Published in final edited form as:

J Int Neuropsychol Soc. 2012 November ; 18(6): 1071–1080. doi:10.1017/S1355617712000859.

Neuropsychological Test Performance and Cognitive Reserve in Healthy Aging and the Alzheimer's Disease Spectrum: A Theoretically-Driven Factor Analysis

Meghan B. Mitchell^{1,2,3}, Lynn W. Shaughnessy^{3,4,5}, Steven D. Shirk^{1,2}, Frances M. Yang^{3,6,7}, and Alireza Atri^{1,2,3}

¹Geriatric Research Education and Clinical Center, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, Massachusetts

²Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts

³Harvard Medical School, Boston, Massachusetts

⁴Department of Psychiatry, Massachusetts Mental Health Center, Boston, Massachusetts

⁵Department of Psychiatry, Beth Israel Deaconess Medical Center, Boston, Massachusetts

⁶Institute for Aging Research, Hebrew Senior Life, Boston, Massachusetts

⁷Department of Internal Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts

Abstract

Accurate measurement of cognitive function is critical for understanding the disease course of Alzheimer's disease (AD). Detecting cognitive change over time can be confounded by level of premorbid intellectual function or cognitive reserve and lead to under or over diagnosis of cognitive impairment and AD. Statistical models of cognitive performance that include cognitive reserve can improve sensitivity to change and clinical efficacy. We used confirmatory factor analysis to test a four-factor model comprised of memory/language, processing speed/executive function, attention, and cognitive reserve factors in a group of cognitively healthy older adults and a group of participants along the spectrum of amnesic mild cognitive impairment to AD (aMCI-AD). The model showed excellent fit for the control group ($\chi^2 = 100$, $df = 78$, CFI = .962, RMSEA = .049) and adequate fit for the aMCI-AD group ($\chi^2 = 1750$, $df = 78$, CFI = .932, RMSEA = .085). Though strict invariance criteria were not met, invariance testing to determine if factor structures are similar across groups yielded acceptable absolute model fits and provide evidence in support of configural, metric, and scalar invariance. These results provide further support for the construct validity of cognitive reserve in healthy and memory impaired older adults.

Keywords

mild cognitive impairment; brain reserve; executive function; memory function; dementia; cognition

Corresponding Author: Meghan Mitchell, 200 Springs Road, 182b, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, Massachusetts 01730, Phone: (781) 687-3650, Fax: (781) 687-2592, Meghan.Mitchell2@va.gov.

The authors have no conflicts of interest with the present study.

Introduction

Alzheimer's disease (AD) is the most common form of dementia with an estimated 5.4 million cases in the United States (Alzheimer's Association, 2011). Measurement of cognitive deficits in AD is crucial for early interventions, which may delay clinical decline (Atri, Rountree, Lopez, & Doody, 2012; DeKosky, 2003). Existing models of cognition in normal aging and AD have identified latent variables representing cognitive domains (Johnson, Storandt, Morris, Langford, & Galvin, 2008). A benefit of such latent variable approaches is that it potentially allows for purer measures of cognition (i.e., latent variables as opposed to a summation of individual indicator variables) that take into account covariance with other latent variables, and can thus provide comprehensive models of cognition with demonstrated convergent and discriminant validity (Satz, Cole, Hardy, & Rassovsky, 2010). Latent variable analysis also simplifies statistical models by reducing the number of indicator variables to their latent constructs, thereby reducing problems with multicollinearity and restrictions of statistical power. A further obstacle to measuring cognition and detecting change over time in AD is the problem of appropriately considering baseline level of premorbid intellectual function or cognitive reserve; if unaccounted for, these may lead to under or over diagnosis of AD and hamper the detection of change or efficacy in clinical trials (Rentz, et al., 2004). There is thus a need for statistical models of cognition that take into account cognitive reserve.

Cognitive reserve is typically defined as the brain's capacity to maintain cognitive function in spite of neurologic damage or disease (Stern, 2009). Thus, an ideal measure of cognitive reserve would quantify the discrepancy between cognitive functioning and the extent of brain disease or damage. While recent advances in neuroimaging techniques enable the quantification of AD-related pathology or correlates of degeneration such as measurement of amyloid load via PET imaging (Roe, et al., 2010) and atrophy via MRI cortical thickness measurement (Dickerson, et al., 2009), these techniques are not widely available, and do not capture all aspects of AD neuropathology (e.g., neurofibrillary tangles). In light of these difficulties with quantifying cognitive reserve, operational definitions typically use "proxy" measures that reflect lifetime experiences in cognitively stimulating activities. Cognitive reserve is typically associated with higher educational and occupational attainment (Stern, et al., 1994), and more involvement in physical, social, and leisure activities (Wilson, Scherr, Schneider, Tang, & Bennett, 2007). Individuals with higher cognitive reserve show more resilience to multiple forms of neurologic insult, slower rate of decline in normal cognitive aging, and delayed onset of decline in AD (Stern, 2009). Interventions aimed at increasing the level of involvement in activities associated with cognitive reserve (e.g., increasing socialization, physical, or cognitive activity), suggest that while some aspects of cognitive reserve may be genetically predetermined (Lee, 2003), other aspects may be modifiable (Wilson et al., 2007).

