

Measuring cognitive reserve based on the decomposition of episodic memory variance

Bruce R. Reed,^{1,2} Dan Mungas,¹ Sarah Tomaszewski Farias,¹ Danielle Harvey,³ Laurel Beckett,³ Keith Widaman,⁴ Ladson Hinton¹ and Charles DeCarli¹

1 Department of Neurology, School of Medicine, University of California, Davis, CA, USA

2 Veterans Administration Northern California Health Care System, Martinez, CA 94553, USA

3 Department of Public Health Sciences, School of Medicine, University of California, Davis, CA 95616, USA

4 Department of Psychology, University of California, Davis, CA 95616, USA

Correspondence to: Bruce Reed,

UC Davis Alzheimer's Centre,

150 Muir Road (127a),

Martinez,

CA 94553, USA

E-mail: brreed@ucdavis.edu

In later adulthood brain pathology becomes common and trajectories of cognitive change are heterogeneous. Among the multiple determinants of late-life cognitive course, cognitive reserve has been proposed as an important factor that modifies or buffers the impact of brain pathology on cognitive function. This article presents and investigates a novel method for measuring and investigating such factors. The core concept is that in a population where pathology is common and variably present, 'reserve' may be defined as the difference between the cognitive performance predicted by an individual's level of pathology and that individual's actual performance. By this definition, people whose measured cognitive performance is better than predicted by pathology have high reserve, whereas those who perform worse than predicted have low reserve. To test this hypothesis, we applied a latent variable model to data from a diverse ageing cohort and decomposed the variance in a measure of episodic memory into three components, one predicted by demographics, one predicted by pathology as measured by structural MRI and a 'residual' or 'reserve' term that included all remaining variance. To investigate the plausibility of this approach, we then tested the residual component as an operational measure of reserve. Specific predictions about the effects of this putative reserve measure were generated from a general conceptual model of reserve. Each was borne from the results. The results show that the current level of reserve, as measured by this decomposition approach, modifies rates of conversion from mild cognitive impairment to dementia, modifies rates of longitudinal decline in executive function and, most importantly, attenuates the effect of brain atrophy on cognitive decline such that atrophy is more strongly associated with cognitive decline in subjects with low reserve than in those with high reserve. Decomposing the variance in cognitive function scores offers a promising new approach to the measure and study of cognitive reserve.

Keywords: cognitive ageing; mild cognitive impairment; dementia; neuropsychological tests; longitudinal change; cognitive reserve

Abbreviations: CDR = Clinical Diagnostic Rating; Mem-D = component of episodic memory related to demographic variables; Mem-B = component of episodic memory related to MRI variables; Mem-R = component of episodic memory unrelated to demographic and MRI variables; SENAS = Spanish and English Neuropsychological Assessment Scale

Introduction

The later decades of life are characterized, at the population level, by worsening cognitive decline (Salthouse, 2003). Individual trajectories, however, are highly heterogeneous, with some people declining rapidly and experiencing severe impairment and others changing little or even improving (Zelinski *et al.*, 1993; Albert *et al.*, 1995; Rubin *et al.*, 1998; Wilson *et al.*, 2002a). Brain pathology, including that of Alzheimer's disease and other neurodegenerative processes, infarcts and other ischaemic damage, and sometimes a cryptic loss of neurons and synapses (White *et al.*, 2005), accumulate to varying degrees in individuals and explain some of the heterogeneity in cognitive ageing (Buckner, 2004). The relationship between measured pathology and cognitive decline, however, is only modest overall (Mungas *et al.*, 2002; Chui *et al.*, 2006) and striking discrepancies are common. For example, ~20% of old people who are cognitively normal and stable prior to death will be found to have sufficient Alzheimer's disease pathology post-mortem to meet neuropathological criteria for Alzheimer's disease (Davis *et al.*, 1999; Schmitt *et al.*, 2000; Bennett *et al.*, 2006; Driscoll *et al.*, 2006). Similarly, population-based studies that use MRI to sensitively study brain structure find that ~20% of older adults have previously unsuspected infarcts (Longstreth *et al.*, 1998; DeCarli *et al.*, 2005).

'Reserve' is the concept that there are characteristics of the brain that buffer the impact of pathology on brain performance (Satz, 1993; Stern, 2002). These characteristics may be structural, e.g. 'extra' neurons and synapses, or may be functional: a degree of excess capacity or compensatory mechanisms such that the brain could continue to perform well despite damage (Stern, 2009). Reserve is postulated to differ between people and to be determined by both biology and experience, with cognitively demanding activities playing a key role. It is an important concept in theories of how the brain responds to insult and is frequently invoked to explain discrepancies between the extent of brain pathology and the extent of cognitive impairment.

A critical question for the study of reserve is how to quantify this abstract construct. The importance of how to measure reserve seems self evident, but the issue has received little formal attention. The most common approach is to use other variables as markers of, or proxies for, reserve; the most common such variable being years of education. Thus, for example, multiple studies have found that level of education is inversely correlated with risk of prevalent (Zhang *et al.*, 1990; Katzman, 1993; Canadian Study of Health and Ageing Working Group, 1994; Ott *et al.*, 1995,) or incident (Cobb *et al.*, 1995; Karp *et al.*, 2004) dementia and have concluded that cognitive reserve plays a role in protecting people against dementing illness. However, education is a highly complex variable with wide ranging effects on people's lives, and thus education may reduce the risk of dementia through mechanisms having nothing to do with reserve. With a few exceptions, higher education confers lower risk of disease (Grossman and Kaestner, 1997). Education correlates with income, access to health care, health promoting behaviours and many other variables that predict health outcome (Kaplan *et al.*, 1987; Grossman and Kaestner, 1997; Deaton and Paxson, 1999).

Lower education is associated with increased risk of heart disease, stroke and diabetes (Liu *et al.*, 1982; Wong *et al.*, 2002; Borrell *et al.*, 2006), hypertension, smoking and obesity (Kaplan and Keil, 1993), each of which is associated with increased risk of dementia in general and Alzheimer's disease in particular. Education may indeed correlate with reserve but it does not measure reserve directly or exclusively.

Other variables have also been used as markers of reserve, for example occupational complexity (Andel *et al.*, 2006; Potter *et al.*, 2008) and extent of intellectual activities during leisure time (Wilson *et al.*, 2002b; Verghese *et al.*, 2003). However, these variables are themselves correlated with education and suffer from the same fundamental problems of being measures of factors thought to contribute to reserve as opposed to being direct measures of reserve itself and of introducing multiple potential confounds into any correlational analysis.

The aim of this article is to present an alternative approach to the measurement of reserve along with a series of analyses that support the utility of this approach. The rationale begins with the premise that a simple way to define reserve is the difference between an individual's expected cognitive performance, given a particular level of brain pathology, and their actual cognitive performance. So defined, a person whose observed cognitive performance is better than predicted—given the extent of pathology present in his or her brain—has high reserve. Conversely, a person with low reserve performs worse than expected.

The conceptual basis for our approach to quantify reserve rests on the observation that cognition is determined by multiple factors and the independent contributions of these determinants can be estimated. One determinant is neuropathology. Neuropsychological tests are sensitive to brain injury and dysfunction, although measures of brain injury or pathology usually leave substantial test variance unexplained. The particular measures of pathology used here are MRI-based measures of brain structure. A second determinant is demographic variables, which typically have strong relationships to cognitive test scores. Our approach to measuring reserve involves an empirical decomposition of variance in a cognitive variable (episodic memory in this case) into independent components explained by measures of brain structure, demographic variables and a component that is independent of brain structure and demographic effects. The latter, residual, component captures individual differences in cognition that are not explained by measured brain variables and provides our putative measure of reserve. That is, individuals with high scores on this component perform better than expected and those with low scores perform worse than expected; this corresponds to the previously presented definition of reserve. It is a measure of the current reserve of an individual, as opposed to an estimate of peak reserve or some sort of general adult level.

This approach, in a sense, defines reserve as prediction error. That is reserve is the sum of unmeasured sources of systematic variance in cognition, which is not explained by a specific set of known brain variables. A potential problem with this approach is that the measure of reserve that is derived will be influenced by the specific brain variables that are used and consequently, different measures of reserve are possible for a person at a specific time. The question is not so much whether this approach yields

a singular measure of reserve, but rather whether it provides a useful method for operationalizing reserve. There are several potential benefits of this approach, which has the core advantage of measuring reserve directly: (i) it provides an operational measure of current reserve that is quantitative, continuous and individually specific; (ii) it defines reserve *a priori* and concretely; (iii) it allows for measurement of change in reserve in an individual over time; and (iv) hypotheses regarding both the determinants and effects of reserve can be tested without circularity. These are very useful features and make it worth investigating as an approach to reserve.

In order to investigate this idea empirically, we used a latent variable modelling approach to decompose the variance of an episodic memory measure and to test relationships between those variance components and several key outcomes that were defined by the reserve hypothesis. For these initial analyses, we used episodic memory as the cognitive dependent measure. Other measures could have been used, but episodic memory changes substantially with age, is strongly affected by multiple brain disorders of ageing and is arguably the most sensitive cognitive indicator in a variety of diseases of ageing that cause cognitive decline. In order to determine whether the results hinged critically on using episodic memory as the measure to decompose, we performed a set of secondary analyses using a semantic memory measure as the dependent cognitive measure. Semantic memory can be conceptualized as a cumulative information store and can thus be viewed as reflective of the intellectual resources accumulated to that point in life. Therefore, one would predict that the extent to which semantic memory performance differed from the level predicted by brain pathology would be a measure of reserve.

