suggest that cognitively normal adults who progress to MCI within a few years also have smaller-than-expected hippocampi at baseline as compared to cognitively normal adults who remain so over time (Jack et al., 2000; Marquis et al., 2002; see also Golomb et al., 1996). Because an age-adjusted residual is the difference between an individual and his/her like-age peers, we might then speculate that studies which use this procedure and obtain large positive correlations between hippocampal volumes and memory scores are strongly influenced by subjects who are on the cusp of clinically significant cognitive decline. This possibility would, of course, be good news for the medical endeavor of helping to identify and aid such individuals before the decline occurs. In contrast however, pooling the results of studies without some form of age-adjustment led to no suggestion of a relationship between hippocampal volume

and memory performance in older adults. For the endeavor of understanding basic structure/function relationships in normal aging, this latter result is less comforting.

5.5. Adjusting for head or body size

The studies reviewed here were divided according to three methods of adjusting hippocampal volumes for overall brain, head, or body size: (1) no such adjustment; (2) division of hippocampal volume by volume of the brain or intracranial vault; and (3) the use of regression or ANCOVA to remove variance in hippocampal volume that is shared with overall brain, vault, or body size. When all the studies were evaluated regardless of age, normalization method had no significant impact. For studies of children and young adults, the initial heterogeneity test was nonsignificant, and Fig. 2 suggests no obvious difference between studies using different methods. For studies of older adults, the method of normalization proved influential, with only the studies using the division method yielding substantial positive correlations between adjusted hippocampal volumes and memory performance (Table 2). It should be noted that splitting the elderly studies by normalization method resulted in relatively small sample sizes, so that the observed differences among methods must be regarded as less conclusive than some of the other analyses based on larger numbers of participants. However, consideration of normalization method may prove critical for future research.

The adjustment of regional brain volumes for head or body size has always been a thorny issue for MR morphometry studies. The nature of the correction factor used depends, in part, on one's assessment of the causal relations between different parameters of body size (height, weight, etc.), skull size, brain size, and the size of individual structures. These have been discussed elsewhere and are not taken up here (Peters et al., 1998; Raz, Gunning-Dixon, et al., 2004). Instead, I summarize some measurement and statistical issues which have also been well-covered elsewhere, but have perhaps not permeated the MR morphometry literature as thoroughly as they should.

Although proportional measures of structure size—structure X considered as a percent of the brain or total intracranial space—have some intuitive appeal, it is well established that such ratio measures do not eliminate the influence of head (or brain) size. Instead, ratio measures have just the opposite effect of producing a relationship between the ratio and its denominator if none exists. Karl Pearson (1897) first commented on the intrinsic relationship between two ratios that share some variable, and the spurious correlations that arise when comparing such ratios. Nearly 30 years ago, Atchley, Gaskins, and Anderson (1976) addressed exactly the issue of interest here—the relationship between a ratio (e.g., X/Y) and its denominator (Y)—by means of large sets of simulated data. When X and Y had a correlation of zero, the correlations between X/Y and Y ranged from -0.50 to -1.0. The magnitude of the spurious correlation depended on the relative variance of X and Y. The spurious correlation grew larger as the standard deviation of Y exceeded that of X, corresponding to the case of a small part (hippocampus) divided by a larger whole (vault size or brain volume).

Atchley's work addressed the extreme case of no pre-existing relationship between the two components of a ratio (Atchley et al., 1976). In morphometry, the volume of any given structure will typically have some true positive correlation with whole brain or vault volume. Real data sets illustrate that ratios tend to change the sign of these correlations from positive to negative. Arndt, Cohen, Alliger, Swayze, and Andreasen (1991) report a correlation between hippocampal and whole-brain volume of 0.26, but a correlation between the hippocampus/brain ratio and brain volume of -0.36. Similarly, Mathalon et al. (1993) report a correlation between intracranial vault and cortical gray volume of 0.43, but a correlation between vault and the gray/vault ratio of -0.31. Data from our group included a modest positive correlation between hippocampal volume and cerebral volume of 0.26, but a large negative correlation (-0.54) between hippocampus/cerebrum and cerebrum (Van Petten et al., 2004).

In contrast to ratio measures, residual methods do remove variance associated with head size: the residual variance in hippocampal volume after regression on head size will (by definition) be perfectly uncorrelated with head size. If one's goal is to eliminate the average influence of head size on the volume of a given structure, regression or ANCOVA methods are strongly preferred to ratio measures. Given the general lack of clarity about the relationship between head size, size of a given brain structure, and cognitive abilities, it will also be useful to analyze these relationships as simple correlations prior to any statistical adjustment.

Having seen that ratio and residual methods of "correcting" for head size are very different, should we care? Specifically, is there some scenario in which the use of a ratio measure could produce a spurious correlation between hippocampal size and memory performance that would not be present when a residual method was used?[4] A priori, it seems that there might be cause for concern under either of two conditions: (1) head size was correlated with memory in the sample; or (2) the relationship between hippocampal size and head size was different for subjects with better versus worse memory performance. One or both of these conditions might emerge as a random coincidence in a small sample, or due to genuine but poorly-understood relationships among the three variables. In the second category, some findings hint at causes for concern. One such hint is

[4] One can also wonder about the converse case: can removing the influence of head size via the residual method eliminate a genuine correlation? If the "bigger is better" hypothesis were true in an absolute sense (larger raw hippocampal volumes associated with better memory), correcting for head size via the residual method would tend to reduce a true positive correlation toward zero. However, the residual method would not yield a negative (reversed) correlation, merely a null effect.

