Neuropsychology
2009, Vol. 23, No. 5, 541–550

© 2009 American Psychological Association
0894-4105/09/$12.00 DOI: 10.1037/a0016161

The Prevalence of Cortical Gray Matter Atrophy May Be Overestimated In the Healthy Aging Brain

Saartje Burgmans, Martin P. J. van Boxtel, Eric F. P. M. Vuurman, Floortje Smeets, and Ed H. B. M. Gronenschild
Maastricht University

Harry B. M. Uylings
Maastricht University and VU University Medical Center Amsterdam

Jelle Jolles
Maastricht University and VU University Medical Center Amsterdam

Prevailing opinion holds that normal brain aging is characterized by substantial atrophy of cortical gray matter. However, this conclusion is based on earlier studies whose findings may be influenced by the inclusion of subjects with subclinical cognitive disorders like preclinical dementia. The present magnetic resonance imaging study tested this hypothesis. Cognitively healthy subjects (mean age 72 years, range 52–82) who remained cognitively stable over a 3-year period were compared to subjects with significant cognitive decline. Subjects who developed dementia within 6 years after the scan session were excluded. The gray matter volumes of seven cortical regions were delineated on T1-weighted magnetic resonance imaging scans. Participants without cognitive decline did not exhibit an age effect on the gray matter volume. Conversely, participants with cognitive decline exhibited a significant age effect in all the seven areas. These results suggest that cortical gray matter atrophy may have been overestimated in studies on healthy aging, since most studies were unable to exclude participants with a substantial atypical cognitive decline or preclinical dementia. Our results underscore the importance of establishing stringent inclusion criteria for future studies on normal aging.

Keywords: healthy aging, gray matter atrophy, cognitive decline, cerebral cortex, MRI

Volumetric studies of healthy aging consistently report that the normal aging human brain is characterized by gray matter atrophy. Both cross-sectional (Sowell et al., 2003; Tisserand et al., 2002) and longitudinal studies (Raz et al., 2005; Resnick, Pham, Kraut, Zonderman, & Davatzikos, 2003) found a significant decrease of gray matter tissue in several brain areas namely in the frontal, parietal and medial temporal areas (Raz, 2000). As a result, a gradual atrophy of the cortical gray matter is generally considered to be a part of the normal aging process. However, this prevailing model of the aging brain may not be completely accurate. Studies of healthy aging usually exclude "nonhealthy" participants based on medical history or of short, medical or cognitive, screening instruments. Extensive neuropsychological test batteries are seldom used to exclude participants with atypical cognitive decline, and the cognitive status of participants has rarely been tested longitudinally for this reason. Clearly, repeated evaluation is imperative to verify whether participants are, and remain, cognitively healthy. Moreover, several authors have suggested that their study samples may have included preclinically ill subjects (Raz et al., 2004; Resnick et al., 2003; Sliwinski & Buschke, 1999). This problem may particularly apply to samples that are composed of highly educated participants, as these subjects are more likely to pass cognitive screening. According to Sliwinski et al. (Sliwinski & Buschke, 1999; Sliwinski, Lipton, Buschke, & Stewart, 1996), preclinically ill participants have been routinely included in normative aging studies. Sliwinski et al. estimated the percentage of participants with preclinical dementia in studies of "healthy aging" at 20% (Sliwinski et al., 1996). This raises the question whether cortical gray matter atrophy really occurs in healthy aging, or whether it is an artifact because of the inclusion of subjects with preclinical cognitive pathology, such as an early stage of dementia.

We investigated whether the prevalence of cortical gray matter atrophy is overestimated in healthy brain aging and hypothesized

Saartje Burgmans, Martin P. J. van Boxtel, Eric F. P. M. Vuurman, Floortje Smeets, Ed H. B. M. Gronenschild, and Jelle Jolles, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, European Graduate School of Neuroscience (EURON), Maastricht University; Harry B. M. Uylings, Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience and European Graduate School of Neuroscience (EURON), Maastricht University and Department of Anatomy and Neuroscience, VU University Medical Center Amsterdam. Jelle Jolles is also with AZIRE Research Institute Faculty of Psychology and Education, VU University Medical Center, Amsterdam.

We thank Danielle Tisserand (Ph.D.) for her help in acquiring and preprocessing the MRI images, and Argonde van Harten (M.Sc.) and Jennifer Reijnders (M.Sc.) for their skillful technical assistance in tracing gray matter. We also thank Frans Verhey (Ph.D. and neuropsychiatrist at Maastricht University Hospital) for evaluating the neuropsychiatric profiles of the cases, and we thank Marjan van den Akker (Ph.D. at the Department of General Practice, University Maastricht) for retrieving the relevant medical information regarding the participants from the Registration Network Family database. Finally, we like to thank the reviewers for their stimulating comments. They have led to substantial improvements of our paper.

Correspondence concerning this article should be addressed to Saartje Burgmans, School for Mental Health and Neuroscience, Maastricht University, PO Box 616, 6200 MD Maastricht, the Netherlands. E-mail: s.burgmans@np.unimaas.nl

that cognitively healthy aging participants do not exhibit a substantial decline of gray matter volume in brain areas that are highly associated with cognition. Measurements were made of the gray matter volume of multiple brain structures that are known to be highly associated with cognition, that is, the hippocampus, parahippocampal gyrus, inferior prefrontal cortex, orbital prefrontal cortex, dorsolateral prefrontal cortex, anterior cingulate gyrus, and posterior cingulate gyrus (see Figure 1).

Method

Participants

Eighty-eight healthy participants were selected from a larger longitudinal study on determinants of cognitive aging, namely the Maastricht Aging Study (MAAS). The aims, population characteristics and design of MAAS have been described in detail elsewhere (Jolles, Houx, Van Boxtel, & Ponds, 1995; Van Boxtel et al., 1998). In short, MAAS encompasses approximately 1,900 healthy subjects aged 25 to 80 years. The data used for the present study were derived from the baseline assessment (t_0), the 3-year follow-up (t_3), the 6-year follow-up (t_6), and the 9-year follow-up (t_9). Selection of the participants for the present study took place based on difference in performance on two extensive neuropsychological assessments, which were performed 3 years apart (at t_0 and t_3).