The basic approach was to decompose the variance in episodic memory into three components, specifically the portion associated with demographic factors, the portion associated with MRI measures of brain structure and the residual or remaining portion; i.e. each subject's memory score was broken down into three scores—a demographic memory score, an MRI memory score and a residual memory score. The residual memory score was defined as a putative measure of reserve and tested using a set of *a priori* hypotheses.

While the basic conceptualization of reserve refers to systematic variance not explained by brain variables, we also incorporated demographic variables into our empirical decomposition of episodic memory. This has the practical result of removing demographic influences from the reserve measure and specifically removes any contributions of education. To the extent that education is central to the development of reserve, this choice will degrade the explanatory power of the reserve measure. Operationalizing reserve in this way presents a stringent and conservative test of the feasibility of this approach.

The general concept of reserve generates several specific hypotheses, which we tested using longitudinal data from a demographically diverse longitudinal cohort of older adults. The first two hypotheses reflect cross-sectional associations. The first is that mean levels of current reserve will diminish with current global cognitive status. Thus, normal > mild cognitive impairment > dementia, and continuous measures of global clinical status will correlate positively with reserve. This reflects the assumption that

the amount of reserve that an individual has will decline as that reserve is overwhelmed by pathology. The second hypothesis is that reserve, as measured by this approach, will be positively associated with other possible indices of reserve, even accounting for education. Thus, for example, reading ability should be positively associated with the residual/reserve variable. The next three hypotheses address the degree to which current reserve modifies our predictions of future decline. The third is that reserve modifies the risk of conversion from mild cognitive impairment to dementia; individuals with more reserve will be less likely to convert than those with lower reserve. Fourthly, we tested whether reserve modifies rates of cognitive decline such that greater reserve is associated with less longitudinal decline. The final hypothesis is that reserve modifies the effect of brain atrophy on cognitive function such that a given loss of brain volume will produce more cognitive decline in individuals with low reserve compared to those with higher reserve.

It should be noted that these hypotheses vary in the degree to which they are substantive tests of the proposed measurement approach to reserve. Hypothesis 1, in particular, is unlikely to be false and might be viewed as simply descriptive. The proposed reserve measure is a memory score adjusted for the effects of MRI and demographics; memory declines markedly across the cognitive syndromes of normal, mild cognitive impairment and dementia and thus it is highly likely that the adjusted score will too. Hypothesis 2 is also cross sectional and therefore less informative regarding reserve, which is fundamentally about change. The remaining hypotheses, which all test reserve as a modifier of longitudinal course, are, however, more substantial and therefore more crucial in testing conjectures about reserve.

Materials and methods

Participants

All 305 participants were evaluated by the University of California at Davis Alzheimer's Disease Centre as part of an ongoing longitudinal study of cognitive impairment in an educationally and ethnically diverse sample of older adults. Participants are recruited into the study through two routes: (i) memory clinic referrals and (ii) community outreach. Approximately 74% of participants were recruited through previously described community outreach protocols (Carter *et al.*, in press) designed to enhance both the racial and ethnic diversity and spectrum of cognitive dysfunction of the sample. The remaining participants were undergoing clinical evaluation at the University of California at Davis Alzheimer's Disease Centre and were recruited for this study. The participant sample included 101 African Americans, 78 Hispanics (35 tested in English, 43 tested in Spanish) and 126 Caucasians. Table 1 shows characteristics of the study sample.

Regardless of recruitment source, the inclusion criterion was that subjects were aged >60. Exclusion criteria included unstable major medical illness, major primary psychiatric disorder (history of schizophrenia, bipolar disorder or recurrent major depression) and substance abuse or dependence in the last five years. Participants who could not undergo MRI were excluded. All participants signed informed consent, and all human subject involvement was overseen by Institutional Review Boards at University of California at Davis, the Veterans

Table 1 Sample characteristics according to baseline diagnosis

	Normal (n = 162)	Mild cognitive impairment (n = 100)	Demented (n = 43)	All (n = 305)
Gender				
n (%) female	111 (69)	57 (57)	25 (58)	193 (63)
Education (years)				
mean (SD)	12.8 (4.4)	13.1 (4.6)	11.3 (4.6)	12.7 (4.5)
Age (years)***				
mean (SD)	73.1 (7.1)	75.1 (6.6)	78.1 (7.4)	74.5 (7.2)
Ethnicity*				
n (%) African American	58 (36)	33 (33)	10 (23)	101 (33)
n (%) Hispanic	49 (30)	17 (17)	12 (28)	78 (26)
n (%) Caucasian	55 (34)	50 (50)	21 (49)	126 (41)
Follow-up time (years) mean (SD)**	3.5 (1.3)	3.2 (1.4)	2.8 (1.4)	3.0 (1.3)
Number of evaluations				
n (%) 2 evaluations	28 (17)	24 (24)	13 (30)	65 (21)
n (%) 3 evaluations	41 (25)	29 (29)	12 (28)	82 (27)
n (%) 4 evaluations	46 (29)	17 (17)	10 (23)	73 (24)
n (%) ≥ 5 evaluations	47 (29)	30 (30)	8 (19)	85 (28)

Note that diagnostic groups differ. * $P=0.03$; ** $P=0.01$; *** $P<0.001$.

Administration Northern California Health Care System and San Joaquin General Hospital in Stockton, California.

Clinical evaluation

All participants received a multidisciplinary diagnostic evaluation through the University of California at Davis Alzheimer's Disease Centre. These evaluations included detailed medical history, physical and neurological examinations. A physician fluent in Spanish examined subjects who spoke only Spanish. A family member or other informant with close contact with the participant was interviewed to obtain information about level of independent functioning. Clinical neuropsychological evaluation using standard neuropsychological tests was given to each subject (these diagnostic tests are distinct from the outcome measures used in analyses). Routine dementia work-up laboratory tests were obtained for all participants.

Diagnosis of cognitive syndrome (normal, mild cognitive impairment, dementia) and, for individuals with dementia, underlying aetiology was made according to standardized criteria and methods. Each case was initially diagnosed at a consensus conference by the clinical team evaluating the participant. Those appearing likely to be eligible for this study were then reviewed at a second, multidisciplinary University of California at Davis Alzheimer's Disease Centre-wide case adjudication conference. Dementia was diagnosed using Diagnostic and Statistical Manual of Mental Disorders-III R (American Psychiatric Association, 1987) criteria for dementia modified to exclude the requirement of memory impairment. Mild cognitive impairment was diagnosed according to standard clinical criteria, modified to include the four different subtypes (Petersen, 2004). Normal cognitive function was diagnosed if there was no clinically significant cognitive or functional impairment. All subject diagnoses were made blind to research neuropsychological testing and quantitative brain image analysis. Of the 43 demented cases, 33 were diagnosed with Alzheimer's disease and five had mixed Alzheimer's and cerebrovascular disease; the others had vascular or other degenerative conditions. None had large cortical strokes or lobar atrophies.

The Clinical Diagnostic Rating (CDR) (Morris, 1993) was completed on the basis of a standardized interview with the identified participant and an informant; the sum of individual items or boxes (CDRSum) was used as a continuous measure of clinical status.

Research neuropsychological tests

All participants were administered the Spanish and English Neuropsychological Assessment Scale (SENAS), which is a comprehensive neuropsychological test battery that has undergone extensive development (Mungas *et al.*, 2000, 2004, 2005b, c; Gonzalez *et al.*, 2001, 2002). The SENAS was developed to be a battery of cognitive tests relevant to diseases of ageing that were psychometrically matched across sub-scales and across English and Spanish versions. Psychometric properties of the resulting scales have been published (Mungas *et al.*, 2004). This study used a subset of SENAS tests that were averaged within domains to create composite measures of Episodic Memory (Word List Learning I and Word List Learning II) and Executive Function (Category Fluency, Phonemic Fluency, Working Memory). The SENAS Episodic Memory composite was used in the decomposition. Secondary analyses were performed using a Semantic Memory composite from the SENAS as the cognitive variable for decomposition. This measure is a composite of an Object Naming task and a non-verbal, Picture Association task that is highly correlated with Object Naming (Mungas *et al.*, 2004). The correlation of Semantic Memory with Episodic Memory in this data set is 0.47.

Because cognitive decline was of particular interest and memory scores were used in the decomposition, a cognitive measure distinct from memory, the Executive Function composite, was used as the outcome. Other outcomes could have been used but executive function is an important indicator of brain changes and diseases of ageing, and is also closely tied to clinically relevant real world daily function (Royall *et al.*, 1998, 2004; Cahn-Weiner *et al.*, 2002, 2007; Tomaszewski Farias *et al.*, 2009). In order to test whether the results were critically dependent on using executive function as the outcome measure, a secondary analysis was performed using the CDRSum as the longitudinal outcome.

A subgroup of participants was tested with measures of word reading ability. Reading measures have been shown to be strong indicators of premorbid cognitive ability and appear to be proxy measures for quality of educational experience (Manly *et al.*, 2002, 2003, 2004) and have also been shown to be indicators of cognitive reserve (Manly *et al.*, 2005). Individuals tested in English were administered the American version of the National Adult Reading Test (Grober and Sliwinski, 1991). Those tested in Spanish were administered the Word Accentuation Test (Del Ser *et al.*, 1997). Using methods previously described, a composite reading measure was constructed by converting American version of the National Adult Reading Test scores of those tested in English and Word Accentuation Test scores of those tested in Spanish to z-scores (Cosentino *et al.*, 2007).