Cognitive reserve is also associated with higher levels of executive functioning (Siedlecki, et al., 2009), leading some to question cognitive reserve's uniqueness as a theoretical construct. Siedlecki and colleagues (2009) investigated the construct validity of cognitive reserve in a series of latent variable models comparing the factor loadings of neuropsychological tests and cognitive reserve variables on the constructs of memory, processing speed, executive function, and cognitive reserve, and found reasonably good indices of model fit across three samples of cognitively normal adults. There were, however, less consistent findings across samples with regard to the intercorrelations between the cognitive reserve and executive functioning indicator variables and constructs, with one out of the three samples showing a strong intercorrelation between cognitive reserve and executive functioning constructs ($r = .90$), and two of three samples showing cognitive reserve variables to have strong factor loadings not only on the cognitive reserve construct,

but also on the executive functioning construct (factor loadings ranging from .54–.93). The authors concluded that their results support convergent validity of the cognitive reserve construct, as their indicator variables of cognitive reserve consistently loaded on the cognitive reserve factor. However, given the divergence in their findings, discriminant validity was not entirely supported and the authors suggested the possibility that executive functioning and cognitive reserve are highly related constructs that may not be distinct.

In a critical review of the literature on the construct validity of cognitive reserve, Satz and colleagues (2011) suggest that construct validity be assessed with factor analysis to determine if indicator variables commonly used as proxies for cognitive reserve (e.g., years of education, premorbid IQ, engagement in cognitively challenging activities) load onto the same or separate factors as other constructs previously suggested to be potentially overlapping (Satz, Cole, Hardy, & Rassovsky, 2011). The authors propose a hypothetical model that includes four subcomponents: 1) executive function (e.g., measures of response inhibition, fluency, error monitoring, selective attention, cognitive switching, reasoning); 2) processing resources (e.g., measures of divided attention, processing speed, and working memory); 3) complex mental activity (e.g., engagement in cognitively stimulating activities, social networks, education, literacy, occupational history; and 4) intelligence or *g*, indicated by measures of fluid and crystallized intelligence (Satz, et al., 2011). It is unclear to us from their discussion of this proposed model if their assertion is that executive functioning, processing resources, mental activity, and general intellectual function are subcomponents of cognitive reserve, or if their model is intended to define cognitive reserve as complex mental activities, and test cognitive reserve's validity in a model with these cognitive constructs. Nonetheless, their hypothetical model underscores the importance of evaluating overlap in the constructs of cognitive reserve and the related constructs of executive function, processing resources, and intellectual functioning.

The aim of our study was to establish a latent variable model of neuropsychological test performance that incorporates cognitive reserve and may later serve as a basis to study longitudinal change in normal aging and along the spectrum of aMCI to AD (aMCI-AD; Locascio & Atri, 2011). We sought to examine the factor structure of cognitive reserve variables in combination with neuropsychological measures to evaluate cognitive reserve as a distinct construct, and to validate a model of cognitive domains and cognitive reserve, the CDCR model, in a group of cognitively normal older adults, and a group with memory impairment along the aMCI-AD spectrum. While we were unable to fully evaluate all possible constructs that potentially overlap with the construct of cognitive reserve, we hypothesized that indicators of cognitive reserve would load onto a different factor than indicators of processing resources and executive function, consistent with both Siedlecki et al. (2009) and Satz et al.'s (2011) hypotheses.

Method

Participants

A total of 559 participants from the Massachusetts Alzheimer's Disease Research Center (MADRC) longitudinal cohort study of memory and aging were administered the standard battery of tests implemented by the Alzheimer Disease Center (ADC) Uniform Data Set (UDS; Beekly, et al., 2007; Morris, et al., 2006). The MADRC implemented the UDS in September 2005, and initial visits ranging from September 2005 to August 2009 were included. Inclusion criteria were: 1.) Age $\geq$ 50 years, and 2) UDS clinical research diagnosis of cognitively normal with Clinical Dementia Rating score (CDR) equal to 0 (Morris, 1993), amnesic Mild Cognitive Impairment (aMCI) with CDR=0.5, or Possible or Probable Alzheimer's disease (PrAD) with CDR $\geq$ 0.5. Participants were then assigned to two groups according to their CDR: a control group comprised of clinically cognitively normal older

adult participants ($n = 294$) with $CDR=0$, and an aMCI-AD group comprised of participants with cognitive impairment across the AD-spectrum ($n = 265$) ranging from aMCI to PrAD and with $CDR = 0.5$.

Procedure

Participants underwent standardized UDS clinical research evaluations and were accompanied by a reliable informant. The cognitive test battery was administered by a trained psychometrician and consisted of the standard UDS neuropsychological battery and supplemental tests (see next paragraph) administered after completion of the UDS tests. Informed consent was obtained for all participants according to an approved protocol by the Partners Healthcare Human Subjects Research Institutional Review Board.

Measures

Measures from the UDS neuropsychological battery included the Wechsler Memory Scale-Revised (WMS-R) subtests Logical Memory IA and IIA (LMI and LMII), Digit Span Forward and Backward (Digit Fwd and Digit Bkwd; Wechsler, 1987), Semantic Fluency (Animals and Vegetables: SEMFLU; Morris, et al., 1989), Boston Naming Test (BNT, 30 item – odd numbered; Mack, Freed, Williams, & Henderson, 1992), Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Symbol Coding subtest (Dig Sym; Wechsler, 1987), and Trail Making Test Parts A and B (Trails A and Trails B; Reitan & Wolfson, 1985). Additional measures were the American National Adult Reading Test (AMNART; Grober & Sliwinski, 1991), and the Free and Cued Selective Reminding Test (FCSRT; Grober, Lipton, Hall, & Crystal, 2000). The AMNART, a single word reading test, correlates with overall intellectual functioning, and is used as an estimate of premorbid functioning, as it is often relatively preserved despite cognitive decline (Grober & Sliwinski, 1991). The FCSRT examines explicit memory using both free and cued recall over three trials, following a study phase of 16 pictures, and summarized by the total number of freely recalled items (FCSRT Free) and the total number of recalled items in both free and cued conditions (FCSRT Total; maximum performance = 48; Grober, Sanders, Hall, & Lipton, 2010).