Exclusion criteria at t_0 were: cerebrovascular pathology (CVA); a history of transient ischemic attacks (TIA); brain surgery, all tumors or (congenital) malformations of the nervous system; multiple sclerosis; parkinsonism; epilepsy; dementia; organic psychosis (other than dementia); schizophrenia; affective psychosis; coma; mental retardation; renal failure; electroconvulsive therapy; and the regular use of psychotropic drugs. Individuals were excluded on the basis of medical information provided by their general practitioners, which was based on a standard morbidity coding system (Lamberts & Wood, 2002), and on the basis of a semistructured medical questionnaire administered at baseline.

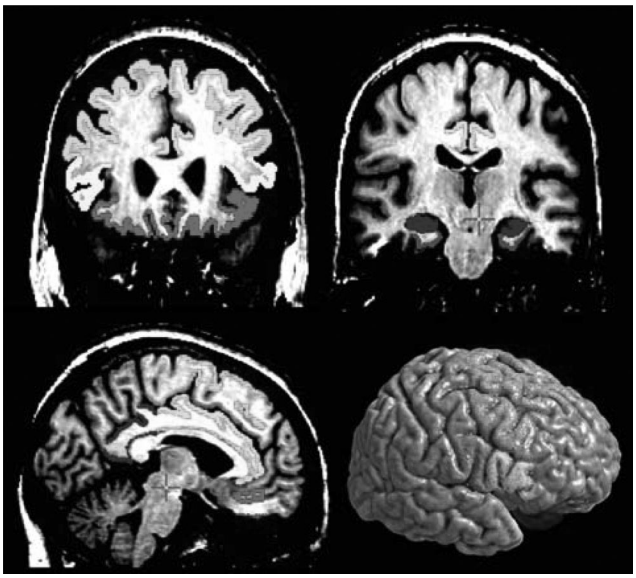


Figure 1. Measured brain areas: inferior prefrontal cortex (yellow), orbital prefrontal cortex (red), dorsolateral prefrontal cortex (light green), anterior cingulate gyrus (light blue), posterior cingulate gyrus (brown), hippocampus (dark blue), and parahippocampal gyrus (purple).

At t_3, 44 cognitively healthy individuals (i.e., not exhibiting a significant cognitive decline in the last 3 years) were selected and matched on the basis of age, sex, and educational level to 44 individuals who exhibited a significant cognitive decline between t_0 and t_3. Differences in performance on two extensive neuropsychological assessments performed 3 years apart (t_0 and t_3) were used to determine the criteria for decline. Cognitive decline was defined as follows: (1) a score of 24 or lower, or a decline of at least three points on the Mini-Mental State Examination (MMSE), or (2) a decline of at least 30% on two or more of the six core tests that were used in MAAS to probe different cognitive domains, that is, verbal memory (Verbal Learning Test, VLT, immediate and delayed recall); verbal fluency (animal naming); basic processing speed (Letter Digit Substitution Test, LDST); and complex information processing (Concept Shifting Test, CST, and Stroop interference) (Jolles et al., 1995; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005, 2006a, 2006b, 2006c, 2006d).

All of the 88 individuals selected were invited to visit the Maastricht University Hospital for a structural magnetic resonance imaging (MRI) scan within 4 weeks after the second cognitive screening (t_3). Six years after the baseline assessment (t_6, i.e., 3 years after the scan session) the participants again took part in a cognitive assessment as part of the MAAS. The t_6 measurements were used in the present study to verify a potential change in the cognitive status of participants after the scan session. Nine years after the baseline assessment (t_9, i.e., 6 years after the scan session) the participants completed a medical questionnaire and were screened for the presence of dementia. Ten participants were excluded, because they were diagnosed with dementia by their general practitioner according to the standard criteria (APA, 1994; Lamberts & Wood, 2002). In addition, an experienced neuropsychiatrist from the Maastricht University Hospital made an etiological diagnosis based on the available neuropsychological data, the available information about daily life functioning and each participant's medical profile (i.e., medical history, comorbidity, course of symptoms and complaints, and evaluation of the MRI scan). The profile of nine participants was compatible with probable Alzheimer's disease (AD) or possible AD/mixed type dementia (McKhann et al., 1984). The profile of one participant matched possible frontotemporal dementia.

Fourteen individuals exhibiting decline and nine controls were excluded from further analysis because of MRI movement artifacts ($n = 8$), neuroanatomical abnormalities ($n = 5$), or because they were diagnosed with dementia at t_9 ($n = 10$; seven cognitive decliners and three controls) (APA, 1994; McKhann et al., 1984). The movement artifacts were critical because we used a semiautomatic tracing method, which required a good contrast between gray and white matter for the tissue classification procedure. The neuroanatomical abnormalities that we observed were: an abnormal large left lateral ventricle (in two subjects); absence of the left temporal lobe; a large lesion in the thalamus; and severe lesions in the basal ganglia. For the present study, two groups of participants were compared: (1) cognitively healthy participants who did not exhibit a significant cognitive decline and did not develop dementia ($n = 35$; 19 women; mean age at $t_0 = 69.1$ year; mean

educational level = low to intermediate); and (2) participants with substantial cognitive decline who did not develop dementia (n = 30; 14 women; mean age at t_0 = 69.2 years; mean educational level = low to intermediate). Twenty-eight participants of the first group (cognitively healthy) remained stable 3 years after the scan session. Seven participants of this group met our criteria for cognitive decline at t_6. Written informed consent was obtained from all participants. The MAAS protocol and the additional scan protocol were approved by the local Medical Ethics Committee. See Figure 2 for the study's time line and assignment of participants to the study groups.

The profiles of cognitive decline represented in the second group were diverse: 18 individuals were included based on the MMSE criterion only, 11 were based on the cognitive test criterion only, and one individual met both criteria. Of the 18 individuals who were included based on the MMSE criterion only, 14 participants had a decline of at least three points on the MMSE, and 7 participants had an MMSE score of 24 or lower at t_3. Most of these participants also had a decline of more than 30% on one of the other cognitive test scores. With regard to the 11 individuals who were included only on the basis of the cognitive test criterion, 9 subjects had more than a 30% decline on the CST subtest C; 4 subjects had more than a 30% decline on the fluency task; 6 subjects had more than a 30% decline on the Stroop Interference; 1 subject had more than a 30% decline on the VLT immediate recall; and 2 subjects had more than a 30% decline on the VLT delayed recall. The participant who met both criteria had 3 points decline on the MMSE and a MMSE score of 24 at t_3. In addition, this participant had more than a 30% reduction on the VLT delayed recall, LDST and CST.