that women (with smaller head sizes on average) tend to perform better than men in several standardized memory tests using verbal material (e.g., California Verbal Learning Test, Rey Auditory Verbal Learning Test, word-pair associate learning, see Lezak, 1995 for review) and perhaps in other episodic memory tests (Beatty, Mold, & Gontkovsky, 2003; Kimura, 1999; McGivern et al., 1997), while men tend to perform better on some tests of spatial memory (Astur, Ortiz, & Sutherland, 1998; Kimura, 1999). The presence of sex-linked differences in both head size and memory performance would implement condition #1 above, if the sample included both sexes and no statistical control for gender. A second such hint begins with considering that the hippocampus is primarily gray matter, so that its size may scale with the amount of gray matter in a brain rather than with intracranial capacity or brain size as a whole. This would not be an issue if we could assume that the proportion of gray matter was constant across different head sizes at the initial state (prior to any age-related changes). The validity of this assumption is unclear. Some recent reports conclude that there is a higher percentage of gray matter in women than men (Allen, Damasio, Grabowski, Bruss, & Zhang, 2003; Gur et al., 1999; Peters et al., 1998; but see Sullivan, Rosenbloom, Serventi, & Pfefferbaum, 2004 for a sex difference in the opposite direction, and Ge et al., 2002 for no sex difference). It is also unclear whether there is some relationship between head size and relative amounts of gray versus white matter independent of gender.[5] Finally, there is universal agreement that adult brain changes prior to age 50 years consist of a decline in gray matter volume but not white matter volume (Courchesne et al., 2000; Ge et al., 2002; Gur et al., 1999; Pfefferbaum et al., 1994; Raz et al., 1998; Sullivan et al., 2004), so that the potential confounds introduced by hippocampus/head size ratios could be different for different age ranges. Overall, there are too many empirical uncertainties to determine the expected influence of a proportional measure of hippocampal size on the hippocampus versus memory correlation in any general sense. Instead, the impact may be idiosyncratic to each data set and add a good deal of noise to the published results considered as a whole.

[5] Some papers suggest this explicitly (Luders, Steinmetz, & Jancke, 2002; Mathalon et al., 1993), but their conclusions appear to be based on negative correlations between gray/ICV ratios and intracranial vault capacity, which is the sort of analysis that produces spurious negative correlations, i.e., if X = gray volume and Y = everything else inside the intracranial cavity (white matter, CSF, etc), then the correlation is between $X/(X+Y)$ and $X+Y$. Because the two terms in the correlation share variables, a spurious negative correlation is to be expected. A more appropriate analysis would examine how gray and white volumes scale up proportional to vault size, i.e., whether the two volumes show the same slope (same regression coefficient) when plotted against vault size. One recent paper shows these relationships graphically (Sullivan et al., 2004; Fig. 3), with an apparent slope of less than 1.0 for gray matter as intracranial volume increases, but a slope of greater than 1.0 for white matter—a pattern corresponding to a higher percentage of gray matter in smaller brains.

5.6. Revisiting the aging-neuropsychology hypothesis

The hypothesis that age-related hippocampal volume loss causes episodic memory impairment stemmed from three premises: (1) episodic memory requires the hippocampus; (2) volumes decline with advancing age in the healthy population; and (3) episodic memory also declines with advancing age in the healthy population. There is little reason to question any of these three basic premises, so that lack of strong evidence for a link between hippocampal volume and memory in older adults raises two sorts of questions: those focusing on the execution of the experiments, and those focusing on conceptualization—the implicit links between the starting premises.

On the execution side, one possible objection is that the memory tasks administered were simply the wrong ones—were not hippocampally dependent. This seems unlikely given that many or all of the wide variety of memory tests employed show impairments in patients with frank hippocampal damage. Moreover, nearly all of the reviewed studies which additionally included a patient group reported significant positive correlations for the patients, suggesting that the tests were capable of detecting hippocampal dysfunction (Cahn et al., 1998; de Toledo-Morrell et al., 2000; Gur et al., 1999; Köhler, Black, et al., 1998; Laakso et al., 2000; O'Driscoll et al., 2001; Sanfilipo et al., 2002; Seidman et al., 2002).

A second sort of potential execution problem is that measurement errors for volume and for memory ability will produce multiplicative noise for a correlation, which could be disastrous for a small effect. Researchers are sensitive to the issue of measurement error for volumetric measures, so that explicit reliability tests are standard, and generally yielded high numbers for the reviewed papers. Reliability of particular memory tests as assessments of memory ability is less often addressed, and some of the reviewed studies used only one memory test, which may not provide a very precise description of an individual's underlying abilities. However, these were a minority of the studies, and inspection of Table 1 shows little commonality among the studies that used a more comprehensive battery of tests. An additional argument against dismissing the null result in older adults as measurement error is, of course, the observation of a significant negative correlation in young subjects.

A third sort of potential pitfall falls midway between execution and conceptualization, and concerns the selection of anatomical region to be correlated with memory performance. Most of the reviewed studies attempted to assess the entire volume of the hippocampal formation. It is possible, however, that age-related shrinkage is more evident in some subregions of the hippocampus than others. Two recent volumetric studies agree that stronger age correlations are observed for the anterior than posterior portions of the hippocampus (Hackert et al., 2002; Driscoll et al., 2003; Pruessner et al., 2001). One of these also report stronger correlations with memory for the head than for other subdi-

visions, but only after residualizing on age, so that the relationship between the aging effect and the memory effect is unclear (Hackert et al., 2002). Further work examining subdivisions of the hippocampus may be useful.[6]

A more general question about the selection of anatomical region concerns the strong emphasis on the medial temporal lobe. A broader restatement of the first premise might be "Episodic memory requires the hippocampus and other medial temporal cortices, and interactions between these regions and the neocortex". Grey matter in many neocortical areas shows age-related volume reductions that are at least as large as those observed in the hippocampus (Coffey, 2000; Raz, 2000). About a third of the studies reviewed here collected temporal lobe measures outside the medial region (de Toledo-Morrell et al., 2000; Gur et al., 2000; Lupien et al., 1998; MacLullich et al., 2002; Raz et al., 1997; Sanfilipo et al., 2002; Sowell et al., 2001; Sullivan et al., 1995; Torres et al., 1997; Van Petten et al., 2004), but few of these examined other lobes, and fewer yet distinguished gray from white matter. It remains possible that additional studies using more comprehensive measures of cortical change with age will reveal stronger relationships to memory abilities.