Magnetic resonance imaging

MRI was used to measure volumes of brain matter, white matter hyperintensities and hippocampus. While brain atrophy and white matter hyperintensities increase with age (Jack *et al.*, 1997; Liao *et al.*, 1997; DeCarli *et al.*, 2005), these are not benign changes and represent a variety of pathological effects (DeCarli *et al.*, 1995; Mungas *et al.*, 2001, 2005a; Kramer *et al.*, 2007). Alzheimer's disease is the single greatest cause of cognitive impairment in the aged, followed by cerebrovascular disease (Lim *et al.*, 1999; Fitzpatrick *et al.*, 2004). Both cause cortical atrophy (Whitwell *et al.*, 2007; Jagust *et al.*, 2008) and both are associated with increased volume of white matter hyperintensity (Jellinger, 2005; Yoshita *et al.*, 2006; Jagust *et al.*, 2008). Hippocampal volume is markedly reduced in Alzheimer's disease and also in hippocampal sclerosis, a condition associated with several aetiologies, including cerebrovascular disease (Jack *et al.*, 2002; Zarow *et al.*, 2005; Chui *et al.*, 2006). Each of these structural brain measures is associated with cognitive loss (Mungas *et al.*, 2001, 2002; Adak *et al.*, 2004; Tullberg *et al.*, 2004; Brickman *et al.*, 2008) and we used them together to estimate brain pathology. Infarcts are another type of brain pathology that is sensitively revealed by MRI. They were not included in initial models because previous analyses had shown their effects to be weak. The core latent variable model was built as described below. The addition of infarcts did not have any effect (independent of the other MRI variables) on episodic memory and therefore was not added to the model.

Research brain imaging was obtained at the University of California at Davis MRI research centre on a 1.5T GE Signa Horizon LX Echospeed system or the Veterans Administration at Martinez on a 1.5T Marconi system. The scan parameters and comparability of images from the two sites has been previously detailed (DeCarli *et al.*, 2008). Image quantification was performed by a rater who was blind to age, gender, race, educational achievement, ethnicity and diagnostic status.

Brain and white matter hyperintensities volumes

Analysis of brain and white matter hyperintensities volumes was based on a fluid attenuated inversion recovery sequence designed to enhance white matter hyperintensity segmentation (Jack *et al.*, 2001). Images were oriented parallel to a hypothetical line connecting the anterior commissure and posterior commissure. Brain and white matter hyperintensity segmentation was performed in a two-step process according to previously reported methods (DeCarli *et al.*, 1992, 1999). In brief, non-brain elements were manually removed from the image by operator guided tracing of the dura matter within the cranial vault,

including the middle cranial fossa, but excluding the posterior fossa and cerebellum. The resulting measure of the cranial vault was defined as the intracranial volume. Image intensity nonuniformities (DeCarli *et al.*, 1996) were then removed from the image and the resulting corrected image was modelled as a mixture of two Gaussian probability functions with the segmentation threshold determined at the minimum probability between these two distributions (DeCarli *et al.*, 1992). This resulted in a quantitation of total brain matter. Once brain matter segmentation was achieved, a single Gaussian distribution was fitted to the image data and a segmentation threshold for white matter hyperintensities was *a priori* determined at 3.5 SDs in pixel intensity above the mean of the fitted distribution of brain parenchyma. Morphometric erosion of two exterior image pixels was also applied to the brain matter image before modelling to remove the effects of partial volume CSF pixels and ventricular ependyma on white matter hyperintensity determination. Intra- and inter-rater reliability for these methods was high and has been published previously (DeCarli *et al.*, 2005).

Hippocampal volumes

Boundaries for the hippocampus were manually traced from the coronal 3D-T₁-weighted images after reorientation along the axis of the left hippocampus. While the borders were traced on the coronal slices, corresponding sagittal and axial views were simultaneously presented to the operator in separate viewing windows in order to verify hippocampal boundaries. The rostral end of the hippocampus was identified using the sagittal view to distinguish between amygdala and the head of the hippocampus. The axial view was used as a separate check. In anterior sections, the superior boundary of the hippocampus was the amygdala. In sections in which the uncus lies ventral to caudal amygdala, the uncus was included in the hippocampus. In more posterior sections that do not contain amygdala, the hippocampal (choroid) fissure and the superior portion of the inferior horn of the lateral ventricle formed the superior boundary. The fimbria were excluded from the superior boundary of the hippocampus. The inferior boundary of the hippocampus was the white matter of the parahippocampal gyrus. The lateral boundary was the inferior (temporal) horn of the lateral ventricle, taking care in posterior sections to exclude the tail of the caudate nucleus. The posterior boundary of the hippocampus was the first slice in which the fornices were completely distinct from any grey/white matter of the thalamus. Intrarater reliability determined for both right and left hippocampus volumes was good with intraclass correlation coefficients of 0.98 for right hippocampus and 0.96 for left hippocampus.

Statistical analysis

Overview

Data were analysed in two phases. In the first phase, a latent variable model was developed to decompose baseline episodic memory into three person-specific components: brain pathology, demographic variables and a latent (unmeasured) person-specific factor that captures difference from the baseline test performance predicted for an average person with similar brain pathology and demographics. We conceptualized this latent variable as cognitive reserve; the features that enable individuals to perform better (or worse) than the brain structure should allow. In the second phase, the latent variable model was expanded to test the five specific predictions, listed in the introduction, about how the latent variable for cognitive reserve should behave.

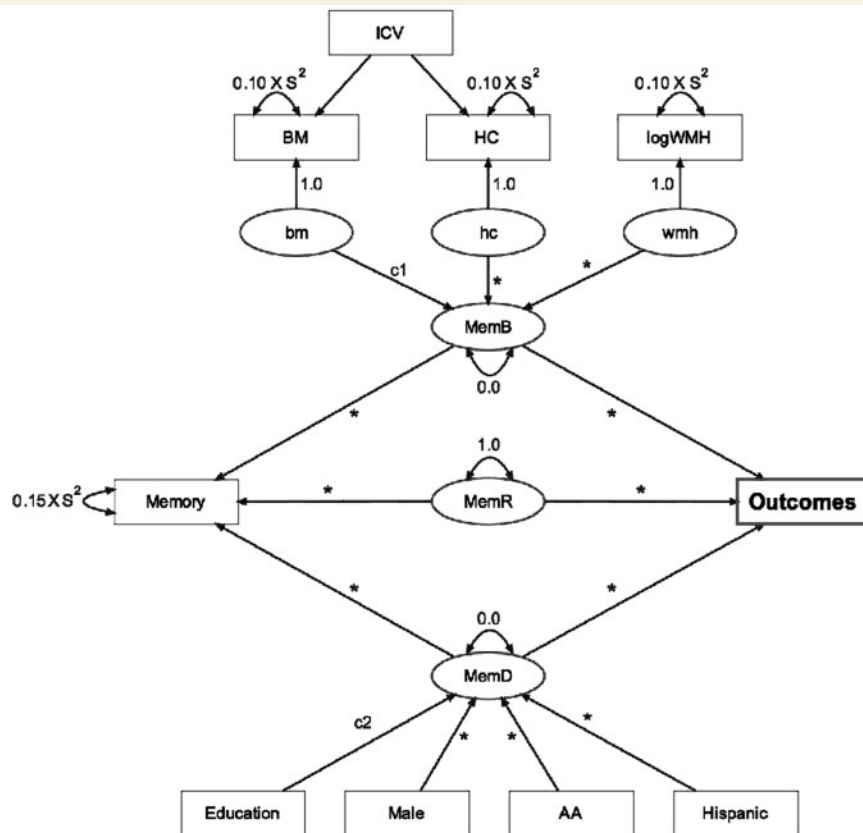


Figure 1 Analytic model for decomposing episodic memory into independent components and relating these components to external variables. Rectangles represent observed variables and ovals represent latent variables. Observed demographic and MRI variables were allowed to correlate freely (paths not shown). Freely estimated parameters are indicated by 'asterisk'. S^2 refers to sample variance. c1 and c2 are scaling constants selected to set variances at 1.0 for the MemB and MemD latent variables.

Latent variable model

A latent variable modelling framework was used to decompose the baseline SENAS episodic memory measure into components corresponding to demographics, MRI variables and a residual. These analyses were implemented with the Mplus application (Muthén and Muthén, 2007). Figure 1 shows the general analytic model. Rectangles refer to observed variables and ovals represent latent variables. The values for fixed model parameters are presented in Fig. 1 and freely estimated parameters are indicated by asterisks (*).

Observed demographic variables included in the model, shown at the bottom of Fig. 1, were years of formal education (quantitative) and indicators for gender (female as reference) and Hispanic or African American ethnicity (Caucasian as reference). Age was not included in the model as a predictor because prior work with this cohort indicated that the effects of age on cognition are entirely mediated through changes in brain volume (Mungas *et al.*, 2009). To verify that this assumption was correct, we examined the correlation of age with the residual term resulting from the model described and found it to be very small and statistically non-significant. Thus age does not contribute to the residual term. Observed MRI variables are shown at the top of Fig. 1. Raw brain matter and hippocampus volumes were regressed on total intracranial volume (intracranial volume) to obtain latent variables (brain matter and hippocampus volume) that are adjusted for intracranial volume and inversely correspond to degree of atrophy of these brain components. White matter hyperintensity was log transformed to normalize its distribution. It was also modelled through a latent variable, but was not adjusted for intracranial

volume. Residual variances of brain matter, hippocampus volume and white matter hyperintensity were fixed at 0.10 times the sample variances for these variables to correspond to conservatively estimated reliability of 0.90.