MRI Acquisition and Analysis

MRI scans were acquired at t_3 with a 1.5 T Gyroscan NT MRI scanner (Philips, Best, The Netherlands). T1-weighted images were obtained in the coronal plane (perpendicular to the anterior commisure—posterior commisure [AC-PC] line), using a 3D-gradient fast field echo (FFE) sequence (TR = 35 milliseconds, TE = 7 milliseconds, FA = 35, FOV = 240 mm, slice thickness = 1.5 mm, matrix size = 256 × 256, voxel size = 0.94 × 0.94 × 1.5 mm). The image volumes were corrected for MR signal nonuniformities caused by magnetic field inhomogeneities in the scanner (Sled, Zijdenbos, & Evans, 1998). As our tracing method was semiautomatic, it was necessary to keep the contrast and brightness of the displayed images constant. Therefore, a standardization of the grayscale intensities of the images was performed using the procedure developed by Nyúl and Udupa (Nyúl & Udupa, 1999).

Semiautomatic Tracing Method

The three-dimensional boundaries (contours) of the brain areas were drawn manually using the DISPLAY software package (Montreal Neurological Institute; http://www.bic.mni.mcgill.ca/software/). The defined contours were used to trace gray matter using a tool of the custom software package GIANT, developed at the Maastricht School for Mental Health and Neuroscience (EHBMG). This tool labels all voxels inside a manually drawn contour that are classified as gray matter. This classification is performed by the INSECT package developed at the Montreal Neurological Institute (Zijdenbos, Forghani, & Evans, 2002). The results were then displayed in GIANT as overlays on top of the original stack in a triplanar view, together with a 3D rotatable outer surface display for sulci identification. All the results were extensively verified slice-by-slice, and corrected manually when necessary. This manual correction was performed very accurately so that the result of the tracing was very similar to the full manual tracing method that is widely used in studies on volumetry and aging (John et al., 2006; Raz, Rodrigue, Kennedy, & Acker, 2007). The great advantage of this semiautomatic approach is that it is faster than the full manual method and leads to lower intraobserver variability. Ten randomly selected brains were measured twice, and these yielded high test–retest reliability (intraclass correlation coefficients, ICC [Shrout & Fleiss, 1979], of >0.95 were attained for all brain areas).

Seven brain areas (see Figure 1) were traced by using the following criteria:

Inferior PFC (~Brodmann areas 44 and 45). The posterior border was defined by the precentral sulcus and the anterior border by the frontal pole region. The frontal pole region lacks evident macroscopical anatomical landmarks. Therefore, a straight line was drawn to mark the posterior border of the frontal pole. This straight line was drawn upward from the anterior top of the cingulate sulcus. The inferior frontal sulcus was taken as the dorsal border, and the horizontal ramus of the Sylvian fissure as the ventral border (Uylings et al., 2005).

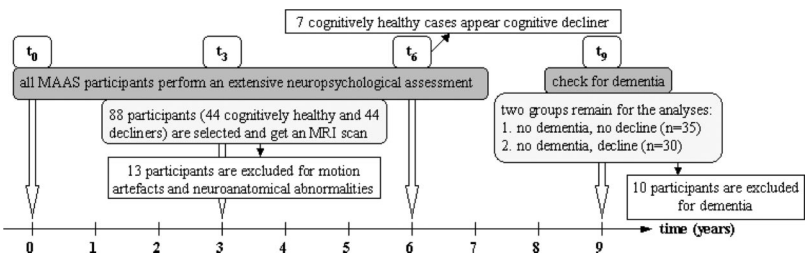


Figure 2. The study's time line and assignment of participants to the study groups. Neuropsychological assessments was done at t_0, t_3, and t_6. At t_3, the participants were selected and assigned to the two groups and MRI scanning took place. At t_6, seven participants of the healthy group met our criteria for cognitive decline. At t_9, participants were screened for dementia and 10 participants were excluded because of dementia.

Orbital PFC (~Brodmann areas 47, 11, and 12). Based upon cytoarchitectonic experience (HBMU), the posterior border was defined by a straight line that was drawn downward from the inner curvature of the corpus callosum. The anterior border was determined by the posterior border of the frontal pole region. Medially, this region was bounded by the ventral part of the anterior cingulate region. The lateral border was the inferior frontal gyrus (Tisserand et al., 2002).

"Dorsolateral PFC" (~Brodmann areas 8, 9, 46, and part of 6). This region has as the lateral ventral border the inferior frontal sulcus, and as the medial border the (interrupted) paracingulate sulcus. The posterior border was defined by the precentral sulcus, and the anterior border by the frontal pole region (Tisserand et al., 2002).

Anterior cingulate gyrus (~Brodmann area 24 and 33) and Posterior cingulate gyrus (~Brodmann area 23, 30, 29). Cingulate gyrus borders were set between the ventral and dorsal banks of both the callosal sulcus and the cingulate sulcus (Jones et al., 2006). The most posterior coronal slice on which the posterior commisure was visible was used to subdivide the anterior cingulate gyrus from the posterior cingulate gyrus (Jones et al., 2006).

Hippocampus. The volume of the hippocampus included the hippocampus proper (including the dentate gyrus), the alveus, and the subiculum. The anterior and posterior borders were based on both sagittal and coronal sections of the brain (Insausti et al., 1998; Visser et al., 1999).

Parahippocampal gyrus. Tracing of the parahippocampal gyrus was performed on every coronal slice on which the hippocampus was visible. The subiculum was taken as the dorsal border and the collateral sulcus as the ventral border (Insausti et al., 1998; Thangavel, Van Hoesen, & Zaheer, 2008).

All the criteria used were based on knowledge of human cortical cytoarchitectonics (HBMU). The criteria for the frontal areas were also used in earlier studies in which HBMU has participated (Burgmans et al., 2008; Jones et al., 2006; Tisserand et al., 2002; Uylings et al., 2005). To define the boundaries of our regions of interest, the clearest landmarks with the lowest individual gyral variability were used. To correct for head size, intracranial volumes were measured by an automated method (FSL Brain Extraction Tool) developed at Oxford Centre for Functional MRI of the Brain (Smith, 2002). The inner-skull contours of our scans, which were generated by this tool, were checked by a neuroanatomist (HBMU).