Statistical Analysis

Under a latent variable modeling (LVM) framework, confirmatory factor analysis (CFA) using Blom-transformed scores for each indicator variable was conducted. Blom transformations are normative percentile ranks that are based on the normal distribution of responses for each score (Blom, 1958). An analytic approach grounded in modern measurement theory (Embretson & Reise, 2000) and item response theory (IRT)-based structural equation modeling, a specific type of LVM, (Gallo, Anthony, & Muthén, 1994; Muthén, 1989) was employed. This method was originally presented as a model of unobserved quantities measured by formative causes and reflective indicators (Hauser & Goldberger, 1971).

Our general approach was to test a series of theoretically-based models using Mplus version 6.0 (Muthén & Muthén, 2010) and following standard practice in the factor analytic field, we modified our model based on model diagnostics (modification indices; Hayden, et al., 2011). Specifically, our initial model separated memory and language into two separate constructs, but we found that the intercorrelations between indicator variables across these two theoretical constructs were high, leading to cross-factor loadings. We thus hypothesized a four-factor CDCR model, which included memory/language, attention, processing speed/executive function, and cognitive reserve factors.

We first conducted CFA on each group separately to test for model fit and examine factor loadings in the two samples. We next conducted a series of two-group CFAs in order to provide an assessment of the stability of our model across diagnostic group (factorial invariance) by systematically increasing the constraints on the model in order to test for four increasingly rigorous types of invariance: configural, metric, scalar, and strict invariance (Hayden, et al., 2011). Configural invariance is met when indicator variables load on the same factors across groups. Metric invariance is met with adequate model fit when factor loadings are held constant across groups. Scalar invariance is met when model fit is adequate when both factor loadings and intercepts are held constant across group. Strict invariance is met when model fit is adequate when factor loadings, intercepts, and residual variances are all constrained to be equal across groups. Model fit was assessed with the root mean square error of approximation (RMSEA) (Browne & Cudek, 1993; Muthén, Khoo, & Francis, 1998) and the comparative fit index (CFI) (Bentler & Chou, 1988; Muthén, 1998). The RMSEA provides a measure of discrepancy per model degree of freedom and approaches zero as fit improves. For the CFI, values greater than 0.90 generally indicate adequate fit (Bentler, 1990; Muthén, 1989). Browne and Cudek (1993) recommended rejecting models with RMSEA values greater than 0.1; Hu and Bentler (1998) suggested that values close to 0.06 or less indicate adequate model fit. In addition to evaluating absolute model fit, we tested for relative fit using the Satorra-Bentler χ^2 test, which provides a statistical comparison of relative fit of nested models with increasing levels of constraints from the respective, less constrained model (Satorra & Bentler, 2001), as well as examining change in CFI following Cheung and Rensvold's (2002) general guideline that a change in CFI of more than 0.01 may indicate worse model fit.

Results

Participant demographics and characteristics were similar to the larger NACC UDS data set (Weintraub et al., 2009). Consistent with typical group comparisons of older adults with normal cognition versus those with memory impairment, the aMCI-AD group tended to be older (mean age for aMCI-AD = 76, controls = 70, $p < 0.001$) but level of education was not significantly different between groups (mean education for aMCI-AD = 15.8, controls = 16.4, $p > 0.05$). The control group recalled, on average, 13 units from Story A on LMII, while the aMCI-AD group recalled an average of 6 units. Descriptive statistics further summarizing demographic, neuropsychological, and cognitive reserve variables are provided in Table 1.

Confirmatory Factor Analysis

As depicted in Figure 1, we tested a hypothesized CDCR model with a latent variable structure in which measures of memory and language (i.e., measures associated with semantic processing: LMI, LMII, FCSRT Free, FCSRT Total, BNT, and SEMFLU) loaded on one factor, measures of attention and verbal working memory loaded on a second factor (Digit Fwd and Digit Bkwd), measures of processing speed and executive function loaded on a third factor (Trails A, Trails B, and Dig Sym), and proxy measures of cognitive reserve loaded on a fourth factor (AMNART and Education). We did not control for demographic covariates such as age or sex in our latent variable models.

Allowing for residual correlations between LMI and LMII, and between FCSRT summary scores, the hypothesized four-factor CDCR model demonstrated reasonable fit across both samples. The CDCR model showed excellent fit for the control group ($\chi^2 = 100.12$, $df = 78$, CFI = .962, RMSEA = .049) and approached good fit for the aMCI-AD group ($\chi^2 = 1750.02$, $df = 78$, CFI = .932, RMSEA = .085).

In the control group, the memory and language factor indicator variables varied in the strength of their factor loadings, with some indicators showing relatively small factor loadings (LMI, LMII, and FCSRT Total loadings ranging from .21–.38) and the remainder indicator variables showing relatively stronger loadings. In contrast to the variability in memory and language measure factor loadings, the factor loadings for all indicator variables in the attention, processing speed/executive function, and cognitive reserve factors were consistently strong, ranging from .67–.81. We additionally modeled the correlations between latent variables in order to examine the shared variance between hypothetically distinct constructs. Based on Cohen's interpretation of correlation magnitudes, generally a correlation greater than 0.5 is large, 0.3–0.5 is medium, 0.1–0.3 is small, and less than 0.1 is trivial (Cohen, 1988). The memory and language factor showed large correlations with all other factors, with correlations ranging from $r = .53$ for the cognitive reserve factor, $r = .61$ for the attention factor, and $r = .71$ for the processing speed/executive function factor. The remainder of the inter-factor correlations were in the small-to-medium range, and the cognitive reserve factor was distinct from the processing speed/executive function factor ($r = .43$). Figure 1 depicts the factor structure, factor loadings, and intercorrelations between factors in the control group.