The centre of Fig. 1 shows three latent variables that represent person-specific characteristics assumed to account independently for episodic memory: Mem-B, Mem-D and Mem-R. A formative model (Bollen and Lennox, 1991; Edwards and Bagozzi, 2000) was used for Mem-B and Mem-D; these latent variables are essentially linear combinations of their observed indicators and conceptually, are considered to be formed or caused by their indicators. Mem-B in this formulation is a linear combination of the three MRI variables, with parameters representing regression coefficients of Mem-B on the three indicators. Mem-D is analogously related to the observed demographic variables. The observed variable, memory, in turn, is a reflective indicator of component of episodic memory related to MRI variables (Mem-B), component of episodic memory related to demographic variables (Mem-D) and component of episodic memory unrelated to demographic and MRI variables (Mem-R). Identification of this model was achieved by constraining the residual variances of Mem-B and Mem-D to 0.0, fixing the variance of Mem-R to 1.0 and by fixing the regression coefficient of Mem-B on one MRI variable and fixing the regression coefficient of Mem-D on one demographic variable. The specific values for the fixed regression coefficients for Mem-B and Mem-D were chosen to establish the variances of these latent variables at 1.0 to facilitate direct comparison of effects of the Mem-B, Mem-D and Mem-R variables in subsequent analyses. The residual variance of memory was

set at 0.15 times the sample variance, to account for measurement error and within-person variation, and to correspond to previously estimated reliability of 0.85 for this measure (Mungas *et al.*, 2004).

Mem-D was regressed on demographic variables and was constrained to have no residual correlation with MRI variables, thus it captures aspects of demographic variables that predict memory independent of brain pathology. Mem-B was linked to MRI variables but constrained to have no residual correlation with demographic variables. Residual correlations of memory with observed demographic and MRI variables were constrained to be zero, so that all relationships with these variables were incorporated into the paths through Mem-D and Mem-B. Correlations of demographic and MRI variables were freely estimated to reflect the natural correlation of these variables within the study sample. Consequently, Mem-D and Mem-B were not strictly orthogonal because of their relationships with potentially correlated MRI and demographic observed variables. Finally, Mem-R, our putative measure of reserve, was constrained to be uncorrelated with observed demographic and MRI variables, so that it accounts for otherwise unexplained systematic deviations of the memory scores. This variable was orthogonal to both Mem-D and Mem-B.

A preliminary set of analyses examined relationships among observed demographic and MRI variables. When variables were significantly related in preliminary analyses, non-zero paths for these variables were freely estimated in subsequent analyses (not shown in Fig. 1) but paths for variables that were not significantly related were constrained to zero.

The Mplus input file used to generate the decomposition of episodic memory into Mem-B, Mem-D and Mem-R is included in the [Supplementary material](#).

Tests of reserve hypotheses

To test each of the five conceptual hypotheses regarding cognitive reserve, additional variables were incorporated into the measurement model presented in Fig. 1 and their relationships to Mem-R, Mem-B and Mem-D were assessed. The criterion variables used in testing these hypotheses are labelled 'Outcomes' in Fig. 1. In general, relationships of the three memory components with the outcomes were tested in expanded models that estimated these relationships while simultaneously estimating the other free parameters noted in Fig. 1. Mplus supports a variety of analytic approaches for evaluating structural relationships of latent variables with external variables, including linear regression, ordinal logistic regression and Cox proportional hazard modelling (Muthén and Muthén, 2007).

Ordinal logistic regression was used to model baseline clinical diagnosis (normal, mild cognitive impairment, dementia) as a function of the three memory components (Hypothesis 1). Logistic regression, with a dichotomous dependent variable, models the probability of a specific outcome in terms of the odds, defined as the ratio (probability of the outcome)/(1–probability of the outcome). The regression coefficient for a specific independent variable is the estimated effect of a one-unit increase in that predictor, assumed to be a fixed change in the log odds, that is, a fixed percent change in the odds of the outcome. In ordinal logistic regression, the response levels are ordered and the model assumes that the effect of the predictor on the odds would be the same for any cut-point dividing the responses into two groups (normal versus mild cognitive impairment and dementia, or normal and mild cognitive impairment versus dementia). The model estimates separate intercepts corresponding to the reference odds for each possible cut-point.

Associations with CDR Sum of Boxes (Hypothesis 1) and Reading (Hypothesis 2) were evaluated in linear regression models. A Cox proportional hazards model was used to assess conversion to dementia,

excluding baseline demented cases (Hypothesis 3). Latent growth modelling (Meredith and Tisak, 1990; Muthén and Muthén, 2007; McArdle, 2009) was used to assess the relationship of the latent memory variables to longitudinal change in Executive Function (Hypothesis 4), estimating random effects for baseline performance and rate of change (Muthén and Muthén, 2007). A second model added terms for the interaction of Mem-B and Mem-R on baseline status and change (Hypothesis 5).

Since there were few cases with more than five evaluations, just the first five evaluations were used in the latent growth modelling for those cases. Missing values analyses were performed. Mplus uses full information maximum likelihood estimation with missing data and this provides unbiased estimation if data are missing completely, at random or missing at random. Missing data were primarily due to the study design issue of rolling enrolment of participants.

Model fit and significance testing

A maximum likelihood estimator was applied to a mean and variance-covariance data structure for all analyses except the ordinal logistic regression, where a mean and variance adjusted weighted least squares estimator was used. Latent variable modelling traditionally uses an overall chi square test of model fit, often supplemented by a number of fit indices to characterize model fit better. Commonly used fit indices include the comparative fit index (Bentler, 1990), the Tucker–Lewis index (Tucker and Lewis, 1973), the root mean square error of approximation (Browne and Cudek, 1993) and the standardized root mean squared residual (Bentler, 1995). The measurement model that decomposes memory in Fig. 1 into Mem-B, Mem-D and Mem-R is a just identified model when observed demographic and MRI variables are allowed to correlate freely. That is, the number of parameters in the model is the same as the number of elements in the variance covariance matrix that are being modelled and fit would be expected to be perfect. Constraining some of the correlations among observed demographic and MRI variables to zero results in an over-identified model, but fit should still be near perfect because these constraints were applied to pairs of variables whose empirical correlations were not significant. The structural part of the model in Fig. 1 that relates Mem-B, Mem-D and Mem-R to external outcomes adds regression paths from the three memory components to the outcomes. Statistical significance of these paths was evaluated by dividing a coefficient by its standard error (SE) to obtain a statistic that can be compared to the standard normal distribution to determine statistical significance.

Results

As expected, the measurement part of latent variable that decomposed episodic memory into the three components fit very well [$\chi^2(11) = 14.3$, $P = 0.21$; comparative fit index = 0.995, Tucker–Lewis index = 0.987, root mean square error of approximation = 0.032 (95% confidence interval = 0.000–0.072), standardized root mean squared residual = 0.030]. Brain pathology, measured through Mem-B, accounted for about 20–25% of variation in episodic memory, while demographics (education, gender and ethnicity) accounted for another 20%. The latent variable Mem-R, hypothesized to represent cognitive reserve, independently accounted for about 50% of the variance. Mem-R was uncorrelated by design with Mem-D and Mem-B. The correlation between Mem-D and Mem-B was –0.16. Intracranial volume was

not significantly correlated with Mem-R or either of the other two component terms. The maximum likelihood estimated mean and variance–covariance data structure for this analysis is presented in the [Supplementary material](#). Results of tests of specific hypotheses were as follows.

Hypothesis 1 (current reserve is associated with baseline clinical status)

All three latent variables were related to clinical diagnosis (Table 2). Mem-B had the strongest relationship; a 1 SD increase in Mem-B would reduce the odds of mild cognitive impairment or dementia versus normal by 57% and similarly would reduce the odds of dementia versus mild cognitive impairment or normal by 57%. A SD increase of 1 in Mem-R was associated with a 47% decrease, while a similar increase in Mem-D was associated with only a 19% decrease. Results were similar for the CDR Sum of Boxes (Table 3), a continuous measure of clinical status. Mem-B and Mem-R were independently related to this variable, with similar strengths of effects, while Mem-D was weakly related to it.

Hypothesis 2 (memory components will differentially relate to word reading)

As hypothesized, the reading measure was independently and positively related to Mem-D and Mem-R (Table 3). The relationship with Mem-D was strongest, as would be expected.

Table 2 Relationships of memory components with baseline clinical diagnosis

Memory component	Odds ratio	95% confidence interval	P-value
Mem-D	0.81	0.66–0.99	0.04
Mem-B	0.43	0.27–0.68	0.001
Mem-R	0.53	0.46–0.62	0.001

Tabled values are odds ratios from ordinal logistic regression of baseline diagnosis (0 = normal; 1 = mild cognitive impairment; 2 = demented) on memory components and represent the change in odds of a worse clinical diagnosis (e.g. a diagnosis reflecting greater cognitive impairment) associated with a one SD difference in the memory component.

Table 3 Relationship of memory components with baseline Reading and CDR Sum of Boxes (CDRSum)

Dependent variable	Memory component	Standardized coefficient	P-value
Reading	Mem-D	0.42	0.001
	Mem-B	0.15	NS
	Mem-R	0.20	0.01
CDRSum	Mem-D	−0.12	0.05
	Mem-B	−0.46	0.001
	Mem-R	−0.42	0.001

Tabled values are standardized regression coefficients from regressions of Reading and CDRSum on memory components. Results can be interpreted as correlations with the dependent variable independent of the other memory components. NS, not significant.

Hypothesis 3 (reserve modifies the risk of conversion to dementia)

Forty of 262 non-demented cases at baseline converted to a diagnosis of dementia over the course of follow-up, with an average of ~5% per year. The results in Table 4 show that a baseline Mem-B score that was 1.0 SD higher was associated with a ~80% reduced risk of converting to dementia, while a 1.0 SD difference in Mem-R was associated with a 65% reduced risk of converting, supporting the independent protective effects of greater brain volumes and increased reserve. Mem-D was not associated with conversion to dementia.