Although I have suggested that none of the execution issues raised above can individually account for the empirical weakness of the aging/neuropsychological hypothesis, it is yet possible that each adds cumulative noise that hinders the observation of a positive relationship between hippocampal volume and memory ability in older adults. However, it is also relevant to examine the conceptual relationship between starting premise #2 (volume declines) and premise #3 (memory declines). The implicit link for the aging/neuropsychological hypothesis is, of course, that volume loss impairs function. This link is not as trivial as it appears. It is important to remember that when hippocampal volume and memory performance have positive correlations in samples of patients (e.g., Alzheimer's disease), small volumes are likely to be accompanied by cellular pathology (plaques, tangles). The volume loss across normal aging may arise from different causes, as suggested by the decline in cellular metabolites in AD but not normal aging (Adalsteinsson, Sullivan, Kleinhans, Spielman, & Pfefferbaum, 2000; Pfefferbaum, Adalsteinsson, Spielman, Sullivan, & Kim, 1999). Developmental morphometry studies suggest that volume decline and cognitive decline are not invariably linked, in that normal childhood and adolescent development consists of a decrease in gray matter volume and an increase in white matter volume as a proportion of brain size (Courchesne et al., 2000; Giedd, Blumenthal, et al., 1999; Giedd, Jeffries, et al., 1999; Pfefferbaum et al., 1994; Sowell et al., 2002). During some stages of development, increased connectivity among cortical regions (and cortical/subcortical connectivity) seems to be more beneficial than sheer amount of gray matter. This developmental result is partially paralleled in aging studies, in that there is general agreement that signs of white matter deterioration in healthy elderly participants are accompanied by poor cognitive performance (Deary, Leaper, Murray, Staff, & Whalley, 2003; Gunning-Dixon & Raz, 2000; O'Sullivan et al., 2001; Van Petten et al., 2004).

The hippocampus is primarily gray matter, and the studies of youthful participants reviewed here suggest that smaller ratios of hippocampus/cranial vault (which is closely related to overall brain size in children and young adults) are accompanied by better memory performance. The relationship between the decline in gray matter during development and during aging is deeply puzzling, as there appears to be no obvious demarcation in morphometry studies that span childhood to old age (Courchesne et al., 2000; Pfefferbaum et al., 1994). The pattern of age-related results observed here suggests that there may be such a demarcation, in that volume–memory correlations change from generally negative to extremely variable as the age of the sample increases. The explanation for that variability is unknown at present, but is likely to require consideration of age-related changes outside the hippocampus as well as inside.

[6] Along the same lines, several recent studies have examined the relationship between memory performance and volumes of medial temporal lobe regions outside the hippocampus. For the parahippocampal gyrus considered as a whole, one study reports a positive correlation between volume and memory performance (Tisserand, Visser, van Boxtel, & Jolles, 2000), while four report null effects (Köhler, Black, et al., 1998; Köhler, Moscovitch, et al., 1998; Lupien et al., 1998; Petersen et al., 2000; Van Petten et al., 2004). Four other studies have examined the entorhinal region within this gyrus: One of these includes strong positive correlations between volume and memory (Rosen et al., 2003); three report null effects in cross-sectional designs (Du et al., 2003; Rodrigue & Raz, 2004; Turetsky, Moberg, Roalf, Arnold, & Gur, 2003). Two of these latter studies also examined longitudinal change in entorhinal volume, one reporting a null effect (Du et al., 2003), and the other reporting a significant association between volume decline and poor memory (Rodrigue & Raz, 2004). No significant negative correlations have been reported in this literature, but because the large majority of nonsignificant effects were reported merely as "ns" (without numbers), it is difficult to evaluate whether volume–memory correlations in this literature might be more consistent than those reported for the hippocampus.

Acknowledgements

I am grateful to Patrick Davidson and Elizabeth Glisky for discovering some of the relevant papers in the literature, and to Trudy Kuo for comments on the manuscript. Four anonymous reviewers also provided extremely constructive comments on a previous version of this paper. Financial support was provided by the National Institute of Aging (AG14792).

References

Adalsteinsson, E., Sullivan, E. V., Kleinhans, N., Spielman, D. M., & Pfefferbaum, A. (2000). Longitudinal decline of the neuronal marker *N*-acetyl aspartate in Alzheimer's disease. *Lancet, 355*, 1696–1697.

Aggleton, J. P., & Brown, M. W. (1999). Episodic memory, amnesia and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences, 22*, 425–489.

Allen, J. S., Damasio, H., Grabowski, T. J., Bruss, J., & Zhang, W. (2003). Sexual dimorphism and asymmetries in the gray–white composition of the human cerebrum. *Neuroimage, 18*, 880–894.

Arndt, S., Cohen, G., Alliger, R. J., Swayze II, V. W., & Andreasen, N. C. (1991). Problems with ratio and proportion measures of imaged cerebral structures. *Psychiatry Research: Neuroimaging, 40*, 79–89.

Astur, R. S., Ortiz, M. L., & Sutherland, R. L. (1998). A characterization of performance by men and women in a virtual Morris water maze task: A large and reliable sex difference. *Behavioural Brain Research, 93*, 185–190.

Atchley, W. R., Gaskins, C. T., & Anderson, D. (1976). Statistical properties of ratios. I. Empirical results. *Systematic Zoology, 25*, 137–148.

Barber, R., McKeith, I. G., Ballard, C., Gholkar, A., & O'Brien, J. T. (2001). A comparison of medial and lateral temporal lobe atrophy in dementia with Lewy bodies and Alzheimer's disease: Magnetic resonance imaging volumetric study. *Dementia and Geriatric Cognitive Disorders, 12*, 198–205.

Beatty, W. W., Mold, J. W., & Gontkovsky, S. T. (2003). RBANS performance: Influences of sex and education. *Journal of Clinical and Experimental Neuropsychology, 25*, 1065–1069.

Bourdeaux, I., Bard, C., Noël, B., Leclerc, I., Cordeau, M.-P., & Bélaire, M. et al., (2002). Loss of brain volume in endogenous Cushing's syndrome and its reversibility after correction of hypercortisolism. *Journal of Clinical Endrocrinology and Metabolism, 87*, 1949–1954.

Bremner, J. D., Randall, P., Scott, T. M., Bronen, R. A., Seibyl, J. P., & Southwick, S. M. et al., (1995). MRI-based measurement of hippocampal volume in patients with combat-related posttraumatic stress disorder. *American Journal of Psychiatry, 152*, 973–981.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience, 2*, 51–61.