In the aMCI-AD group, there was more consistency in the strength of indicator variables' factor loadings on all latent variables, with the indicator variables for the memory and language factor demonstrating loadings ranging from .69–.88, attention factor indicators ranging from .76–.87, processing speed/executive function factor indicators ranging from .80–.86, and cognitive reserve factor indicators ranging from .64–.73. The inter-factor correlations between latent variables in the aMCI-AD group showed a similar pattern to that of the control group, with the memory and language factor showing large correlations with the processing speed/executive function factor ($r = .70$) and attention factor ($r = .56$). The cognitive reserve factor showed medium-sized correlations with processing speed/executive function ($r = .43$) and the memory and language ($r = .39$). A full summary of factor structure and loadings in the aMCI-AD group is portrayed in Figure 2.

We next conducted a series of two-group factor analyses across four levels of invariance testing by first evaluating absolute model fit at each level of invariance testing, and then comparing relative fit of each nested model using the Satorra-Bentler scaled χ^2 test to determine if each successive level of invariance testing resulted in model fit that was significantly different from the previous, less restricted model (Satorra & Bentler, 2001) as well as change in CFI (Δ CFI) of 0.01 or greater as suggested by Cheung and Rensvold (2002). Using the criteria of an RMSEA of 0.06 or less or CFI of 0.90 or more to evaluate absolute model fit, we found the hypothesized four-factor CDCR model to demonstrate reasonable fit across both samples assuming configural invariance (RMSEA = 0.067; CFI = 0.944; $\chi^2 = 262.51$, $df = 117$). The configural invariance model produced estimated factor scores in the control group for all factors that were approximately 0.000 with *SDs* ranging from 0.807–0.912, all of which were not significantly different from those for the aMCI-AD group (all $M = 0.000$, *SDs* ranging from 0.838–0.939; all $p > 0.05$).

At each level of invariance testing, relative fit deteriorated significantly (all S-B $\chi^2 p < 0.001$, Δ CFI > 0.01), but absolute model fit remained acceptable assuming metric invariance (RMSEA = 0.076; CFI = 0.923; $\chi^2 = 342.343$, $df = 130$), which again produced estimated factor scores that were not significantly different by group (Control group memory/language $M = 0.000$, $SD = 0.722$; attention $M = 0.000$, $SD = 0.761$; processing speed/executive function $M = 0.000$, $SD = 0.845$; and cognitive reserve $M = 0.000$, $SD = 0.778$; aMCI-AD group memory/language $M = 0.000$, $SD = 1.066$; attention $M = 0.000$, $SD = 0.996$; processing speed/executive function $M = 0.000$, $SD = 1.002$; and cognitive reserve $M = 0.000$, $SD = 0.868$, all $p > 0.05$).

Absolute fit for the model testing scalar invariance was also acceptable (RMSEA = 0.080; CFI = 0.905; $\chi^2 = 385.764$, $df = 139$), and produced estimated factor scores that were all significantly different by group (Control group memory/language $M = 0.000$, $SD = 0.714$; attention $M = 0.000$, $SD = 0.764$; processing speed/executive function $M = 0.000$, $SD = 0.845$; and cognitive reserve $M = 0.000$, $SD = 0.778$; aMCI-AD group memory/language $M = -2.180$, $SD = 1.053$; attention $M = -0.851$, $SD = 0.999$; processing speed/executive function $M = -1.532$, $SD = 1.001$; and cognitive reserve $M = -0.424$, $SD = 0.864$, all $p < 0.001$).

Indices of both relative and absolute model fit did not support strict invariance (RMSEA = 0.087; CFI = 0.877; $\chi^2 = 471.145$, $df = 151$; $\Delta CFI = 0.028$). Estimated factor scores for the strict invariance model were significantly different between groups for all but the cognitive reserve factor (Control group memory/language $M = 0.000$, $SD = 0.725$; attention $M = 0.000$, $SD = 0.824$; processing speed/executive function $M = 0.000$, $SD = 0.855$; and cognitive reserve $M = 0.389$, $SD = 0.310$; aMCI-AD group memory/language $M = -2.179$, $SD = 1.038$; attention $M = -0.857$, $SD = 0.957$; processing speed/executive function $M = -1.525$, $SD = 0.988$; and cognitive reserve $M = -0.391$, $SD = 0.393$). Examination of model modification indices suggested that the memory and language indicators were contributing most to model misfit starting with the model testing for metric invariance. This suggests that factor loadings for memory and language measures were different in controls versus the aMCI group and model misfit arose by constraining these loadings to be equivalent across groups. We thus tested partial metric invariance by allowing memory and language factor loadings to vary by group and found that model fit was similar to the model testing configural invariance (RMSEA = 0.069; CFI = 0.936; $\chi^2 = 291.177$, $df = 125$, $\Delta CFI = 0.008$). Partial scalar invariance (allowing memory and language loadings and intercepts to vary by group), similarly showed little change from the model testing partial metric invariance, (RMSEA = 0.073; CFI = 0.924; $\chi^2 = 331.284$, $df = 134$, $\Delta CFI = 0.012$), but the model testing for partial strict invariance (allowing memory and language loadings, intercepts, and residual variances to vary by group) showed significant deterioration in model fit (RMSEA = 0.080; CFI = 0.899; $\chi^2 = 410.000$, $df = 147$, $\Delta CFI = 0.025$). A summary of standardized factor loadings, absolute model fit, and relative model fit statistics across levels of invariance testing are summarized in Table 2.