Hypothesis 4 (reserve affects longitudinal decline)

An initial random effects model included baseline Mem-D, Mem-B and Mem-R as independent variables and evaluated their effects on baseline executive function and change in executive function. Table 5 shows results from this analysis. All three components were strongly related to the estimated baseline executive function. Change in executive function was independently related to both Mem-B and Mem-R such that higher scores on these components at baseline predicted slower decline.

Hypothesis 5 (the association between brain atrophy and cognitive decline will be stronger in persons with low reserve than in persons with high reserve)

A second random effects model was estimated to address whether Mem-R moderated the effect of baseline Mem-B on longitudinal change in executive function. This model included baseline Mem-B and Mem-R as independent variables explaining baseline status and change, and also included terms representing the effects of the interaction of these two variables on baseline and change. The effects on longitudinal change were of primary interest. Mem-B ($\beta=0.072$, $SE=0.015$, $P<0.001$) and Mem-R ($\beta=0.068$, $SE=0.011$, $P<0.001$) were independently related to change, but in addition, there was a significant interaction effect of these two variables on longitudinal change ($\beta=-0.091$, $SE=0.018$, $P<0.001$). Mem-B had a stronger effect on executive function change in individuals with low Mem-R scores. Figure 2 shows

Table 4 Relative risk of converting to dementia associated with a 1.0 SD higher score for baseline episodic memory components, entered jointly as predictors

Memory component	Relative risk (confidence interval)
Mem-D	1.09 (0.74–1.59)
Mem-B	0.21 (0.13–0.36)
Mem-R	0.35 (0.25–0.49)

Results are statistically significant for Mem-B and Mem-R.

Table 5 Effects of baseline memory components on baseline and longitudinal change in executive function, expressed in SD units, based on random effects models for repeated measures

Effect estimated by model	Coefficient	SE	P-value
Mean at baseline for average person	−0.600	0.141	0.001
Effect at baseline of 1 SD increase in:			
Mem-D	0.418	0.070	0.001
Mem-B	0.352	0.079	0.001
Mem-R	0.287	0.038	0.001
Mean change/year for average person	−0.062	0.014	0.001
Effect on change/year of 1 SD increase in:			
Mem-D	−0.010	0.011	NS
Mem-B	0.067	0.016	0.001
Mem-R	0.049	0.011	0.001

Results show the expected difference in the executive function baseline score and annual change associated with a 1.0 unit difference in memory components. NS, not significant.

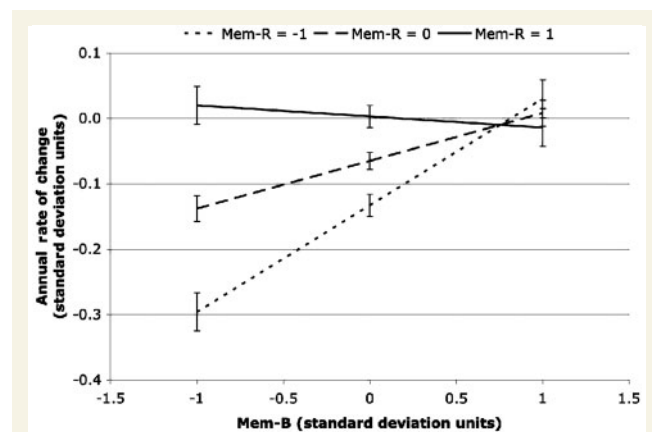


Figure 2 Relationship of baseline Mem-B to longitudinal change in executive function as a function of different levels of Mem-R. Mem-B scores reflect brain status as measured by structural MRI measures of brain matter, hippocampal volume and white matter hyperintensity. Positive Mem-R scores indicate that episodic memory performance was better than expected on the basis of demographic and MRI variables (high reserve), negative scores indicate worse than expected performance (low reserve). Mem-B was a stronger determinant of executive function change in individuals with lower Mem-R.

model derived linear relationships between Mem-B and executive function change for selected values of Mem-R, and illustrates how Mem-B is a more important determinant of executive function change in individuals with low Mem-R than in persons with high Mem-R.

A final model examined whether the interaction of Mem-B and Mem-R on executive function change would still be present after controlling for clinical diagnosis at baseline (normal, mild cognitive impairment, demented). Terms were added to the model to account for baseline diagnosis effects on executive function baseline and change. Mean estimated baseline executive function was significantly related to diagnosis, independent of the memory components and Mem-B by Mem-R interaction, but diagnosis was not independently related to change (not shown). The Mem-B by

Mem-R interaction effect on executive function was unchanged ($\beta = -0.093$, $SE = 0.018$, $P < 0.001$).

Secondary analyses were performed decomposing the SENAS semantic memory composite score into Sem-D, Sem-B and Sem-R components, using the same basic model applied to episodic memory in the primary analyses. The Sem-D and Sem-R components each accounted for 40–45% of the semantic memory variance, while Sem-B was smaller, accounting for about 10%. Results with respect to Hypotheses 1–5 were virtually identical in pattern and significance. However, the relationships of Sem-R with diagnosis, CDR and conversion to dementia, while statistically significant, were weaker than for Mem-R (not shown). The most critical test of whether the residual term has the effects predicted for a measure of reserve is Hypothesis 5 (reserve will modify the effect of brain atrophy on longitudinal change). For that reason, results from that analysis are presented in Fig. 3, which shows model-derived linear relationships between Sem-B and change in executive function for selected values of Sem-R. Executive function change was significantly related to Sem-B ($\beta = 0.055$, $SE = 0.013$, $P < 0.001$), Sem-R ($\beta = 0.027$, $SE = 0.012$, $P = 0.02$) and the interaction of these two variables ($\beta = -0.047$, $SE = 0.016$, $P < 0.002$). The Sem-R and Sem-R \times Sem-B interaction effects were weaker than for Mem-R and Mem-R \times Mem-B.

An additional secondary analysis used the CDR sum of boxes as the longitudinal outcome measure in the model for Hypothesis 5. Results showed that CDR change was associated with Mem-B ($\beta = -0.245$, $SE = 0.033$, $P < 0.001$), Mem-R ($\beta = -0.457$, $SE = 0.055$, $P = 0.02$) and the Mem-B \times Mem-R interaction ($\beta = 0.361$, $SE = 0.046$, $P < 0.001$). These results show that Mem-R moderates the effect of Mem-B on CDR sum of boxes; CDR scores are more strongly related to Mem-B in individuals with low Mem-R.

Discussion

The basic hypothesis of this investigation was that statistically decomposing scores on a test of episodic memory into

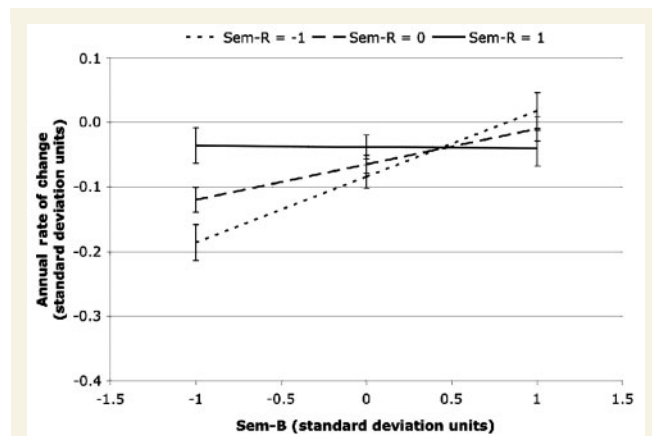


Figure 3 Relationship of baseline Sem-B to longitudinal change in executive function as a function of different levels of Sem-R. Sem-B scores reflect brain status as measured by structural MRI measures of brain matter, hippocampal volume and white matter hyperintensity. Positive Sem-R scores indicate that semantic memory performance was better than expected on the basis of demographic and MRI variables (high reserve), negative scores indicate worse than expected performance (low reserve). Sem-B was a stronger determinant of executive function change in individuals with lower Sem-R.

uncorrelated component scores could be used to create a useful measure of cognitive reserve. Baseline data from a diverse longitudinal ageing cohort were used to generate three component scores: Mem-D, the component score determined by demographic variables; Mem-B, the component score determined by basic brain volumes; and Mem-R, the component determined by all other factors and which we tested as a potential measure of current reserve. The analyses showed that the latter component, which accounted for the largest proportion of episodic memory performance, captured clinically meaningful person-specific deviations from predicted memory performance that correlate with a previously proposed indicator of reserve and that help to explain important longitudinal clinical outcomes consistent with the concept of cognitive reserve. Thus the results support treating the residual component as a substantive and useful, if imprecise and impure, measure of cognitive reserve.

Reserve is generally conceptualized as a capacity that is built through cognitively demanding and stimulating experiences, the chief example of which is education. Here, the residual component Mem-R was calculated in a way that made it statistically independent of the variable years of education. Word recognition ability correlates strongly with education but also adds important information about quality of education (Manly *et al.*, 2002, 2004) and is commonly used as an estimate of 'premorbid' cognitive ability (Grober and Sliwinski, 1991; Crawford *et al.*, 2001). Word recognition itself has been used as an indicator of cognitive reserve (Schmand *et al.*, 1997; Manly *et al.*, 2003; Richards *et al.*, 2004). However, education's effects on episodic memory are generally weak (e.g. 10% of variance) even when range of education in the study sample is wide (Mungas *et al.*, 2005b). Thus, the finding that the residual measure of reserve correlates with reading ability is not simply a matter of correlation between highly related

cognitive domains but can be reasonably interpreted as a correlation between different indicators of reserve. In contrast, Mem-B, determined by brain pathology, was not associated with reading measures.