Statistical Analyses

Statistical analyses were performed using the Statistical Package for Social Sciences (SPSS Inc., Chicago), version 15.0 for Windows. First, the characteristics of the two groups were calculated. The characteristics of both groups were compared at t_0, t_3, and t_6 with independent *t* tests (continuous variables) and chi-square tests (categorical variables). Second, the effect of age on gray matter volume was calculated. The univariate ANOVA under the General Linear Model (GLM) menu was used to test the age effect on the individual brain areas, and to calculate age by group interactions. All individual tests were checked for normality, linearity, independency (of the values of the outcome variables), and influential outliers. The volumes were calculated as the sum of the left and right hemisphere in native space. All ANOVA analyses were adjusted for intracranial volume by expressing all volumes as a percentage of the intracranial volume. In addition, the effect sizes (Pearson's *r* correlation) of the age effects were calculated and a power analyses was performed with G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). Finally, nonlinear (quadratic) effects were investigated. To this end, the variables age and age^2 were tested within the same GLM analyses. Both variables were centered (the mean was set at 0) to prevent collinearity.

Results

Characteristics of the Healthy and the Cognitive Decline Group

Table 1 shows the characteristics of the two groups studied, that is, the cognitively healthy group and the cognitive decline group, at three time points (at baseline t_0, after 3 years t_3, and after 6 years t_6). The two groups did not differ significantly with respect to age, gender, educational level, and intracranial volume. In addition, there were no significant group differences with respect to cardiovascular status, such as hypertension, diabetes, other cardiovascular diseases, and use of medication.

At t_0, there were no significant differences with respect to cognitive test performance, except for one memory subtask: the Verbal Learning Test–delayed recall, on which the decline group performed significantly worse than the healthy group (7.93 vs. 9.26 remembered words, $t = -2.01$, $p = .05$, *Cohen's* $d = 0.54$, $df = 63$). At t_3, the decline group performed significantly worse than the healthy group on the Mini-Mental State Examination (26.27 vs. 28.23, $t = -4.49$, $p < .01$, $d = 1.07$, $df = 63$); the Fluency task (18.30 vs. 21.43 named animals, $t = -2.51$, $p = .02$, $d = 0.61$, $df = 63$); and the Verbal Learning Test–delayed recall (7.60 vs. 9.86 remembered words, $t = -3.70$, $p < .01$, $d = 0.94$, $df = 63$).

At t_6, the decline group performed significantly worse than the healthy group on the Mini-Mental State Examination (27.5 vs. 28.4, $t = -2.05$, $p = .04$, $d = 0.50$, $df = 61$); the Verbal Learning Test–immediate recall (9.7 vs. 11.9 remembered words, $t = -3.66$, $p < .01$, $d = 0.90$, $df = 61$); and the Verbal Learning Test–delayed recall (7.0 vs. 10.2 remembered words, $t = -4.29$, $p < .01$, $d = 1.10$, $df = 61$). In addition, seven cases of the 35 "cognitively healthy" participants met our criteria for cognitive decline at t_6. The other 28 participants remained cognitively stable.

Effect of Age on the Gray Matter Volume of the Brain Areas Studied

Table 2, Table 3, and Figure 3 show the effect of age on the volumes of each brain area studied. The cognitively healthy group did not exhibit a significant age effect on any of the measured brain structures, except for the hippocampus ($F = 6.33$, $p = .017$, $r = .40$, $df = 1$). However, when we excluded the seven participants of the cognitively healthy group who were no longer cognitively

Table 1
Group Characteristics

	Cognitively healthy (n = 35)	Cognitive decline (n = 30)	Group difference				
	Mean (*SD*) or %	Mean (*SD*) or %	t	d	df	χ^2	p
t_0							
Age (years)	69.1 (7.7) range 49.3–79.3	69.2 (8.1) range 50.0–81.0	0.06	0.01	63		0.96
Female sex (%)	54.3 %	46.7 %			1	0.38	0.54
Educational level (CBS)	2.5 (1.7)	2.3 (1.2)	−0.61	0.14	63		0.55
MMSE	27.9 (1.6)	28.0 (1.8)	0.13	0.06	63		0.90
CST, subtest C (seconds)	44.6 (13.0)	43.3 (19.5)	−0.51	0.08	63		0.76
Fluency (named animals)	21.9 (5.9)	21.1 (6.6)	−0.51	0.13	63		0.61
LDST (items after 90 seconds)	39.8 (7.7)	38.6 (8.5)	−0.63	0.15	63		0.53
Stroop, Interference (seconds)	113.4 (25.0)	117.7 (26.3)	−0.67	0.17	61		0.50
VLT immediate recall (words)	10.9 (2.8)	10.5 (1.8)	−0.61	0.17	63		0.55
VLT delayed recall (words)	9.3 (2.7)	7.9 (2.5)	−2.01	0.54	63		0.05*
t_3							
Intracranial volume (cm^3)	1463.7 (177.6)	1487.1 (152.7)	0.57	0.14	63		0.57
MMSE	28.2 (1.5)	26.3 (2.0)	−4.49	1.07	63		<0.01**
CST, subtest C (seconds)	41.7 (13.5)	45.6 (12.1)	1.21	0.30	62		0.23
Fluency (named animals)	21.4 (4.6)	18.3 (5.5)	−2.51	0.61	63		0.02*
LDST (items after 90 seconds)	40.1 (7.8)	38.7 (9.8)	−0.64	0.16	63		0.53
Stroop, Interference (seconds)	118.2 (29.0)	128.9 (32.9)	1.38	0.35	61		0.17
VLT immediate recall (words)	11.5 (1.9)	10.8 (2.0)	−1.43	0.36	63		0.16
VLT delayed recall (words)	9.9 (2.5)	7.6 (2.4)	−3.70	0.94	63		<0.01**
Hypertension treated	7	5			1	0.12	0.73
Hypertension untreated	7	4			1	0.51	0.48
Other vascular disease	9	13			1	2.24	0.14
Diabetes insulin dependent	0	1			1	1.19	0.28
Diabetes insulin independent	3	6			1	1.77	0.18
Medication cardiovascular	10	9			1	0.02	0.90
Medication diabetes	2	4			1	1.12	0.29
Medication other	9	8			1	0.01	0.93
t_6							
MMSE	28.4 (1.6)	27.5 (2.0)	−2.05	0.50	61		0.04*
CST, subtest C (seconds)	49.6 (17.1)	53.1 (21.4)	0.73	0.18	60		0.47
Fluency (named animals)	19.8 (3.8)	18.7 (5.8)	−0.96	0.22	61		0.34
LDST (items after 90 seconds)	40.0 (8.6)	37.7 (9.7)	−1.01	0.25	60		0.32
Stroop, Interference (seconds)	116.4 (24.5)	134.3 (47.4)	1.91	0.47	59		0.06
VLT immediate recall (words)	11.9 (2.3)	9.7 (2.6)	−3.66	0.90	61		<0.01**
VLT delayed recall (words)	10.2 (2.9)	7.0 (2.9)	−4.29	1.10	61		<0.01**