Cahn, D. A., Sullivan, E. V., Shear, P. K., Marsh, L., Fama, R., & Lim, K. O. et al., (1998). Structural MRI correlates of recognition memory in Alzheimer's disease. *Journal of the International Neuropsychological Society, 4*, 106–114.

Chantôme, M., Perruchet, P., Hasboun, D., Dormont, D., Sahel, M., & Sourour, N. et al., (1999). Is there a negative correlation between explicit memory and hippocampal volume? *NeuroImage, 10*, 589–595.

Chetelat, G., & Baron, J. C. (2003). Early diagnosis of Alzheimer's disease: Contribution of structural neuroimaging. *Neuroimage, 18*, 525–541.

Christensen, H., McKinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., & Jacomb, P. et al., (1999). An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychology and Aging, 14*, 365–379.

Coffey, C. E. (2000). Anatomic imaging of the aging human brain. In: C. E. Coffey, J. L. Cummings (Eds.), *Textbook of geriatric neuropsychiatry* (2nd ed., pp. 181–238). Washington, DC: American Psychiatric Press.

Cohen, R. M., Small, C., Lalonde, F., Friz, J., & Sunderland, T. (2001). Effect of apolipoprotein E genotype on hippocampal volume loss in aging healthy women. *Neurology, 57*, 2223–2228.

Convit, A., Wolf, O. T., Tarshish, C., & de Leon, M. J. (2003). Reduced glucose tolerance is associated with poor memory performance and hippocampal atrophy among normal elderly. *Proceedings of the National Academy of Sciences of the United States of America, 100*, 2019–2022.

Courchesne, E., Chisum, H. J., Townsend, J., Cowles, A., Covington, J., & Egaas, B. et al., (2000). Normal brain development and aging: Quantitative analysis at in vivo MR imaging in healthy volunteers. *Radiology, 216*, 672–682.

Cowan, W. M., Fawcett, J. W., O'Leary, D. D., & Stanfield, B. B. (1984). Regressive events in neurogenesis. *Science, 225*, 1258–1265.

Damasio, A. R. (1989). Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition. *Cognition, 33*, 25–62.

de Toledo-Morrell, L., Dickerson, B., Sullivan, M. P., Spanovic, C., Wilson, R., & Bennett, D. A. (2000). Hemispheric differences in hippocampal volume predict verbal and spatial memory performance in patients with Alzheimer's disease. *Hippocampus, 10*, 136–142.

Deary, I. J., Leaper, S. A., Murray, A. D., Staff, R. T., & Whalley, L. J. (2003). Cerebral white matter abnormalities and lifetime cognitive change: A 67-year followup of the Scottish Mental Survey of 1932. *Psychology and Aging, 18*, 140–148.

den Heijer, T., Geerlings, M. I., Hofman, A., de Jong, F. H., Launer, L. J., & Pols, H. A. P. et al., (2003). Higher estrogen levels are not associated with larger hippocampi and better memory performance. *Archives of Neurology, 60*, 213–219.

Devolder, P. A. (1991). Comparison of within-group variabilities across older and younger adults: A meta-analysis. *Experimental Aging Research, 17*, 113–117.

Deweer, B., Lehéricy, S., Pillon, B., Baulac, M., Chiras, J., & Marsault, C. et al., (1995). Memory disorders in probable Alzheimer's disease: The role of hippocampal atrophy as shown with MRI. *Journal of Neurology, Neurosurgery and Psychiatry, 58*, 590–597.

Driessen, M., Herrmann, J., Stahl, K., Zwaan, M., Meier, Z., & Hill, A. et al., (2000). Magnetic resonance imaging volumes of the hippocampus and amygdala in women with borderline personality disorder and early traumatization. *Archives of General Psychiatry, 57*, 1115–1122.

Driscoll, I., Hamilton, D. A., Petropoulos, H., Yeo, R. A., Brooks, W. M., & Baumgartner, R. N. et al., (2003). The aging hippocampus: Cognitive, biochemical and structural findings. *Cerebral Cortex, 13*, 1344–1351.

Du, A. T., Schuff, N., Xhu, X. P., Jagust, W. J., Miller, B. L., & Reed, B. R. et al., (2003). Atrophy rates of entorhinal cortex in AD and normal aging. *Neurology, 60*, 481–486.

Edwards, A. L. (1963). *Experimental design in psychological research*. New York: Holt, Rinehart and Winston.

Fennema-Notestine, C., Stein, M. B., Kennedy, C. M., Archibald, S. L., & Jernigan, T. L. (2002). Brain morphometry in female victims of intimate partner violence with and without posttraumatic stress disorder. *Biological Psychiatry, 51*, 1089–1101.

Foster, J. K., Meikle, A., Goodson, G., Mayes, A. R., Howard, M., & Sünram, S. I. et al., (1999). The hippocampus and delayed recall, bigger is not necessarily better? *Memory, 7*, 715–732.

Ge, Y., Grossman, R. I., Babb, J. S., Rabin, M. L., Mannon, L. J., & Kolson, D. L. (2002). Age-related total gray and white matter changes in normal adult brain. Part I: Volumetric MR imaging analysis. *American Journal of Neuroradiology, 23*, 1327–1333.

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

Giedd, J. N., Jeffries, N. O., Blumenthal, J., Castellanos, F. X., Vaituzis, A. C., & Fernandez, T. et al., (1999b). *Childhood-onset schizophrenia: Progressive brain changes during adolescence Biological Psychiatry, 46*, 892–898.

Giedd, J. N., Vaitzus, C., Hamburger, S. D., Lange, N., Rajapakse, J. C., & Kaysen, D. et al., (1996). Quantitative MRI of the temporal lobe, amygdala, and hippocampus in normal human development: Ages 4–18 years. *Journal of Comparative Neurology, 366*, 223–230.

Golomb, J., Kluger, A., de Leon, M. J., Ferris, S. H., Mittelman, M., & Cohen, J. et al., (1996). Hippocampal formation size predicts declining memory performance in normal aging. *Neurology, 47*, 810–813.

Golomb, J., Kluger, A., de Leon, M. J., Ferris, S. H., Convit, A., & Mittelman, M. S. et al., (1994). Hippocampal formation size in normal human aging: A correlate of delayed secondary memory performance. *Learning & Memory, 1*, 45–54.