Discussion

These results provide support for a four-factor CDCR model of several related aspects of cognition and cognitive reserve in both normal aging and memory impaired older adults along the AD clinical spectrum. There was some evidence for the stability of the model across groups in that indices of absolute model fit were acceptable when assuming configural, metric, and scalar invariance, but not strict invariance. Across all levels of invariance testing, however, there was a significant change in model fit from the previous, less constrained model, and model modification indices suggested that the source of misfit was due to group differences in factor loadings of the memory and language measures. This finding is not surprising, as we were comparing performance from a cognitively normal group to a memory-impaired group of participants. Tests of partial invariance suggested that when allowing memory and language measures to vary by group, there was evidence for partial metric and scalar invariance, but not strict invariance. This suggests that while memory and language performance differs by group in terms of factor loadings and intercepts, the other factors in our model remain relatively stable across groups, but residual variation in test performance was significantly different across groups for all factors included in our model.

Regarding the relationship between the cognitive reserve and processing speed/executive function constructs, our results indicate that these theoretical constructs may have less shared variance than previously observed and provide supporting evidence for the discriminant validity of cognitive reserve as a unique construct. While the CDCR model in both normal and memory impaired samples of older adults demonstrated medium-to-large correlations across all constructs, the correlations in both samples between the cognitive reserve and processing speed/executive function factors were in the medium range. This is illustrated in figures 1 and 2, where it is shown that the intercorrelation between cognitive reserve and processing speed/executive function is 0.431 in the control group, and 0.433 in the aMCI-AD group, both of which constitute what is considered medium-sized correlations (Cohen, 1988). In contrast, intercorrelations between latent variables that have previously been established to be distinct constructs in latent variable models, such as memory and executive functioning (Salthouse, Atkinson, & Berish, 2003), showed even higher inter-factor correlations in both samples ($r = 0.711$ for the control group and $r = 0.697$ for the aMCI-AD group). These results thus support that cognitive reserve is not just processing speed or executive function. Functional neuroimaging evidence also supports this assertion; in a study probing the relationship between task-related activation on a working memory task and measures of cognitive reserve, Stern and colleagues demonstrated a cognitive reserve-related network that clearly included frontal areas, but also included medial temporal areas (Stern, et al., 2008).

There nonetheless remains an interesting theoretical relationship between cognitive reserve and executive function, which may both be subserved by the same, similar, or highly overlapping frontal networks and cognitive systems. For example, it is plausible to conceive that individuals with more cognitive flexibility, a commonly referenced component of executive function, would be able to adapt better to brain dysfunction than those with limited cognitive flexibility. As was argued by Satz et al. (2011), there are likely several cognitive processes that contribute to and are related to cognitive reserve. Although our model was not identical to the hypothetical model presented by Satz et al., both models suggest that cognitive reserve is distinct from the executive function and processing resources factors in their model and from the processing speed/executive function and attention factors in our CDCR model.

This study has several strengths; it assesses a theoretically-driven latent-variable model of the interrelations of cognitive reserve with other cognitive factors in both normal aging and across the aMCI-AD clinical spectrum; uses a large sample of well-characterized participants representative of the UDS and ADC research populations (Weintraub, et al., 2009); and employs known neuropsychological measures, many of which have defined age-, education- and gender-adjusted normative ranges available (Grober & Sliwinski, 1991; Shirk, et al., 2011; Weintraub, et al., 2009). The use of the FCSRT is a particular strength, as it is an increasingly recognized measure to discriminate between retrieval problems associated with normal aging and storage-based memory dysfunction characteristic of AD (Grober, et al., 2010; Rami, et al., 2012). The utilization of the AMNART is also important as it correlates highly with verbal IQ and general intelligence, or g , is relatively stable through adulthood, and is relatively resistant to major decline until the later stages of dementia (Patterson, Graham, & Hodges, 1994). This study is also the first to examine the construct validity of cognitive reserve in the UDS framework – a framework and dataset that is large, longitudinal, expanding, and can be utilized in the future to further validate and refine the CDCR model.

The study also has several inherent limitations, some of which may be remedied or assessed through future research. Participants were a majority White and highly educated, and while consistent with the UDS sample nationally, this limits generalizability and likely limited the

model's fit due to the relatively narrow range of variation in the indicator variables for cognitive reserve (i.e., AMNART and education), and, to a lesser degree, for the cognitive domain indicator variables. Having only two indicator variables each for cognitive reserve and attention in our model also places substantial burden on the data-model covariance fit structure and may have limited model fits; the model fits were nonetheless acceptable. This does not prove that a four-factor model, or this four-factor CDCR model, is either necessary or the correct model; it does, however, provide evidence that the four-factor CDCR model is sufficient to explain the observed data. Future datasets with appropriate additional indicator variables for these constructs, preferably consisting of tests with low inter-correlations within each domain, should improve accuracy and model fits.

Another set of limitations relates to the battery of tests included in our analysis. Ideally, we would have had a battery that included sufficient indicators of processing speed, executive function, memory, and language to model each of these as their own, distinct constructs. While less than ideal, in the context of this dataset's limitations, we hypothesized a joint processing speed/executive function construct, and a joint memory/language construct. These combinations were driven by the particular nature of the available tests. For the processing speed/executive function construct, all tests included require speed of information processing, but two of the tests (Trails B and Dig Sym) have additional executive components including maintaining cognitive set. For the memory and language construct, all tests included require verbal processing or semantic knowledge and performance heavily relies on integrated medial temporal and frontal network hubs. It is noteworthy that in both the controls and AD-spectrum groups semantic fluency was the highest loaded test on this combined (memory/language) latent construct – a measure that arguably best represents and relies on such an amalgamated construct of memory and language, and their shared frontal and medial temporal networks. It is also important to note that when performance range and variability increase, as in the AD-Spectrum group, the loadings of this combined memory/language latent variable onto memory tests increase substantially to the 0.7–0.8 range.