Note. The characteristics of the two groups were compared at t_0, t_3 and t_6 with independent t-tests (continuous variables) and Chi-square tests (categorical variables). Cognitively healthy = participants with no cognitive decline in the 3 years before MRI and who did not develop dementia within 6 years after MRI (n = 35); Cognitive decline = participants with significant cognitive decline who did not develop dementia (n = 30); CBS = 7-point score for educational level (1 = elementary school; 7 = University) (Van den Brandt et al., 1990); MMSE = Mini-Mental State Examination; CST = Concept Shifting Test; LDST = Letter Digit Substitution Test; VLT = Verbal Learning Test; t_0 = baseline; t_3 = 3 years after baseline; t_6 = 6 years after baseline; d = Cohen's d (effect size); df = degrees of freedom; t = t value. Significant difference between the two groups: * $p < .05$. ** $p < .01$.

stable 3 years after scanning, the age effect in the hippocampus disappeared (F = 2.43, p = .131, r = .29, df = 1). The cognitive decline group exhibited a significant age effect on all measured brain structures. In addition, there was a significant age by group interaction in the inferior prefrontal cortex (F = 6.45, p = .014, df = 1) and orbital prefrontal cortex (F = 4.54, p = .037, df = 1), which indicates that the two groups differed significantly in their age effect on these volumes. Moreover, when we excluded the seven "decliners" from the cognitively healthy group, namely those who were no longer cognitively stable 3 years after scanning, there was also a significant age by group interaction in the hippocampus (F = 4.35, p = .042, df = 1).

To gain more insight into the differences between the two groups, the effect sizes of the age effects were calculated. The effect size in the cognitively healthy group was medium in the hippocampus (r = .40) and small in the other regions (r = .08 to 0.20, see Table 2). The effect size of the age effects in the cognitive decline group was large in all regions (r = .43 to .65). Power analyses (a priori power analyses in G*Power 3, with a power of 0.8, an α-level of 0.05, and the effect sizes presented in Table 2) demonstrated that we would have needed a very large number of participants to obtain significant volume loss in these regions of the cognitively healthy group (parahippocampal gyrus: 425; inferior prefrontal cortex: 962; orbital prefrontal cortex: 150; dorsolateral prefrontal cortex: 311; anterior cingulate gyrus: 270; posterior cingulate gyrus: 167).

Table 2 also includes the means of the individual brain areas for each group. For most brain areas, the cognitive decline group had

Table 2
Effect of Age On Gray Matter Volume in the Cognitively Healthy and Cognitive Decline Group. Age Effects in the Cognitively Healthy Group Were Performed Twice: With (N = 35) and Without (N = 28) the 7 Participants Who Were No Longer Cognitively Stable 3 Years After Scanning

Cortical region	Volume in ml (*SD*)	Volume as % of ICV (*SD*)	Age effect			
			F	*r*	*df*	*p*
Cognitively healthy group (n = 35)						
Hippocampus	7.40 (1.12)	0.51 (0.06)	6.33	0.40	1	0.017*
Parahippocampal gyrus	3.46 (0.83)	0.24 (0.05)	0.46	0.12	1	0.502
Inferior prefrontal cortex	10.38 (2.86)	0.71 (0.17)	0.18	0.08	1	0.671
Orbital prefrontal cortex	8.57 (2.28)	0.59 (0.14)	1.46	0.20	1	0.236
Dorsolateral prefrontal cortex	28.15 (6.61)	1.92 (0.38)	0.67	0.14	1	0.419
Anterior cingulate gyrus	7.48 (1.92)	0.51 (0.12)	0.77	0.15	1	0.386
Posterior cingulate gyrus	3.04 (0.98)	0.21 (0.06)	1.25	0.19	1	0.271
Cognitive decline group (n = 30)						
Hippocampus	6.84 (1.26)	0.46 (0.08)	16.97	0.61	1	<0.001**
Parahippocampal gyrus	3.21 (0.76)	0.22 (0.05)	10.53	0.52	1	0.003**
Inferior prefrontal cortex	10.67 (2.54)	0.72 (0.19)	20.01	0.65	1	<0.001**
Orbital prefrontal cortex	8.32 (2.44)	0.56 (0.16)	20.60	0.65	1	<0.001**
Dorsolateral prefrontal cortex	27.18 (7.25)	1.84 (0.51)	7.49	0.47	1	0.011*
Anterior cingulate gyrus	7.68 (2.18)	0.52 (0.15)	6.38	0.43	1	0.018*
Posterior cingulate gyrus	3.08 (1.15)	0.21 (0.08)	9.17	0.50	1	0.005**
Cognitively healthy group minus seven participants who were no longer cognitively stable 3 years after scanning (n = 28)						
Hippocampus	7.47 (1.05)	0.51 (0.05)	2.43	0.29	1	0.131
Parahippocampal gyrus	3.46 (0.88)	0.24 (0.05)	0.72	0.16	1	0.403
Inferior prefrontal cortex	10.48 (2.93)	0.71 (0.18)	0.11	0.06	1	0.744
Orbital prefrontal cortex	8.79 (2.34)	0.60 (0.14)	0.32	0.11	1	0.575
Dorsolateral prefrontal cortex	28.52 (7.01)	1.94 (0.39)	0.25	0.10	1	0.620
Anterior cingulate gyrus	7.36 (1.93)	0.50 (0.11)	0.57	0.15	1	0.456
Posterior cingulate gyrus	3.01 (1.07)	0.20 (0.06)	0.83	0.18	1	0.370

Note. The univariate ANOVA under the General Linear Model (GLM) menu was used to test the age effect on the individual brain areas. The mean volumes of the brain areas are presented in ml. They were calculated as the sum of the left and right hemisphere and as percentage of the intracranial volume. Cognitively healthy group = participants with no cognitive decline in the 3 years before MRI who did not develop dementia within 6 years after MRI; Cognitive decline group = participants with significant cognitive decline who did not develop dementia; r = Pearson's correlation coefficient (effect size); df = degrees of freedom; F = F value. Significant outcome: * $p < .05$. ** $p < .01$.

smaller mean volumes than the cognitively healthy group. However, this was not the case in the inferior prefrontal cortex and cingulate gyrus. The reason for this is that the three youngest participants (age ≤60 years) in the decline group had relatively large volumes. When we excluded all participants who were 60 years old and younger, all brain volumes were smaller in the decline group than in the cognitively healthy group.