The core of the reserve concept is that reserve modifies the impact of acquired pathology. The common age-related brain pathologies are cumulative and thus reserve ought to modify rates of change. For this reason, longitudinal tests of the hypothesis are especially critical. We performed three such tests. The reserve concept predicts that persons with higher reserve have a lower risk of incident dementia; here, the residual term was shown to modify the risk of converting from non-demented to demented such that those with higher residual values had a lower risk of converting. In addition, we found that the reserve/residual term modified rates of decline in executive function, such that higher Mem-R was associated with less decline in this important, second cognitive domain. Finally, the analyses show that brain status (Mem-B) had less of an effect on decline in executive function in individuals with high reserve (high Mem-R) than in subjects with low reserve (Fig. 2). The latter finding was replicated in an analysis that used global cognitive status, as measured by the CDR sum of boxes, as the outcome measure. Thus, in this series of analyses, the variable Mem-R showed the same pattern and direction of significant effects on a variety of clinically relevant outcomes as would be expected of a direct measure of reserve.

The fact that the residual component is rather strongly related to cross-sectional indicators of clinical status (diagnosis, CDR) is not surprising given that it is residual variance in episodic memory. However, the finding does illustrate an interesting property of this measure, which is that it is dynamic. Reserve is conceptualized as a capacity that is eroded by brain pathology; pathology tends to impair cognition and thus, people who are more impaired should average less reserve than people less impaired. The residual term behaves in this way.

While each of the *a priori* hypotheses was confirmed, one might question how strongly they support labelling the residual term 'reserve'. The residual term can be accurately described as a measure of episodic memory that has been statistically adjusted for the effect of MRI and demographics. This restatement makes it clear that some of the results are nearly a foregone conclusion. This is most true of Hypothesis 1; we hypothesized that reserve would correlate with diminished global cognitive status and it would be remarkable if an adjusted measure of episodic memory did not do so. Whilst Hypotheses 3 and 4, where the residual term will modify conversion and clinical course, are not self-fulfilling in this same way, the results are not surprising given that baseline episodic memory has previously been shown to predict conversion to dementia from mild cognitive impairment (Tabert *et al.*, 2006; Albert *et al.*, 2007; Dickerson *et al.*, 2007), even controlling for MRI measured brain volumes (Decarli *et al.*, 2004). If these results are unsurprising for an adjusted measure of episodic memory, why call it 'reserve'?

Parameterizing the unexplained variance in episodic memory facilitates asking what the effects of this unexplained variance are and challenges us to investigate its determinants. Those advantages are also some of the reasons that theoretical constructs are useful. The residual term has an almost literal, and perhaps

deceptively simple, correspondence to the concept of reserve. However, framing the residual term in the theoretical context facilitates investigation of reserve, a construct that is thought to be useful but which is difficult to measure and hence to study.

The results of the present analyses illustrate the advantages of mapping this variable to the construct of reserve. Hypotheses 3 and 4 (regarding conversion and rates of decline in executive ability) in a sense replicate prior reports that episodic memory is a predictor of decline. However, viewing the outcomes as an effect of reserve leads to additional predictions, such as Hypothesis 5. Studies that test reserve by analysing cognitive change in relation to measured pathology are rare, and have not always shown supportive results (Del Ser *et al.*, 1999). Thus, the finding that the residual modifies the impact of brain atrophy on cognition is a strong test of the reserve hypothesis and is an analysis that results from viewing the residual as a measure of reserve.

It should also be noted that each of the three component measures created in this model are derived from and correlated with episodic memory, yet these components have differential relationships with external variables that have theoretical, practical and prognostic significance. The demographic component was strongly related to concurrent cognitive variables (Reading and Executive Function), but was weakly or not related to cross-sectional clinical variables (clinical diagnosis, CDR Sum) or to longitudinal outcomes (conversion to dementia, change in executive function). The component related to brain structure, not surprisingly, was most strongly related to clinical status and longitudinal change. However, the residual component still had important relations with these outcomes. These differential patterns of statistical association suggest that the three should also be differentiated conceptually.

The Mem-R component undoubtedly represents factors other than reserve. While we explicitly modelled measurement error, the Mem-R component is influenced by systematic sources of variance that are unmeasured by the variables that were incorporated in these statistical models. For instance, many aspects of brain pathology are not captured by the volumetric MRI variables used in this study. These could include different structural measures (e.g. regional volumes, cortical thickness, hippocampal shape measures, alternative measures of white matter integrity), as well as measures of functional networks and molecular markers of pathology. It is also true that certain brain volumes may capture part of what reserve is. As such variables are identified and incorporated into models to explain cognitive function, one would expect that the pathology-related component would become larger and a more powerful determinant of clinical outcomes. However, it is unlikely that any set of pathology measures, no matter how inclusive and refined, will fully predict the changes in cognitive function that occur as a consequence of that pathology. That, at least, is the concept of reserve.

The finding that analyses using the residual term Sem-R, derived from a semantic memory measure, replicate but are weaker than those obtained with Mem-R makes two points. First, the findings are not highly sensitive to the particular cognitive measure that is decomposed to estimate reserve. Presumably, there is a degree of generality to cognitive reserve that is captured in a variety of cognitive measures. Secondly, there may also be a degree of

domain specificity to reserve and people may differ in the extent to which they have reserve with respect to different cognitive functions. For example, one could imagine that artistic and verbal abilities could be maintained to very different degrees in the face of dementia (Soricelli, 2006).

The demographic component, Mem-D was derived from models using only sex, ethnicity and education to predict baseline performance. Obviously other variables might be added, and as this component explains more variance, Mem-R would explain less. However, determining which variables to use as indicators of Mem-D requires careful thought. The point of Mem-R is to capture variance that modifies rates of change in cognitive function. If variables that determine reserve are incorporated into Mem-D the explanatory value of Mem-R is diminished.

We chose to use education as a demographic factor in Mem-D to separate education effects from Mem-R as a matter of proof of principle. Moreover, that Mem-D did not affect longitudinal change or clinical outcomes while Mem-R was associated with these outcomes is interesting with respect to the role of education in reserve. It may be that education contributes to reserve, but that those effects are entirely mediated by brain volumes. It could also be that certain aspects of education's effects on cognition raise a person's level of cognitive performance but do not protect against pathology-induced decline. It also invites the question of what factors, associated with education (e.g. lifetime cognitive activity), might relate to reserve. Finally, it suggests that reserve is more than a protective factor that simply provides higher baseline function.

Whether or not the residual episodic memory component would continue to behave as a measure of cognitive reserve as additional variables are identified and incorporated into the basic model to explain memory function is ultimately an empirical question. But the major importance of this study is not that it provides an indisputable measure of cognitive reserve, but rather, that it demonstrates an approach that can be used to measure cognitive reserve and to study how it interacts with brain disease and injury to determine extent of cognitive impairment. This approach offers an operational measure of reserve that is explicit, quantitative and individually specific. Such measures are convenient as dependent variables. This approach also permits measuring change in reserve over time, something that is nearly impossible when demographic covariates are used as markers of reserve. One of the appealing features of Mem-R is that once defined, it is relatively easy to test whether or not other ostensible determinants of reserve are related to it. While it is important to avoid being overly literal about the label applied to the residual component, it appears to provide an operational measure that is clearly defined, relatively simple to derive, and that can then be used to test hypotheses about numerous factors that potentially modify the impact of brain pathology on cognitive performance.

A notable strength of this study is the wide range of education, ethnicity, life experience and cognitive performance of the participants, suggesting that there is probably substantially more variation in cognitive reserve than in a typical research clinic population. The participants were well characterized by uniform ascertainment of clinical diagnosis and use of cognitive measures that have excellent psychometric properties across the full

spectrum of performance, education and ethnic backgrounds. Nonetheless, replication of our basic findings in independent data sets is needed. The methods used for analysis do not depend on a pre-specified set of proxy measures, but rather parameterize unexplained variance and then attempt to determine its potential impact independent of pathology and demographics. Future work, therefore, will attempt to identify personal or life-style characteristics associated with this residual variance and incorporate these findings into a more complete model to explain cognitive function thereby extending our knowledge of the role played by these factors.

Decomposing the variance in cognitive test scores appears to be a useful approach to investigating the complex determinants of cognitive function in older adults. Demographic variables have strong effects on cross-sectional cognitive test scores but are less important as determinants of longitudinal change. Diseases of ageing like Alzheimer's disease are strong determinants of changes in brain structure and consequently, volumetric brain measures have major effects on both cross sectional cognition and on rates of change. Even when combined, however, these classes of variables still leave considerable variance in the cognitive function of older adults unexplained. Explicitly defining a latent variable accounting for residual, or otherwise unexplained differences in episodic memory performance appears to be a useful strategy for investigating this unexplained variance. Because this variable behaves as a measure of reserve might be expected to, most critically because it modifies the effect of brain pathology on cognitive decline, we believe it is reasonable to treat this residual term as a measure of reserve.

Whatever the measure is named, however, is much less important than the empirical outcomes that may result from its use. In studying cognitive ageing the most useful constructs are those that offer new ways to explain variability in change. Reserve is often viewed as such a construct. However useful it may be, it is desirable to eventually dispense with it in large measure, and instead to replace it with a set of empirically based understandings of what specific factors—at what point in life, and by what mechanisms—increase the individual's capacity to mitigate the effects of brain disease. Parameterizing residual variance in episodic memory performance appears to be a useful strategy for investigating this unexplained variance, its determinants, and how it modifies the impact of disease on cognition.