The interplay of AD pathology and disease burden in modulating the relationship of neuropsychological domains and cognitive reserve merits further investigation within the UDS framework. Given the large scope of this national database, our model could be further validated and extended longitudinally (Locascio & Atri, 2011) in this larger sample, particularly when complemented by similar supplementary measures to the basic UDS battery collected by several UDS sites including proxies of cognitive reserve. The cognitive constructs in the current study can be used to produce latent trait scores for future extensions of the CDCR model to potentially predict and relate to other outcomes of interest. While these loadings demonstrate magnitude of inter-correlations, a limitation of a cross-sectional model is that it cannot indicate directionality of cause-effect changes dynamically as participants remain stable or experience cognitive decline. However, with increasing evidence for cognitive domain specific differences in relation to dementia and response to treatments (Persson, Wallin, Levander, & Minthon, 2009; Tinklenberg, et al., 1990; Zec, et al., 1992), latent trait scores for each domain can be used in future research as starting points in longitudinal models in the presence of multiple covariates and outcomes.

The use of latent variables for cognition in a model that also includes cognitive reserve provides a promising approach for tracking cognitive changes in AD to assess treatment outcomes (Atri, Shaughnessy, Locascio, & Growdon, 2008) or early detection of deviation from normal aging (Stern, 2009). The memory/language and processing speed/executive function latent variables may be particularly sensitive measures for tracking AD-related cognitive decline, and the latent cognitive reserve variable could be tested in a longitudinal model to determine how it moderates rate of cognitive decline. These types of measurement

and analytic factors may play an even more crucial role in successful trial design as the AD clinical trial field focuses on earlier stages of disease and extends the duration of studies, where increasingly sensitive measures are needed to quantify differences in group trajectories due to treatment-related disease modification. Such methods could also allow for earlier detection, by providing a more individualized approach to determining relative decline while adjusting for estimated premorbid level of cognitive functioning and taking into account how potential proxies for cognitive reserve may not only modify cognitive test performance but also moderate levels of observed neuroimaging biomarkers of AD such as PIB amyloid load (Rentz, et al., 2010). In summary, this model further supports the construct validity of cognitive reserve, that it shares similarities but is not the same as processing resources or executive functions, and lays the groundwork to test the CDCR model longitudinally.

Acknowledgments

We would like to acknowledge the Clinical, Administrative, and Biostatistics and Bioinformatics Cores of the Massachusetts Alzheimer's Disease Research Center (NIA 5 P50AG05134 Growdon and Hyman), and the Bedford Division of the New England Geriatric Research Education and Clinical Center (GRECC) at the ENRM Veterans Administration (VA) Hospital. The contents of this study do not represent the views of the Department of Veterans Affairs or the United States Government. We would also like to particularly acknowledge Dr. Daniel Mungas and the Friday Harbor Advanced Psychometric Methods for Cognitive Aging Research workshop (R13AG030995, PI: Dan Mungas), Dr. Joseph Locascio, Dr. Rebecca England, Dr. Dorene Rentz, Dr. Alden Gross, Dr. Richard Jones, and Dr. Liang Yap for providing significant assistance, guidance, teaching, resources and/or helpful comments. Finally, and most importantly, we express our deep gratitude for the commitment of the MADRC UDS Longitudinal Cohort study participants without whose generous contribution and dedication this research would not be possible. This study was funded by NIA K23AG027171 (Atri).

References

- Alzheimer's Association. 2011 Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. 2011; 4(2):110–133.
- Atri A, Rountree SD, Lopez OL, Doody RS. Validity, significance, strengths, limitations, and evidentiary value of real-world clinical data for combination therapy in Alzheimer's disease: comparison of efficacy and effectiveness studies. *Neuro-degenerative Diseases*. 2012; 10:170–174. [PubMed: 22327239]
- Atri A, Shaughnessy LW, Locascio JJ, Growdon JH. Long-term course and effectiveness of combination therapy in Alzheimer disease. *Alzheimer Disease and Associated Disorders*. 2008; 22(3):209–221. [PubMed: 18580597]
- Beekly DL, Ramos EM, Lee WW, Deitrich WD, Jacka ME, Wu J, Kukull WA. The National Alzheimer's Coordinating Center (NACC) database: the Uniform Data Set. *Alzheimer Disease and Associated Disorders*. 2007; 21(3):249–258. [PubMed: 17804958]
- Bentler, P.; Chou, C. Practical issues in structural equation modeling. In: Long, J., editor. *Common Problems/Proper Solutions: Avoiding Error In Quantitative Research*. Newbury Park, CA: Sage; 1988.
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107(2):238–246. [PubMed: 2320703]
- Blom, G. *Statistical Estimates and Transformed Beta Variables*. New York: John Wiley & Sons, Inc; 1958.
- Browne, M.; Cudek, R. Alternative ways of assessing model fit. In: Bollen, K.; Long, J., editors. *Testing Structural Equation Models*. Thousand Oaks, CA: Sage; 1993. p. 136–162.
- Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*. 2002; 9(2):233–255.
- Cohen, JD. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, N.J: L. Erlbaum Associates; 1988.
- DeKosky S. Early intervention is key to successful management of Alzheimer disease. *Alzheimer Disease and Associated Disorders*. 2003; 17(Suppl 4):S99–104. [PubMed: 14512815]