The fact that the "young" individuals in the cognitive decline group had larger volumes than the "young" individuals in the cognitively healthy group could have biased the results. To test whether or not this could explain the age effects in the cognitive decline group, all GLM tests were also done without the participants who were 60 years old or younger (n = 6). The age effects in the cognitive decline group remained significant in six out of seven measured brain structures. Only the age effect on the dorsolateral PFC volume (F = 3.85, p = .061, r = .37, df = 1) did not reach significance. Excluding the youngest participants did not change the results in the cognitively healthy group. We found that the mean volume of the three "young" decliners (age ≤60 years) was larger than the mean volume of the three "young" nondecliners in all measured brain areas. However, the total brain volume (measured by the FSL Brain Extraction Tool, expressed as percentage of the intracranial volume) was slightly smaller in the "young" decliners (67.3% in the young decliners vs. 69.1% in the young nondecliners).

We did not find indications for quadratic age effects. Furthermore, we did not find a relationship between intracranial volume and age. All data met the assumptions of normality, linearity and independency (i.e., the values of the outcome variables were independent). There was only one outlier present in the data: one participant of the healthy group had a very large orbital PFC. However, our findings did not change when this outlier was excluded.

Discussion

The present study did not reveal an age effect on the gray matter volume of the parahippocampal, cingulated, and frontal areas in cognitively healthy participants who remained cognitively stable over 3 years; whereas a significant age effect in these brain areas

Table 3
Age by Group Interactions. Interactions of the Cognitively Healthy Group Were Calculated Twice: With (N = 35) and Without (N = 28) the Seven Participants Who Were No Longer Cognitively Stable 3 Years After Scanning

	Age by group interaction			Age by group interaction		
	Healthy (n = 35); decliners (n = 30)			Healthy (n = 28); decliners (n = 30)		
Cortical region	F	df	p	F	df	p
Hippocampus	2.65	1	0.109	4.35	1	0.042*
Parahippocampal gyrus	2.69	1	0.106	1.68	1	0.201
Inferior prefrontal cortex	6.45	1	0.014*	6.11	1	0.017*
Orbital prefrontal cortex	4.54	1	0.037*	5.96	1	0.018*
Dorsolateral prefrontal cortex	2.81	1	0.099	2.88	1	0.095
Anterior cingulate gyrus	1.96	1	0.166	1.89	1	0.175
Posterior cingulate gyrus	2.58	1	0.113	2.29	1	0.136

Note. The univariate ANOVA under the General Linear Model (GLM) menu was used to calculate age by group interactions. Cognitively healthy grou$p$ = participants with no cognitive decline in the 3 years before MRI who did not develop dementia within 6 years after MRI; Cognitive decline grou$p$ = participants with significant cognitive decline who did not develop dementia; df = degrees of freedom; F = F value. Significant outcome: * $p < .05$.

was found in participants with substantial cognitive decline. The hippocampus was the only structure in which an age effect was present in both groups. However, the age effect in the hippocampus of the cognitively healthy group disappeared when we excluded the participants who were no longer cognitively stable 3 years after the scanning.

These findings suggest that gray matter atrophy in the hippocampus, parahippocampal, cingulated, and frontal areas is to a lesser extent associated with the healthy aging process, but more likely with brain processes underlying significant cognitive decline. In other words, as long as people stay cognitively healthy, there may be no substantial gray matter atrophy in several brain areas that are highly associated with cognition. These results indicate that the amount of cortical gray matter atrophy in the aging brain may be overestimated in a large number of studies on healthy aging (Raz et al., 2005; Resnick et al., 2003; Sowell et al., 2003; Tisserand et al., 2002), because these studies were not able to exclude subjects with an atypical cognitive decline or preclinical dementia. Nevertheless, we should remain cautious about this conclusion, because longitudinal MRI data obtained from a larger number of subjects are required to be certain. Our results are in line with research that takes into account substantial longitudinal changes in cognition, such as the study of Kramer et al. (Kramer et al., 2007).

Given the number of participants studied and the fact that the MRI data was cross-sectional, we cannot rule out that there may be a small age effect in all of the assessed brain areas of the cognitively healthy group. Nevertheless, the data does support the conclusion that the age effect in previous studies may partly have been caused by the inclusion of participants with subclinical cognitive disorders. First, the effect sizes of six brain structures in the cognitively healthy group were small, whereas the effect sizes of all brain structures in the cognitive decline group were large. Power analyses demonstrated that we would have needed a very large number of participants to render the volume loss significant in these regions in the cognitively healthy group. This indicates that the age effect in these regions is very weak or absent in cognitively healthy individuals, whereas it is very significant in the cognitive decliners. Second, there were significant age by group interactions in the inferior prefrontal and orbital prefrontal cortex; there was a trend in the dorsolateral prefrontal cortex; and there was a significant interaction in the hippocampus when we excluded the seven "prodromal decliners" from the healthy group. Although the age by group interactions did not reach significance in all brain areas, their presence supports our conclusion that the two groups differ with respect to age effect on volume.

In summary, our findings stress the importance of inclusion criteria for studies on normal brain aging. The fact that our results do not match those of several earlier studies on healthy aging may partly be a result of how "normal" or "healthy" aging is defined. Most of the earlier studies defined "healthy" as 'being free from diseases,' whereas in our study "healthy" also means an absence of preclinical dementia and other subclinical pathological brain processes that result in a substantial cognitive decline or a low MMSE score.