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S. J. (2001). A voxel-based morphometric study of aging in 465 adult human brains. *Neuroimage, 14*, 21–36.

Gould, E., Beylin, A., Tanapat, P., Reeves, A., & Shors, T. J. (1999a). Learning enhances neurogenesis in the hippocampal formation. *Nature Neuroscience, 2*, 260–265.

Gould, E., Reeves, A. J., Fallah, M., Tanapat, P., Gross, C. G., & Fuchs, E. (1999b). Hippocampal neurogenesis in Old World primates. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 5263–5267.

Gunning-Dixon, F. M., & Raz, N. (2000). The cognitive correlates of white matter abnormalities in normal aging: A quantitative review. *Neuropsychology*, *14*, 224–232.

Gur, R. E., Turetsky, B. I., Cowell, P. E., Finkelman, C., Maany, V., & Grossman, R. I. et al., (2000). Temporolimbic volume reductions in schizophrenia. *Archives of General Psychiatry*, *57*, 769–775.

Gur, R. C., Turetsky, B. I., Matsui, M., Yan, M., Bilker, W., & Hughett, P. et al., (1999). Sex differences in brain gray and white matter in healthy young adults: Correlations with cognitive performance. *Journal of Neuroscience*, *19*, 4065–4072.

Hackert, V. H., den Heijer, T., Oudkerk, M., Koudstaal, P. J., Hofman, A., & Breteler, M. M. B. (2002). Hippocampal head size associated with verbal memory performance in nondemented elderly. *Neuroimage*, *17*, 1365–1372.

Hale, S., Myerson, J., Smith, G. A., & Poon, L. W. (1988). Age, variability, and speed: Between-subjects diversity. *Psychology and Aging*, *3*, 407–410.

Hampel, H., Teipel, S. J., Bayer, W., Alexander, G. E., Schwarz, R., & Schapiro, M. B. et al., (2002). Age transformation of combined hippocampus and amygdala volume improves diagnostic accuracy in Alzheimer's disease. *Journal of the Neurological Sciences*, *194*, 15–19.

Harding, A. J., Halliday, G. M., & Kril, J. J. (1998). Variation in hippocampal neuron number with age and brain volume. *Cerebral Cortex*, *8*, 710–718.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego: Academic Press.

Hunt, A. L., Orrison, W. W., Yeo, R. A., Haaland, K. Y., Rhyne, R. L., & Garry, P. J. et al., (1989). Clinical significance of MRI white matter lesions in the elderly. *Neurology*, *39*, 1470–1474.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage Publications.

Huttenlocher, P. R. (1993). Morphometric study of human cerebral cortex development. In: M. H. Johnson (Ed.), *Brain development and cognition* (pp. 112–124). Oxford: Blackwell.

Insausti, R., Insausti, A. M., Sobreviela, M. T., Salina, A., & Martínez-Peñuela, J. M. (1998a). Human medial temporal lobe in aging: Anatomical basis of memory preservation. *Microscopy Research and Technique*, *43*, 8–15.

Insausti, R., Juottonen, K., Soininen, H., Insausti, A. M., Partanen, K., & Vainio, P. et al., (1998b). MR volumetric analysis of the human entorhinal, perirhinal, and temporopolar cortices. *American Journal of Neuroradiology*, *19*, 659–671.

Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., & Ivnik, R. J. et al., (2000). Rates of hippocampal atrophy correlate with change in clinical status in aging and AD. *Neurology*, *55*, 484–490.

Jernigan, T. L., Archibald, S. L., Fennema-Notestine, C., Gamst, A. C., Stout, J. C., & Bonner, J. et al., (2001). Effect of age on tissues and regions of the cerebrum and cerebellum. *Neurobiology of Aging*, *22*, 581–594.

Jernigan, T. L., Ostergaard, A. L., & Fennema-Notestine, C. (2001). Mesial temporal, diencephalic, and striatal contributions to deficits in single word reading, word priming, and recognition memory. *Journal of the International Neuropsychological Society*, *7*, 63–78.

Kantarci, K., Xu, Y., Shiung, M. M., O'Brien, P. C., Cha, R. H., & Smith, G. E. et al., (2002). Comparative diagnostic utility of different MR modalities in mild cognitive impairment and Alzheimer's disease. *Dementia & Geriatric Cognitive Disorders*, *14*, 198–207.

Kemper, T. L. (1994). Neuroanatomical and neuropathological changes during aging and dementia. In: M. L. Albert, J. E. Knoefel (Eds.), *Clinical neurology of aging* (2nd ed., pp. 3–67). New York: Oxford University Press.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: MIT Press.

Klekamp, J., Riedel, A., Harper, C., & Kretschmann, H. J. (1991). Morphometric study on the postnatal growth of the hippocampus in Australian Aborigines and Caucasians. *Brain Research*, *549*, 90–94.

Köhler, S., Black, S. E., Sinden, M., Szekely, C., Kidron, D., & Parker, J. L. et al., (1998a). Memory impairments associated with hippocampal versus parahippocampal-gyrus atrophy: An MR volumetry study in Alzheimer's disease. *Neuropsychologia*, *36*, 901–914.

Köhler, S., Moscovitch, M., Winocur, G., Houle, S., & McIntosh, A. R. (1998b). Networks of domain-specific and general regions involved in episodic memory for spatial location and object identity. *Neuropsychologia*, *36*, 129–142.

Kopelman, M. D., Lasserson, D., Kingsley, D., Bello, F., Rush, C., & Stanhope, N. et al., (2001). Structural MRI volumetric analysis in patients with organic amnesia. 2: Correlations with anterograde memory and executive tests in 40 patients. *Journal of Neurology, Neurosurgery and Psychiatry*, *71*, 23–28.

Kornack, D. R., & Rakic, P. (1999). Continuation of neurogenesis in the hippocampus of the adult macaque monkey. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 5768–5773.

Kretschmann, H-J., Kammradt, G., Krauthausen, I., & Wingert, F. (1986). Growth of the hippocampal formation in man. *Bibliotheca Anatomica*, *28*, 27–52.