- Dickerson BC, Bakkour A, Salat DH, Feczko E, Pacheco J, Greve DN, Buckner RL. The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals. *Cerebral Cortex*. 2009; 19(3):497–510. [PubMed: 18632739]
- Embretson, SE.; Reise, SP. Item Response Theory for Psychologists. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2000.
- Gallo JJ, Anthony JC, Muthén BO. Age differences in the symptoms of depression: a latent trait analysis. *Journals of Gerontology, Psychological Sciences*. 1994; 49(6):251–264.
- Grober E, Lipton RB, Hall C, Crystal H. Memory impairment on free and cued selective reminding predicts dementia. *Neurology*. 2000; 54(4):827–832. [PubMed: 10690971]
- Grober E, Sanders AE, Hall C, Lipton RB. Free and cued selective reminding identifies very mild dementia in primary care. *Alzheimer Disease and Associated Disorders*. 2010; 24(3):284–290. [PubMed: 20683186]
- Grober E, Sliwinski M. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*. 1991; 13(6): 933–949. [PubMed: 1779032]
- Hauser, RM.; Goldberger, AS. The Treatment of Unobservable Variables in Path Analysis. In: Costner, HL., editor. *Sociological Methodology*. San Francisco: Jossey-Bass; 1971. p. 81-117.
- Hayden KM, Jones RN, Zimmer C, Plassman BL, Browndyke JN, Pieper C, Welsh-Bohmer KA. Factor structure of the National Alzheimer's Coordinating Centers Uniform Dataset Neuropsychological Battery: An evaluation of invariance between and within groups over time. *Alzheimer Disease and Associated Disorders*. 2011; 25(2):128–137. [PubMed: 21606904]
- Hu L, Bentler P. Fit indices in covariance structure analysis: Sensitivity to underparameterized model misspecifications. *Psychological Methods*. 1998; 4:424–453.
- Johnson D, Storandt M, Morris J, Langford Z, Galvin J. Cognitive profiles in dementia: Alzheimer disease vs healthy brain aging. *Neurology*. 2008; 71(22):1783–1789. [PubMed: 19029518]
- Lee JH. Genetic evidence for cognitive reserve: Variations in memory and related cognitive functions. *Journal of Clinical and Experimental Neuropsychology*. 2003; 25(5):594–613. [PubMed: 12815498]
- Locascio J, Atri A. An Overview of Longitudinal Data Analysis Methods for Neurological Research. *Dementia and Geriatric Cognitive Disorders Extra*. 2011; 1:330–357. [PubMed: 22203825]
- Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: Shortened versions for use in Alzheimer's disease. *J Gerontol*. 1992; 47(3):P154–158. [PubMed: 1573197]
- Morris JC. The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*. 1993; 43(11):2412–2414. [PubMed: 8232972]
- Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*. 1989; 39(9):1159–1165. [PubMed: 2771064]
- Morris JC, Weintraub S, Chui HC, Cummings J, Decarli C, Ferris S, Kukull WA. The Uniform Data Set (UDS): Clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Disease and Associated Disorders*. 2006; 20(4):210–216. [PubMed: 17132964]
- Muthén, B. The Development of Heavy Drinking and Alcohol Related Problems from Ages 18 to 37 in a US National Sample. Los Angeles: Graduate School of Education and Information Studies; 1998.
- Muthén, B.; Khoo, S-T.; Francis, D. CSE Technical Report No 489. Los Angeles: UCLA Graduate School of Education and Information Studies; 1998. Multi-stage Analysis of Sequential Developmental Processes to Study Reading Progress: New Methodological Developments Using General Growth Mixture Modeling.
- Muthén BO. Latent variable modeling in heterogeneous populations. *Meetings of Psychometric Society* (1989, Los Angeles, California and Leuven, Belgium). *Psychometrika*. 1989; 54(4):557–585.
- Patterson KE, Graham N, Hodges JR. Reading in dementia of the Alzheimer type: A preserved ability? *Neuropsychology*. 1994; 8(3):395–412.

- Persson C, Wallin A, Levander S, Minthon L. Changes in cognitive domains during three years in patients with Alzheimer's disease treated with donepezil. *BMC Neurology*. 2009; 9(1):7. [PubMed: 19208247]
- Reitan, RM.; Wolfson, D. The Halstead-Reitan Neuropsychological Test Battery. 2. Tucson, AZ: Neuropsychology Press; 1985.
- Rentz D, Huh T, Faust R, Budson A, Scinto L, Sperling R, et al. Use of IQ-adjusted norms to predict progressive cognitive decline in highly intelligent older individuals. *Neuropsychology*. 2004; 18:38–49. [PubMed: 14744186]
- Rentz DM, Locascio JJ, Becker JA, Moran EK, Eng E, Buckner RL, Johnson KA. Cognition, reserve, and amyloid deposition in normal aging. *Annals of Neurology*. 2010; 67(3):353–364. [PubMed: 20373347]
- Roe CM, Mintun MA, Ghoshal N, Williams MM, Grant EA, Marcus DS, Morris JC. Alzheimer disease identification using amyloid imaging and reserve variables: Proof of concept. *Neurology*. 2010; 75(1):42–48. [PubMed: 20603484]
- Salthouse T, Atkinson T, Berish D. Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology General*. 2003; 132(4): 566–594. [PubMed: 14640849]
- Satorra A, Bentler P. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*. 2001; 66(4):507–514.
- Satz P, Cole MA, Hardy DJ, Rassovsky Y. Brain and cognitive reserve: Mediator(s) and construct validity, a critique. *Journal of Clinical and Experimental Neuropsychology*. 2011; 33(1):121–130. [PubMed: 20680883]
- Shirk S, Mitchell M, Shaughnessy L, Sherman J, Locascio J, Weintraub S, Atri A. A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. *Alzheimer's Research & Therapy*. 2011; 3(6):32.
- Siedlecki K, Stern Y, Reuben A, Sacco R, Elkind M, Wright C. Construct validity of cognitive reserve in a multiethnic cohort: The Northern Manhattan Study. *Journal of the International Neuropsychological Society*. 2009; 15(04):558–569. [PubMed: 19573274]
- Stern Y. Cognitive reserve. *Neuropsychologia*. 2009; 47:2015–2028. [PubMed: 19467352]
- Stern Y, Gurland B, Tatemichi TK, Tang MX, Wilder D, Mayeux R. Influence of education and occupation on the incidence of Alzheimer's disease. *Journal of the American Medical Association*. 1994; 271(13):1004–1010. [PubMed: 8139057]
- Stern Y, Zarahn E, Habeck C, Holtzer R, Rakitin B, Kumar A, Brown T. A common neural network for cognitive reserve in verbal and object working memory in young but not old. *Cerebral Cortex*. 2008; 18(4):959. [PubMed: 17675368]
- Tinklenberg J, Brooks JO, Tanke ED, Khalid K, Poulsen SL, Kraemer HC, Yesavage JA. Factor analysis and preliminary validation of the Mini-Mental State Examination from a longitudinal perspective. *International Psychogeriatrics*. 1990; 2(02):123–134. [PubMed: 2101301]
- Wechsler, D. Wechsler Adult Intelligence Scale-Revised. San Antonio: Psychological Corporation; 1987.
- Wechsler, DA. Wechsler Memory Scale - Revised. San Antonio: Psychological Corporation; 1987.
- Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, Morris JC. The Alzheimer's Disease Centers' Uniform Data Set (UDS): The neuropsychologic test battery. *Alzheimer Disease and Associated Disorders*. 2009; 23(2):91–101. [PubMed: 19474567]
- Wilson R, Scherr P, Schneider J, Tang Y, Bennett D. Relation of cognitive activity to risk of developing Alzheimer disease. *Neurology*. 2007; 69(20):1911–1920. [PubMed: 17596582]
- Zec R, Landreth E, Vicari S, Belman J, Feldman E, Andrise A, Kumar V. Alzheimer Disease Assessment Scale: a subtest analysis. *Alzheimer Disease and Associated Disorders*. 1992; 6(3): 164–181. [PubMed: 1485931]