A few considerations about the results of the current study need to be noted. First, in Figure 3 the youngest participants of the decline group seem to have larger volumes than the youngest participants of the healthy group. This may have been caused by the relatively small number of "young" participants in our sample. Only three members of the decline group and three healthy participants were younger than 60. Therefore, it is likely that the "young" decliners had larger volumes by coincidence. However, there may be an alternative explanation. It is speculative, but atrophy may be preceded by swelling because of edema as the result of an inflammatory reaction to the degenerative process. This fits in with our finding in the "young" participants (age $\leq$60). The mean gray matter volume of the three "young" decliners was larger than the mean gray matter volume of the three "young" nondecliners in all measured brain areas, but the total brain volume was slightly smaller in the "young" decliners. Thus the three "young" decliners seemed to have average brain volumes, but higher than average gray matter volumes in brain regions that are highly associated with cognition. This might have been caused by

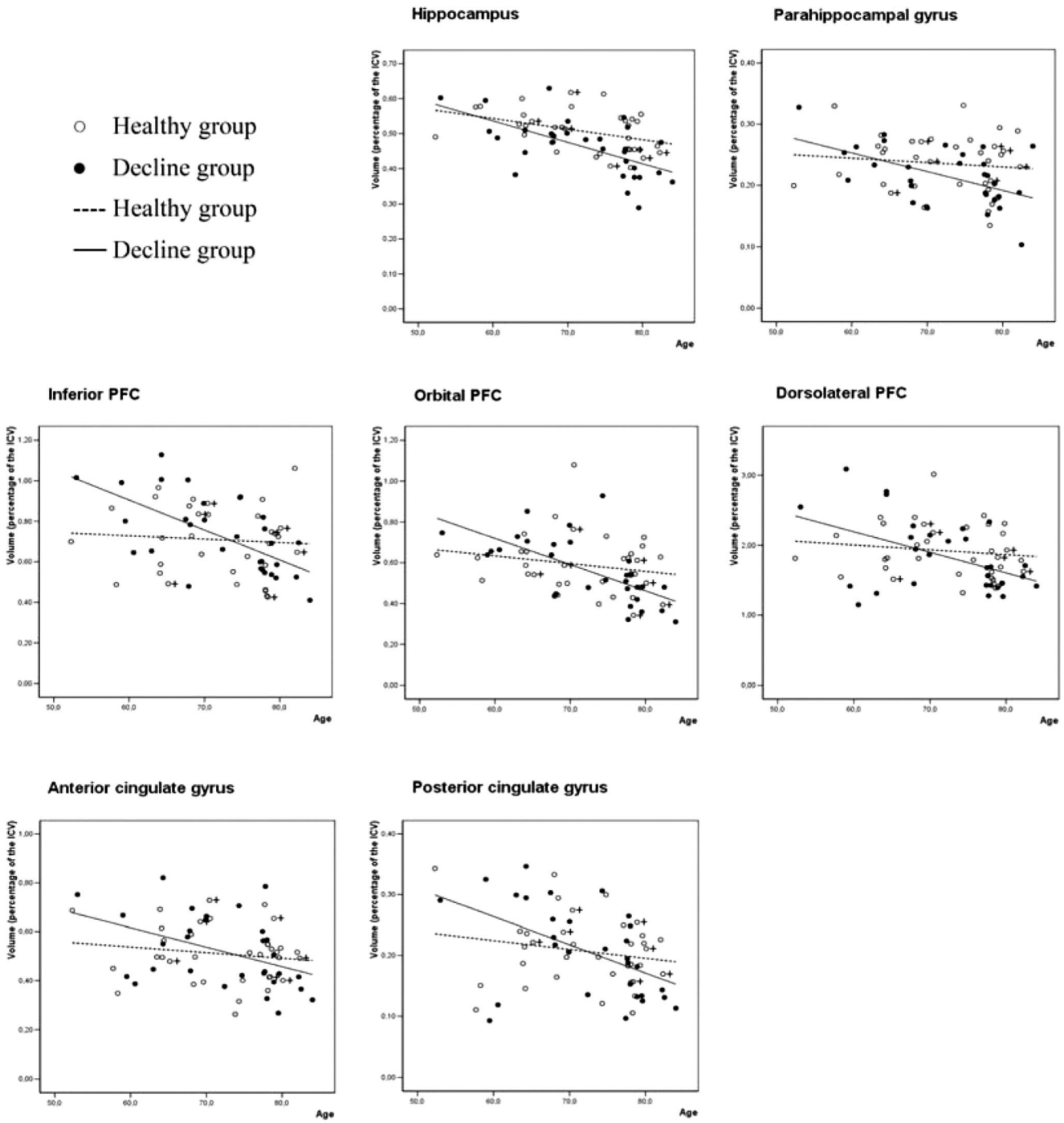


Figure 3. Scatter plots with regression lines demonstrating a significant age effect for the decline group and no age effect for the healthy group. Healthy group = participants with no cognitive decline in the 3 years before MRI who did not develop dementia ($n = 35$); the data points of the seven prodromal decliners (participants of the healthy group who appeared cognitive decliner 3 years after scanning) are marked with a +; Decline group = participants with significant cognitive decline in the 3 years before MRI who did not develop dementia ($n = 30$). Volumes were calculated as the sum of the left and right hemisphere and are expressed as percentage of the intracranial volume.

swelling. Nevertheless, the large volumes of the youngest members of the decline group were not responsible for the age effect in the cognitive decline group. When we excluded all Participants 60 years and younger, the age effects remained significant in six out of seven measured brain structures. Only the age effect on the dorsolateral PFC volume ($F = 3.85$, $p = .061$, $r = .37$) did not reach significance after excluding the youngest participants. The volumes of the youngest members were also not responsible for

the nonsignificant age effects in the cognitively healthy group, because excluding these participants did not change the results.

Second, we must consider the nature of the cognitive change within the decline group. The inclusion criteria for the decline group were based on different cognitive domains. As a result, the type of cognitive decline in this group was diverse. Some participants exhibited a specific memory decline and met the criteria for mild cognitive impairment as defined by Petersen (Petersen et al., 1999), whereas others exhibited a decline in information processing speed. Because of this diversity, the average decline on individual tasks was only moderate. When we compared the cognitive test scores of t_0 and t_3, we found a significant decline in the MMSE, the Fluency task, and the Stroop interference (results not shown), but there was no significant decline with respect to the verbal memory tasks. This could be a result of the diversity of the group, but it might also be caused by a learning effect, which is often present in longitudinal studies on cognitive aging.

In conclusion, our results demonstrated that cognitively healthy aging participants did not exhibit a significant age effect in the volume of the measured brain areas, whereas participants with a substantial cognitive decline did exhibit a significant age effect in the examined brain structures. This finding is in line with the conclusion that substantial cortical gray matter atrophy is not a typical characteristic of the healthy aging process, and that it is more accurately associated with significant cognitive decline. Our results indicate that the age effect found in many studies on healthy aging may have been overestimated, because in these studies it was not possible to exclude subjects with a substantial atypical cognitive decline or preclinical dementia. Finally, our results stress the importance of establishing stringent inclusion criteria for future studies on normal aging.