Laakso, M. P., Hallikainen, M., Hänninen, T., Partanen, K., & Soininen, H. (2000). Diagnosis of Alzheimer's disease: MRI of the hippocampus vs. delayed recall. *Neuropsychologia*, *38*, 579–584.

Lawrie, S. M., Whalley, H. C., Abukmeil, S. S., Kestelman, J. N., Miller, P., & Best, J. J. K. et al., (2002). Temporal lobe volume reduction in people at high risk of schizophrenia with psychotic symptoms. *British Journal of Psychiatry*, *181*, 138–143.

Lawson, J. A., Vogrin, S., Bleasel, A. F., Cook, M. J., Burns, L., & McAnally, L. et al., (2000). Predictors of hippocampal, cerebral, and cerebellar reduction in childhood epilepsy. *Epilepsia*, *41*, 1540–1545.

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.) Oxford: Oxford University Press.

Lieberman, J., Chakos, M., Wu, H., Alvir, J., Hoffman, E., & Robinson, D. et al., (2001). Ziskind Somerfeld Research Award: Longitudinal study of brain morphology in first episode schizophrenia. *Biological Psychiatry*, *49*, 487–499.

Lupien, S. J., de Leon, M., DeSanti, S., Convit, A., Tarshish, C., & Nair, N. P. V. et al., (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience*, *1*, 69–73.

Luders, E., Steinmetz, H., & Jancke, L. (2002). Brain size and grey matter volume in the healthy human brain. *Neuroreport*, *13*, 2371–2374.

MacKay, C. E., Roberts, N., Mayes, A. R., Downes, J. J., Foster, J. K., & Mann, D. (1998). An exploratory study of the relationship between face recognition memory and the volume of medial temporal lobe structures in healthy young males. *Behavioural Neurology*, *11*, 3–20.

MacLullich, A. M. J., Ferguson, K. J., Deary, I. J., Seckl, J. R., Starr, J. M., & Wardlaw, J. M. (2002). Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. *Neurology*, *59*, 169–174.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S. J., & Burgess, N. (2003). Navigation expertise and the human hippocampus: A structural brain imaging analysis. *Hippocampus*, *13*, 250–259.

Marquis, S., Moore, M. M., Howieson, D. B., Sexton, G., Payami, H., & Kaye, J. A. et al., (2002). Independent predictors of cognitive decline in healthy elderly persons. *Archives of Neurology*, *59*, 601–606.

Mathalon, D. H., Sullivan, E. V., Rawles, J. M., & Pfefferbaum, A. (1993). Correction for head size in brain-imaging measurements. *Psychiatry Research: Neuroimaging*, *50*, 121–139.

Mayes, A. R. (2000). Effects on memory of Papez circuit lesions. In: F. Boller, J. Grafman (Eds.), *Handbook of neuropsychology* (2nd ed., pp. 111–132). Amsterdam: Elsevier.

McClelland, J. M., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *109*, 419–457.

McGivern, R. F., Huston, J. P., Byrd, D., King, T., Siegle, G. L., & Reilly, J. (1997). Sex differences in visual recognition memory: Support for a sex-related difference in attention in adults and children. *Brain and Cognition*, *34*, 323–336.

Mirsen, T. R., Lee, D. H., Wong, C. J., Diaz, J. F., Fox, A. J., & Hachinski, V. C. et al., (1991). Clinical correlates of white matter changes on magnetic resonance imaging scans of the brain. *Archives of Neurology*, *48*, 1015–1021.

Mizuno, K., Wakai, M., Takeda, A., & Sobue, G. (2000). Medial temporal atrophy and memory impairment in early stage of Alzheimer's disease: An MRI volumetric and memory assessment study. *Journal of the Neurological Sciences*, *173*, 18–24.

Morse, C. K. (1993). Does variability increase with age? An archival study of cognitive measures. *Psychology and Aging*, *8*, 156–164.

Mueller, E. A., Moore, M. M., Kerr, D. C. R., Sexton, G., Camicioli, R. M., & Howieson, D. B. et al., (1998). Brain volume preserved in healthy elderly through the eleventh decade. *Neurology*, *51*, 1555–1562.

Mungas, D., Reed, B. R., Jagust, W. J., DeCarli, C., Mack, W. J., & Kramer, J. H. et al., (2002). Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. *Neurology*, *59*, 867–873.

Nelson, E. A., & Dannefer, D. (1992). Aged heterogeneity, fact or fiction? The fate of diversity in gerontological research. *Gerontologist*, *32*, 17–23.

Nyberg, L., Habib, R., McIntosh, A. R., & Tulving, E. (2000). Reactivation of encoding-related brain activity during memory retrieval. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11120–11124.

O'Driscoll, G. A., Florencio, P. S., Gagnon, D., Wolff, A.-L. V., Benkelfat, C., & Mikula, L. et al., (2001). Amygdala-hippocampal volume and verbal memory in first-degree relatives of schizophrenic patients. *Psychiatry Research: Neuroimaging*, *107*, 75–85.

O'Sullivan, M., Jones, D. K., Summers, P. E., Morris, R. G., Williams, S. C. R., & Markus, H. S. (2001). Evidence for cortical "disconnection" as a mechanism of age-related cognitive decline. *Neurology*, *57*, 632–638.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, *17*, 299–320.

Pearson, K. (1897). On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London*, *60*, 489–502.

Peters, M., Jäncke, L., Staiger, J. F., Schlaug, G., Huang, Y., & Steinmetz, H. (1998). Unsolved problems in comparing brain sizes in Homo Sapiens. *Brain and Cognition*, *37*, 254–285.

Petersen, R. C., Jack, C. R., Xu, Y.-C., Waring, S. C., O'Brien, P. C., & Smith, G. E. et al., (2000). Memory and MRI-based hippocampal volumes in aging and AD. *Neurology*, *54*, 581–587.

Pfefferbaum, A., Adalsteinsson, E., Spielman, D., Sullivan, E. V., & Lim, K. O. (1999). In vivo spectroscopic quantification of *N*-acetyl moiety, creatine, and choline from large volumes of brain gray and white matter: Effects of normal aging. *Magnetic Resonance in Medicine*, *41*, 276–284.