Funding

National Institutes on Ageing (AG10129 and AG021028 to C.D., AG10220 to D.M. and AG031563 to B.R.); and by the California Department of Public Health Alzheimer's Disease Program Contracts 06-55311 (D.M.) and 6-55312 (B.R.). This material is in part the result of work supported with resources and the use of facilities at the Northern California VA Health Care System.

Supplementary material

Supplementary material is available at *Brain* online.

References

- Adak S, Illouz K, Gorman W, Tandon R, Zimmerman EA, Guariglia R, et al. Predicting the rate of cognitive decline in aging and early Alzheimer disease. *Neurology* 2004; 63: 108–14.
- Albert MS, Jones K, Savage CR, Berkman L, Seeman T, Blazer D, et al. Predictors of cognitive change in older persons: MacArthur studies of successful aging. *Psychol Aging* 1995; 10: 578–89.
- Albert M, Blacker D, Moss MB, Tanzi R, McArdle JJ. Longitudinal change in cognitive performance among individuals with mild cognitive impairment. *Neuropsychology* 2007; 21: 158–69.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-III-R. 3rd edn., rev. edn. Washington, DC: American Psychiatric Association; 1987.
- Andel R, Vigen C, Mack WJ, Clark LJ, Gatz M. The effect of education and occupational complexity on rate of cognitive decline in Alzheimer's patients. *J Int Neuropsychol Soc* 2006; 12: 147–52.
- Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. *Neurology* 2006; 66: 1837–44.
- Bentler PM. Comparative fit indices in structural models. *Psychological Bulletin* 1990; 107: 238–46.
- Bentler PM. EQS structural equations program manual. Encino, CA: Multivariate Software; 1995.
- Bollen K, Lennox R. Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin* 1991; 110: 305–14.
- Borrell LN, Dallo FJ, White K. Education and diabetes in a racially and ethnically diverse population. *Am J Public Health* 2006; 96: 1637–42.
- Brickman AM, Honig LS, Scarmeas N, Tatarina O, Sanders L, Albert MS, et al. Measuring cerebral atrophy and white matter hyperintensity burden to predict the rate of cognitive decline in Alzheimer disease. *Arch Neurol* 2008; 65: 1202–8.
- Browne M, Cudek R. Alternate ways of assessing model fit. In: Bollen K, Long J, editors. *Testing structural equation models*. Thousand Oaks, CA: Sage; 1993. p. 136–62.
- Buckner RL. Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. *Neuron* 2004; 44: 195–208.
- Cahn-Weiner DA, Boyle PA, Malloy PF. Tests of executive function predict instrumental activities of daily living in community-dwelling older individuals. *Appl Neuropsychol* 2002; 9: 187–91.
- Cahn-Weiner DA, Farias ST, Julian L, Harvey DJ, Kramer JH, Reed BR, et al. Cognitive and neuroimaging predictors of instrumental activities of daily living. *J Int Neuropsychol Soc* 2007; 13: 747–57.
- Canadian Study of Health and Aging Working Group. The Canadian Study of Health and Aging: risk factors for Alzheimer's disease in Canada. *Neurology* 1994; 44: 2073–80.
- Carter K, Hinton L, Reed BR, Beckett L, Lara E, DeCarli C, et al. Recruitment of a community-based cohort for research on diversity and risk of dementia. Alzheimer's disease and related disorders. In press.
- Chui HC, Zarow C, Mack WJ, Ellis WG, Zheng L, Jagust WJ, et al. Cognitive impact of subcortical vascular and Alzheimer's disease pathology. *Ann Neurol* 2006; 60: 677–87.
- Cobb JL, Wolf PA, Au R, White R, D'Agostino RB. The effect of education on the incidence of dementia and Alzheimer's disease in the Framingham Study. *Neurology* 1995; 45: 1707–12.
- Cosentino S, Manly J, Mungas D. Do reading tests measure the same construct in multiethnic and multilingual older persons? *J Int Neuropsychol Soc* 2007; 13: 228–36.
- Crawford JR, Deary IJ, Starr J, Whalley LJ. The NART as an index of prior intellectual functioning: a retrospective validity study covering a 66-year interval. *Psychol Med* 2001; 31: 451–8.
- Davis DG, Schmitt FA, Wekstein DR, Markesbery WR. Alzheimer neuropathologic alterations in aged cognitively normal subjects. *J Neuropathol Exp Neurol* 1999; 58: 376–88.

- Deaton A, Paxson C. Mortality, education, income and inequality among American cohorts. In: Wise DA, editor. *Themes in the Economics of Aging*. Chicago, IL: University of Chicago Press; 1999. p. 129–170.
- DeCarli C, Grady CL, Clark CM, Katz DA, Brady DR, Murphy DGM, et al. Comparison of positron emission tomography, cognition, and brain volume in Alzheimer's disease with and without severe abnormalities of white matter. *J Neurol Neurosurg Psychiatry* 1996; 60: 158–67.
- DeCarli C, Maisog J, Murphy DG, Teichberg D, Rapoport SI, Horwitz B. Method for quantification of brain, ventricular, and subarachnoid CSF volumes from MR images. *J Comput Assist Tomogr* 1992; 16: 274–84.
- DeCarli C, Massaro J, Harvey D, Hald J, Tullberg M, Au R, et al. Measures of brain morphology and infarction in the framingham heart study: establishing what is normal. *Neurobiol Aging* 2005; 26: 491–510.
- DeCarli C, Miller BL, Swan GE, Reed T, Wolf PA, Garner J, et al. Predictors of brain morphology for the men of the NHLBI twin study. *Stroke* 1999; 30: 529–36.
- DeCarli C, Mungas D, Harvey D, Reed B, Weiner M, Chui H, et al. Memory impairment, but not cerebrovascular disease, predicts progression of MCI to dementia. *Neurology* 2004; 63: 220–7.
- DeCarli C, Murphy DGM, Tranh M, Grady CL, Haxby JV, Gillette JA, et al. The effect of white matter hyperintensity volume on brain structure, cognitive performance, and cerebral metabolism of glucose in 51 healthy adults. *Neurology* 1995; 45: 2077–84.
- DeCarli C, Reed BR, Jagust W, Martinez O, Ortega M, Mungas D. Brain behavior relationships among African Americans, whites, and Hispanics. *Alzheimer Dis Assoc Disord* 2008; 22: 382–91.
- Del Ser T, Gonzalez-Montalvo JL, Martinez-Espinosa S, Delgado-Villalobos C, Bermejo F. Estimation of premorbid intelligence in Spanish people with the Word Accentuation Test and its application to the diagnosis of dementia. *Brain Cogn* 1997; 33: 343–56.
- Del Ser T, Hachinski V, Merskey H, Munoz DG. An autopsy-verified study of the effect of education on degenerative dementia. *Brain* 1999; 122: 2309–19.
- Dickerson BC, Sperling RA, Hyman BT, Albert MS, Blacker D. Clinical prediction of Alzheimer disease dementia across the spectrum of mild cognitive impairment. *Arch Gen Psychiatry* 2007; 64: 1443–50.
- Driscoll I, Resnick SM, Troncoso JC, An Y, O'Brien R, Zonderman AB. Impact of Alzheimer's pathology on cognitive trajectories in nondemented elderly. *Ann Neurol* 2006; 60: 688–95.
- Edwards JR, Bagozzi RP. On the nature and direction of relationships between constructs and measures. *Psychol Methods* 2000; 5: 155–74.
- Fitzpatrick AL, Kuller LH, Ives DG, Lopez OL, Jagust W, Breitner JC, et al. Incidence and prevalence of dementia in the Cardiovascular Health Study. *J Am Geriatr Soc* 2004; 52: 195–204.
- Gonzalez HM, Mungas D, Haan MN. A verbal learning and memory test for English- and Spanish-speaking older Mexican-American adults. *Clin Neuropsychol* 2002; 16: 439–51.
- Gonzalez HM, Mungas D, Reed BR, Marshall S, Haan MN. A new verbal learning and memory test for English- and Spanish-speaking older people. *J Int Neuropsychol Soc* 2001; 7: 544–55.
- Grober E, Sliwinski M. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *J Clin Exp Neuropsychol* 1991; 13: 933–50.
- Grossman M, Kaestner R. *Effects of education on health*. Ann Arbor, MI: University of Michigan Press; 1997.
- Jack CR Jr, Dickson DW, Parisi JE, Xu YC, Cha RH, O'Brien PC, et al. Antemortem MRI findings correlate with hippocampal neuropathology in typical aging and dementia. *Neurology* 2002; 58: 750–7.
- Jack CR Jr, O'Brien PC, Retzlman DW, Shiung MM, Xu Y, Muthupillai R, et al. FLAIR histogram segmentation for measurement of leukoaraiosis volume. *J Magn Reson Imaging* 2001; 14: 668–76.
- Jack CR Jr, Petersen RC, Xu YC, Waring SC, O'Brien PC, Tangalos EG, et al. Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology* 1997; 49: 786–94.
- Jagust WJ, Zheng L, Harvey DJ, Mack WJ, Vinters HV, Weiner MW, et al. Neuropathological basis of magnetic resonance images in aging and dementia. *Ann Neurol* 2008; 63: 72–80.
- Jellinger KA. Understanding the pathology of vascular cognitive impairment. *J Neurol Sci* 2005; 229–230: 57–63.
- Kaplan G, Haan M, Syme S, Minkler M, Winkelby M. Closing the gap: the burden of unnecessary illness. In: Amler R, Dull H, editors. *Socioeconomic status and health*. New York: Oxford University Press; 1987. p. 125–9.
- Kaplan GA, Keil JE. Socioeconomic factors and cardiovascular disease: a review of the literature. *Circulation* 1993; 88(4 Pt 1): 1973–98.
- Karp A, Karelhult I, Qiu C, Bellander T, Winblad B, Fratiglioni L. Relation of education and occupation-based socioeconomic status to incident Alzheimer's disease. *Am J Epidemiol* 2004; 159: 175–83.
- Katzman R. Education and the prevalence of dementia and Alzheimer's disease. *Neurology* 1993; 43: 13–20.
- Kramer JH, Mungas D, Reed BR, Wetzel ME, Burnett MM, Miller BL, et al. Longitudinal MRI and cognitive change in healthy elderly. *Neuropsychology* 2007; 21: 412–8.
- Liao D, Cooper L, Cai J, Toole J, Bryan N, Burke G, et al. The prevalence and severity of white matter lesions, their relationship with age, ethnicity, gender, and cardiovascular disease risk factors: the ARIC Study. *Neuroepidemiology* 1997; 16: 149–62.
- Lim A, Tsuang D, Kukull W, Nochlin D, Leverenz J, McCormick W, et al. Clinico-neuropathological correlation of Alzheimer's disease in a community-based case series. *J Am Geriatr Soc* 1999; 47: 564–9.
- Liu K, Cedres LB, Stamler J, Dyer A, Stamler R, Nanas S, et al. Relationship of education to major risk factors and death from coronary heart disease, cardiovascular diseases and all causes, findings of three Chicago epidemiologic studies. *Circulation* 1982; 66: 1308–14.
- Longstreth WT Jr, Bernick C, Manolio TA, Bryan N, Jungreis CA, Price TR. Lacunar infarcts defined by magnetic resonance imaging of 3660 elderly people: the Cardiovascular Health Study. *Arch Neurol* 1998; 55: 1217–25.
- McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. *Ann Rev Psychol* 2009; 60: 577–605.
- Manly JJ, Byrd DA, Touradji P, Stern Y. Acculturation, reading level, and neuropsychological test performance among African American elders. *Appl Neuropsychol* 2004; 11: 37–46.
- Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. *J Int Neuropsychol Soc* 2002; 8: 341–8.
- Manly JJ, Schupf N, Tang MX, Stern Y. Cognitive decline and literacy among ethnically diverse elders. *J Geriatr Psychiatry Neurol* 2005; 18: 213–7.
- Manly JJ, Touradji P, Tang M-X, Stern Y. Literacy and Memory Decline Among Ethnically Diverse Elders. *J Clin Exp Neuropsychol* 2003; 25: 680–90.
- Meredith W, Tisak J. Latent curve analysis. *Psychometrika* 1990; 55: 107–22.
- Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology* 1993; 43: 2412–14.
- Mungas D, Harvey D, Reed BR, Jagust WJ, DeCarli C, Beckett L, et al. Longitudinal volumetric MRI change and rate of cognitive decline. *Neurology* 2005a; 65: 565–71.
- Mungas D, Jagust WJ, Reed BR, Kramer JH, Weiner MW, Schuff N, et al. MRI predictors of cognition in subcortical ischemic vascular disease and Alzheimer's disease. *Neurology* 2001; 57: 2229–35.
- Mungas D, Reed BR, Crane PK, Haan MN, Gonzalez H. Spanish and English Neuropsychological Assessment Scales (SENAS): further development and psychometric characteristics. *Psychol Assess* 2004; 16: 347–59.
- Mungas D, Reed BR, Farias ST, DeCarli C. Age and education effects on relationships of cognitive test scores with brain structure in demographically diverse older persons. *Psychol Aging* 2009; 24: 116–28.