References

APA. (1994). *Diagnostic and statistical manual of mental disorders IV.* Washington DC: American Psychiatric Association.

Burgmans, S., van Boxtel, M. P. J., Smeets, F., Vuurman, E. F. P. M., Gronenschild, E. H. B. M., Verhey, F. R. J., et al. (2008). Prefrontal cortex atrophy predicts dementia over a six-year period. *Neurobiology of Aging.*

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, 39,* 175–191.

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

John, J. P., Wang, L., Moffitt, A. J., Singh, H. K., Gado, M. H., & Csernansky, J. G. (2006). Inter-rater reliability of manual segmentation of the superior, inferior and middle frontal gyri. *Psychiatry Research, 148,* 151–163.

Jolles, J., Houx, P. J., Van Boxtel, M. P. J., & Ponds, R. W. H. M. (1995). *Maastricht Aging Study: Determinants of cognitive aging.* Maastricht: Neuropsych Publishers.

Jones, B. F., Barnes, J., Uylings, H. B. M., Fox, N. C., Frost, C., Witter, M. P., et al. (2006). Differential regional atrophy of the cingulate gyrus in Alzheimer Disease: A volumetric MRI study. *Cerebral Cortex, 16,* 1701–1708.

Kramer, J. H., Mungas, D., Reed, B. R., Wetzel, M. E., Burnett, M. M., Miller, B. L., et al. (2007). Longitudinal MRI and cognitive change in healthy elderly. *Neuropsychology, 21,* 412–418.

Lamberts, H., & Wood, M. (2002). The birth of the International Classification of Primary Care (ICPC). Serendipity at the border of Lac Leman. *Family Practice, 19,* 433–435.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34,* 939–944.

Nyúl, L. G., & Udupa, J. K. (1999). On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine, 42,* 1072–1081.

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Archives of Neurology, 56,* 303–308.

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.* Mahwah, NJ: Erlbaum.

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

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., et al. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15,* 1676–1689.

Raz, N., Rodrigue, K. M., Kennedy, K. M., & Acker, J. D. (2007). Vascular health and longitudinal changes in brain and cognition in middle-aged and older adults. *Neuropsychology, 21,* 149–157.

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.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 2,* 420–428.

Sled, J. G., Zijdenbos, A. P., & Evans, A. C. (1998). A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Transactions On Medical Imaging, 17,* 87–97.

Sliwinski, M., & Buschke, H. (1999). Cross-sectional and longitudinal relationships among age, cognition, and processing speed. *Psychology and Aging, 14,* 18–33.

Sliwinski, M., Lipton, R. B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 51,* P217–225.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17,* 143–155.

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6,* 309–315.

Thangavel, R., Van Hoesen, G. W., & Zaheer, A. (2008). Posterior parahippocampal gyrus pathology in Alzheimer's disease. *Neuroscience, 154,* 667–676.

Tisserand, D. J., Pruessner, J. C., Sanz Arigita, E. J., Van Boxtel, M. P. J., Evans, A. C., Jolles, J., et al. (2002). Regional frontal cortical volumes decrease differentially in aging: An MRI study to compare volumetric approaches and voxel-based morphometry. *Neuroimage, 17,* 657–669.

Uylings, H. B. M., Malofeeva, L. I., Bogolepova, I. N., Jacobsen, A. M., Amunts, K., & Zilles, K. (2005). No postnatal doubling of number of neurons in human Broca's areas (Brodmann areas 44 and 45)? A stereological study. *Neuroscience, 136,* 715–728.

Van Boxtel, M. P. J., Buntinx, F., Houx, P. J., Metsemakers, J. F., Knottnerus, A., & Jolles, J. (1998). The relation between morbidity and

cognitive performance in a normal aging population. *Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 53,* M147–154.

Van den Brandt, P. A., Goldbohm, R. A., Van 't Veer, P., Volovics, A., Hermus, R. J., & Sturmans, F. (1990). A large-scale prospective cohort study on diet and cancer in the Netherlands. *Journal of Clinical Epidemiology, 43,* 285–295.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J., & Jolles, J. (2005). Rey's verbal learning test: Normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society, 11,* 290–302.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J., & Jolles, J. (2006a). The Concept Shifting Test: Adult normative data. *Psychological Assessment, 18,* 424–432.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J., & Jolles, J. (2006b). The Letter Digit Substitution Test: Normative data for 1,858 healthy participants aged 24–81 from the Maastricht Aging Study (MAAS): Influence of age, education, and sex. *Journal of Clinical and Experimental Neuropsychology, 28,* 998–1009.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J., & Jolles, J. (2006c). Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society, 12,* 80–89.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J., & Jolles, J. (2006d). The Stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment, 13,* 62–79.

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.

Zijdenbos, A. P., Forghani, R., & Evans, A. C. (2002). Automatic "pipeline" analysis of 3-D MRI data for clinical trials: Application to multiple sclerosis. *IEEE Transactions on Medical Imaging, 21,* 1280–1291.

Received July 7, 2008
Revision received March 16, 2009
Accepted March 30, 2009 ■

New Editors Appointed, 2011–2016

The Publications and Communications Board of the American Psychological Association announces the appointment of 3 new editors for 6-year terms beginning in 2011. As of January 1, 2010, manuscripts should be directed as follows:

- *Developmental Psychology* (http://www.apa.org/journals/dev), **Jacquelynne S. Eccles, PhD,** Department of Psychology, University of Michigan, Ann Arbor, MI 48109
- *Journal of Consulting and Clinical Psychology* (http://www.apa.org/journals/ccp), **Arthur M. Nezu, PhD,** Department of Psychology, Drexel University, Philadelphia, PA 19102
- *Psychological Review* (http://www.apa.org/journals/rev), **John R. Anderson, PhD,** Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213

Electronic manuscript submission: As of January 1, 2010, manuscripts should be submitted electronically to the new editors via the journal's Manuscript Submission Portal (see the website listed above with each journal title).

Manuscript submission patterns make the precise date of completion of the 2010 volumes uncertain. Current editors, Cynthia García Coll, PhD, Annette M. La Greca, PhD, and Keith Rayner, PhD, will receive and consider new manuscripts through December 31, 2009. Should 2010 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 2011 volumes.