Pfefferbaum, A., Mathalon, D. H., Sullivan, E. V., Rawles, J. M., Zipursky, R. B., & Lim, K. O. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Archives of Neurology*, *51*, 874–887.

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

Plassman, B. L., Welsh-Bohmer, K. A., Bigler, E. D., Johnson, S. C., Anderson, C. V., & Helms, M. J. et al., (1997). Apolipoprotein E epsilon 4 allele and hippocampal volume in twins with normal cognition. *Neurology*, *48*, 985–989.

Powers, R. E. (2000). Neurobiology of aging. In: C. E. Coffey, & J. L. Cummings (Eds.), *Textbook of geriatric neuropsychiatry* (2nd ed., pp. 33–79). Washington, DC: American Psychiatric Press.

Pruessner, J. C., Collins, D. L., Pruessner, M., & Evans, A. C. (2001). Age and gender predict volume decline in the anterior and posterior hippocampus in early adulthood. *Journal of Neuroscience*, *21*, 194–200.

Rabbit, P. A. (1993). Does it all go together when it goes? The nineteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*, *46A*, 385–434.

Rakic, P., Bourgeois, J.-P., Eckenhoff, M. F., Zecevic, N., & Goldman-Rakic, P. S. (1986). Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science*, *232*, 232–234.

Rao, S. M., Mittenberg, W., Bernardin, L., Haughton, V., & Leo, G. J. (1989). Neuropsychological test findings in subjects with leukouraisis. *Archives of Neurology*, *46*, 40–44.

Raz, N. (2000). Aging of the brain and its impact on cognitive performance: Integration of structural and functional findings. In: F. I. M. Craik, T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 1–90). Mahwah, NJ: Erlbaum.

Raz, N., Gunning, F. M., Head, D., Dupuis, J. H., McQuain, J., & Briggs, S. D. et al., (1997). Selective aging of the human cerebral cortex observed in vivo: Differential vulnerability of the prefrontal gray matter. *Cerebral Cortex*, *7*, 268–282.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging, Evidence from structural magnetic resonance imaging. *Neuropsychology*, *12*, 95–114.

Raz, N., Rodrigue, K. M., Head, D., Kennedy, K. M., & Acker, J. D. (2004a). Differential aging of the medial temporal lobe: A study of five-year change. *Neurology*, *62*, 433–438.

Raz, N., Gunning-Dixon, F., Head, D., Rodrigue, K. M., Williamson, A., & Acker, J. D. (2004b). Aging, sexual dimorphism, and hemispheric asymmetry of the cerebral cortex: Replicability of regional differences in volume. *Neurobiology of Aging*, *25*, 377–396.

Reiman, E. M., Uecker, A., Caselli, R. J., Lewis, S., Bandy, D., & de Leon, M. J. et al., (1998). Hippocampal volumes in cognitively normal persons at genetic risk for Alzheimer's disease. *Annals of Neurology*, *44*, 288–291.

Rempel-Clower, N. L., Zola, S. M., Squire, L. R., & Amaral, D. G. (1996). Three cases of enduring memory impairment after bilateral damage limited to the hippocampal formation. *Journal of Neuroscience*, *16*, 5233–5255.

Resnick, S. M., Pham, D. L., Kraut, M. A., Zonderman, A. B., & Davatzikos, C. (2003). Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. *Journal of Neuroscience*, *23*, 3295–3301.

Rodrigue, K. M., & Raz, N. (2004). Shrinkage of entorhinal cortex over five years predicts memory performance in healthy adults. *Journal of Neuroscience*, *24*, 956–963.

Rosen, A. C., Prull, M. W., Gabrieli, J. D. E., Stoub, T., O'Hara, R., & Friedman, L. et al., (2003). Differential associations between entorhinal and hippocampal volumes and memory performance in older adults. *Behavioral Neuroscience*, *117*, 1150–1160.

Sanfilipo, M., Lafargue, T., Rusinek, H., Arena, L., Loneragan, C., & Lautin, A. et al., (2002). Cognitive performance in schizophrenia: Relationship to regional brain volumes and psychiatric symptoms. *Psychiatry Research: Neuroimaging*, *116*, 1–23.

Schmidt, R., Fazekas, F., Offenbacher, H., Lytwyn, H., Blematl, B., & Niederkorn, K. et al., (1991). Magnetic resonance imaging white matter lesions and cognitive impairment in hypertensive individuals. *Archives of Neurology*, *48*, 417–420.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery and Psychiatry*, *20*, 11–21.

Seidman, L. J., Faraone, S. V., Goldstein, J. M., Kremen, W. S., Horton, N. J., & Makris, N. et al., (2002). Left hippocampal volume

as a vulnerability indicator for schizophrenia. *Archives of General Psychiatry*, *59*, 839–849.

Senkfor, A. J., Van Petten, C., & Kutas, M. (2002). Episodic action monitoring for real objects, An ERP investigation with Perform, Watch, and Imagine action encoding tasks versus a non-action encoding task. *Journal of Cognitive Neuroscience*, *14*, 402–419.

Sheline, Y. I., Sanghavi, M., Mintun, M. A., & Gado, M. H. (1999). Depression duration but not age predicts hippocampal volume loss in medically healthy women with recurrent major depression. *Journal of Neuroscience*, *19*, 5034–5043.

Simpson, S. W., Baldwin, R. C., Burns, A., & Jackson, A. (2001). Regional cerebral volume measurements in late-life depression: Relationship to clinical correlates, neuropsychological impairment and response to treatment. *International Journal of Geriatric Psychiatry*, *16*, 469–476.

Smith, M. L., & Bigel, M. G. (2000). The temporal lobes and memory In: F. Boller, & J. Grafman (Eds.), *Handbook of neuropsychology* (2nd ed., pp. 49–66). Amsterdam: Elsevier.

Soininen, H. S., Partanen, K., Pitkänen, A., Vainio, P., Hänninen, T., & Hallikainen, M. et al., (1994). Volumetric MRI analysis of the amygdala and the hippocampus in subjects with age-associated memory impairment. *Neurology*, *44*, 1660–1668.

Sowell, E. R., Delis, D., Stiles, J., & Jernigan, T. L. (2001). Improved memory functioning and frontal lobe maturation between childhood and adolescence: A structural MRI study. *Journal of the International Neuropsychological Society*, *7*, 312–322.