- Mungas D, Reed BR, Haan MN, Gonzalez H. Spanish and English neuropsychological assessment scales: relationship to demographics, language, cognition, and independent function. *Neuropsychology* 2005b; 19: 466–75.
- Mungas D, Reed BR, Jagust WJ, DeCarli C, Mack WJ, Kramer JH, et al. Volumetric MRI predicts rate of cognitive decline related to AD and cerebrovascular disease. *Neurology* 2002; 59: 867–73.
- Mungas D, Reed BR, Marshall SC, Gonzalez HM. Development of psychometrically matched English and Spanish language neuropsychological tests for older persons. *Neuropsychology* 2000; 14: 209–23.
- Mungas D, Reed BR, Tomaszewski Farias S, DeCarli C. Criterion-referenced validity of a neuropsychological test battery: equivalent performance in elderly Hispanics and non-Hispanic Whites. *J Int Neuropsychol Soc* 2005c; 11: 620–30.
- Muthén LK, Muthén BO. *Mplus User's Guide*. Fifth Edition. Los Angeles, CA: Muthén & Muthén; 2007.
- Ott A, Breteler MMB, van Harskamp F, Claus JJ, van der Cammen TJM, Grobbee DE, et al. Prevalence of Alzheimer's disease and vascular dementia: association with education. The Rotterdam study. *British Med J* 1995; 310: 970–3.
- Petersen RC. Mild cognitive impairment as a diagnostic entity. *J Intern Med* 2004; 256: 183–94.
- Potter GG, Helms MJ, Plassman BL. Associations of job demands and intelligence with cognitive performance among men in late life. *Neurology* 2008; 70(19 Pt 2): 1803–8.
- Richards M, Shipley B, Fuhrer R, Wadsworth ME. Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study. *British Med J* 2004; 328(7439): 552.
- Royall DR, Cabello M, Polk MJ. Executive dyscontrol: an important factor affecting the level of care received by older retirees. *J Am Geriatr Soc* 1998; 46: 1519–24.
- Royall DR, Palmer R, Chiodo LK, Polk MJ. Declining executive control in normal aging predicts change in functional status: the Freedom House Study. *J Am Geriatr Soc* 2004; 52: 346–52.
- Rubin EH, Storandt M, Miller JP, Kinschler DA, Grant EA, Morris JC, et al. A prospective study of cognitive function and onset of dementia in cognitively healthy elders. *Arch Neurol* 1998; 55: 395–401.
- Salthouse TA. Memory aging from 18 to 80. *Alzheimer Dis Assoc Disord*. 2003; 17: 162–7.
- Satz P. Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology* 1993; 7: 273–95.
- Schmand B, Smit JH, Geerlings MI, Lindeboom J. The effects of intelligence and education on the development of dementia. A test of the brain reserve hypothesis. *Psychol Med*. 1997; 27: 1337–44.
- Schmitt FA, Davis DG, Wekstein DR, Smith CD, Ashford JW, Markesbery WR. "Preclinical" AD revisited: neuropathology of cognitively normal older adults. *Neurology* 2000; 55: 370–6.
- Soricelli RL. Medicine and the arts. A series of self portraits by William Utermohlen. *Commentary. Acad Med* 2006; 81: 996–7.
- Stern Y. What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc* 2002; 8: 448–60.
- Stern Y. Cognitive reserve. *Neuropsychologia* 2009; 47: 2015–28.
- Tabert MH, Manly JJ, Liu X, Pelton GH, Rosenblum S, Jacobs M, et al. Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Arch Gen Psychiatry* 2006; 63: 916–24.
- Tomaszewski Farias S, Cahn-Weiner DA, Harvey DJ, Reed BR, Mungas D, Kramer JH, et al. Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *Clin Neuropsychol* 2009; 23: 446–61.
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika* 1973; 38: 1–10.
- Tullberg M, Fletcher E, DeCarli C, Mungas D, Reed BR, Harvey DJ, et al. White matter lesions impair frontal lobe function regardless of their location. *Neurology* 2004; 63: 246–53.
- Verghese J, Lipton RB, Katz MJ, Hall CB, Derby CA, Kuslansky G, et al. Leisure activities and the risk of dementia in the elderly. *N Engl J Med*. 2003; 348: 2508–16.
- White L, Small BJ, Petrovitch H, Ross GW, Masaki K, Abbott RD, et al. Recent clinical-pathologic research on the causes of dementia in late life: update from the Honolulu-Asia Aging Study. *J Geriatr Psychiatry Neurol* 2005; 18: 224–7.
- Whitwell JL, Przybelski SA, Weigand SD, Knopman DS, Boeve BF, Petersen RC, et al. 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain* 2007; 130(Pt 7): 1777–86.
- Wilson RS, Beckett LA, Barnes LL, Schneider JA, Bach J, Evans DA, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002a; 17: 179–93.
- Wilson RS, Mendes De Leon CF, Barnes LL, Schneider JA, Bienias JL, Evans DA, et al. Participation in cognitively stimulating activities and risk of incident Alzheimer disease. *JAMA* 2002b; 287: 742–8.
- Wong MD, Shapiro MF, Boscardin WJ, Ettner SL. Contribution of major diseases to disparities in mortality. *N Engl J Med* 2002; 347: 1585–92.
- Yoshita M, Fletcher E, Harvey D, Ortega M, Martinez O, Mungas DM, et al. Extent and distribution of white matter hyperintensities in normal aging, MCI, and AD. *Neurology* 2006; 67: 2192–8.
- Zarow C, Vinters HV, Ellis WG, Weiner MW, Mungas D, White L, et al. Correlates of hippocampal neuron number in Alzheimer's disease and ischemic vascular dementia. *Ann Neurol* 2005; 57: 896–903.
- Zelinski EM, Gilewski MJ, Schaie KW. Individual differences in cross-sectional and 3-year longitudinal memory performance across the adult life span. *Psychol Aging* 1993; 8: 176–86.
- Zhang MY, Katzman R, Salmon D, Jin H, Cai GJ, Wang ZY, et al. The prevalence of dementia and Alzheimer's disease in Shanghai, China: impact of age, gender, and education. *Ann Neurol* 1990; 27: 428–37.