Sowell, E. R., Trauner, D. A., Gamst, A., & Jernigan, T. L. (2002). Development of cortical and subcortical brain structures in childhood and adolscence: A structural MRI study. *Developmental Medicine and Child Neurology*, *44*, 4–16.

Starkman, M. N., Giordani, B., Gebarski, S. S., Berent, S., Schork, M. A., & Schteingart, D. E. (1999). Decrease in cortisol levels reverses human hippocampal atrophy following treatment of Cushing's disease. *Biological Psychiatry*, *46*, 1595–1602.

Stein, M. B., Koverola, C., Hanna, C., Rochia, M. G., & McClarty, B. (1997). Hippocampal volume in women victimized by childhood sexual abuse. *Psychological Medicine*, *27*, 951–959.

Sullivan, E. V., Marsh, L., Mathalon, D. H., Lim, K. O., & Pfefferbaum, A. (1995). Age-related decline in MRI volumes of temporal lobe gray matter but not hippocampus. *Neurobiology of Aging*, *16*, 591–606.

Sullivan, E. V., Rosenbloom, M., Serventi, K. L., & Pfefferbaum, A. (2004). Effects of age and sex on volumes of the thalamus, pons, and cortex. *Neurobiology of Aging*, *25*, 185–192.

Sullivan, E. V., Pfefferbaum, A., Swan, G. E., & Carmelli, D. (2001). Heritability of hippocampal size in elderly twin men: Equivalent influences from genes and environment. *Hippocampus*, *11*, 754–762.

Terry, R. D., & Katzman, R. (2001). Life span and synapses: Will there be a primary senile dementia? *Neurobiology of Aging*, *22*, 347–348.

Tisserand, D. J., Visser, P. J., van Boxtel, M. P. J., & Jolles, J. (2000). The relation between global and limbic brain volumes on MRI and cognitive performance in healthy individuals across the age range. *Neurobiology of Aging*, *21*, 569–576.

Torres, I. J., Flashman, L. A., O'Leary, D. S., Swayze, V. I., & Andreasen, N. C. (1997). Lack of an association between delayed memory and hippocampal and temporal lobe size in patients with schizophrenia and healthy controls. *Biological Psychiatry*, *42*, 1087–1096.

Tupler, L. A., Coffey, C. E., Logue, P. E., Djang, W. T., & Fagan, S. M. (1992). Neuropsychological importance of subcortical white matter hyperintensity. *Archives of Neurology*, *49*, 1248–1252.

Turetsky, B. I., Moberg, P. J., Roalf, D. R., Arnold, S. E., & Gur, R. E. (2003). Decrements in volume of anterior ventromedial temporal lobe and olfactory dysfunction in schizophrenia. *Archives of General Psychiatry*, *60*, 1193–2000.

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

Uttal, W. R. (2001). *The new phrenology. The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.

VanLandingham, K. E., Heinz, E. R., Cavazos, J. E., & Lewis, D. V. (1998). Magnetic resonance imaging of hippocampal injury after prolonged focal febrile convulsions. *Annals of Neurology*, *43*, 413–426.

Van Petten, C., Plante, E., Davidson, P. S. R., Kuo, T. Y., Bajuscak, L., & Glisky, E. L. (2004). Memory and executive function in older adults: Relationships with temporal and prefrontal gray matter volumes and white matter hyperintensities. *Neuropsychologia*, *42*(10) 1313–1335.

Verhaeghen, P. (1990). Memory aging and the increasing diversity hypothesis: Empirical evidence. *Psychologica Belgica*, *30*, 167–176.

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood. Estimates of linear and non-linear age effects and structural models. *Psychological Bulletin*, *122*, 231–249.

Visser, P. J., Scheltens, P., Verhey, F. R. J., Schmand, B., Launer, L. J., & Jolles, J. et al., (1999). Medial temporal lobe atrophy and memory dysfunction as predictors for dementia in subjects with mild cognitive impairment. *Journal of Neurology*, *246*, 477–485.

Wheeler, M. E., Peterson, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 11125–11129.

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., & Evans, D. A. et al., (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging*, *17*, 179–194.

Wilson, R. S., Sullivan, M., de Toledo-Morrell, L., Stebbins, G. T., Bennett, D. A., & Morrell, F. (1996). Association of memory and cognition in Alzheimer's disease with volumetric estimates of temporal lobe structures. *Neuropsychology*, *10*, 459–463.

Wolf, F. M. (1986). *Meta-analysis. Quantitative methods for research synthesis*. Beverly Hills, CA: Sage Publications.

Wolf, H., Hensel, A., Kruggel, F., Riedel-Heller, S. G., Arendt, T., Wahlund, L. -O., & Gertz, H. -J. (in press). Structural correlates of mild cognitive impairment. *Neurobiology of Aging*.

Wolf, H., Jelic, V., Gertz, H.-J., Nordberg, A., Julin, P., & Wahlund, L. O. (2003). A critical discussion of the role of neuroimaging in mild cognitive impairment. *Acta Neurologica Scandinavica*, *107*, 52–76.

Woodruff-Pak, D. S., Goldenberg, G., Downey-Lamb, M. M., Boyko, O. B., & Lemieux, S. K. (2000). Cerebellar volume in humans related to magnitude of classical conditioning. *Neuroreport*, *11*, 609–615.

Ylikoski, R., Salonen, O., Mäntylä, R., Ylikoski, A., Keskivaara, P., & Leskelä, M. et al., (2000). Hippocampal and temporal lobe atrophy and age-related decline in memory. *Acta Neurologica Scandinavica*, *101*, 273–278.

Yurgelun-Todd, D., Killgore, W. D., & Cintron, C. B. (2003). Cognitive correlates of medial temporal lobe development across adolescence: A magnetic resonance imaging study. *Perceptual and Motor Skills*, *96*, 3–17.

Zola, S. M., & Squire, L. R. (2000). The medial temporal lobe and the hippocampus. In: E. Tulving, & F. I. M. Craik (Eds.), *The Oxford handbook of memory* (pp. 485–500). Oxford: Oxford University Press.