Figure 1.

The CDCR (cognitive domains and cognitive reserve) model results: Results of confirmatory factor analysis (CFA) of cognitive reserve and neuropsychological measures in a group of healthy older adults ($n = 294$). $\chi^2 = 100.12$, $df = 78$, CFI = .962, RMSEA = .049. Variables in boxes represent observed measures, variables in ovals represent latent variables, numbers along straight arrows represent factor loadings for each indicator variable on the respective latent variable, and numbers along curved arrows represent correlations between latent variables and correlations between indicator variables. LMI & LMII = Wechsler Memory Scale-Revised (WMS-R) subtests Logical Memory IA and IIA; FCSRT Free & FCSRT Total = Free and Cued Selective Reminding Test Free Recall and Total Recall; BNT = Boston Naming Test (30 item – odd numbered); SEMFLU = Semantic Fluency (Animals and Vegetables); Digit Fwd & Digit Bkwd = WMS-R subtests Digit Span Forward and Backward; Trails A & Trails B = Trail Making Test Parts A and B; Dig Sym = Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Symbol Coding subtest; AMNART = American National Adult Reading Test; CFI = comparative fit index; RMSEA = root mean square error of approximation; aMCI = amnesic mild cognitive impairment; AD = Alzheimer's disease.

Figure 2.

The CDCR (cognitive domains and cognitive reserve) model results in a group with aMCI or AD ($n=265$). $\chi^2 = 1750.02$, $df = 78$, CFI = .932, RMSEA = .085. LMI & LMII = Wechsler Memory Scale-Revised (WMS-R) subtests Logical Memory IA and IIA; FCSRT Free & FCSRT Total = Free and Cued Selective Reminding Test Free Recall and Total Recall; BNT = Boston Naming Test (30 item – odd numbered); SEMFLU = Semantic Fluency (Animals and Vegetables); Digit Fwd & Digit Bkwd = WMS-R subtests Digit Span Forward and Backward; Trails A & Trails B = Trail Making Test Parts A and B; Dig Sym = Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Symbol Coding subtest; AMNART = American National Adult Reading Test; CFI = comparative fit index; RMSEA = root mean square error of approximation; aMCI = amnesic mild cognitive impairment; AD = Alzheimer's disease.

Table 1

Descriptive statistics.

Variable	Controls			aMCI-AD			Entire Sample		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Age**	294	69.58	8.907	265	76.21	8.790	559	72.72	9.444
Years Education	294	16.44	2.514	265	15.80	3.215	559	16.13	2.883
Mini Mental State Examination**	293	29.19	1.034	265	25.46	4.247	558	27.42	3.547
Logical Memory Ia**	292	14.17	3.443	265	7.76	4.974	557	11.12	5.311
Logical Memory IIa**	292	13.31	3.492	265	5.71	5.028	557	9.69	5.730
Free & Cued Free Recall**	132	32.84	5.045	115	17.15	10.011	247	25.53	11.024
Free & Cued Total Recall**	132	47.43	1.285	115	38.45	11.860	247	43.25	9.285
Boston Naming Test, 30-odd items**	292	9.14	1.872	264	8.01	2.282	556	25.38	5.433
Semantic Fluency**	293	7.35	2.216	265	5.75	2.112	558	30.84	10.644
Digit Span Forward**	293	36.73	7.956	265	24.33	9.380	558	8.60	2.150
Digit Span Backward**	293	27.84	2.038	265	22.66	6.601	558	6.59	2.309
Trail Making Test A**	293	30.23	12.147	260	51.69	32.721	553	40.32	26.370
Trail Making Test B**	291	78.30	40.402	250	158.39	88.326	541	115.31	77.918
Digit Symbol Coding**	293	50.46	11.631	258	34.50	14.463	551	42.98	15.268
AMNART*	171	123.09	7.118	133	118.92	10.594	304	121.26	9.034

* Indicates significant differences between groups ($p = 0.001$);** Indicates $p < 0.001$.

AMNART = American National Adult Reading Test.

Table 2

Standardized factor loadings and indices of model fit for confirmatory factor analysis (CFA) models for control participants (n = 294) and participants with memory impairment along the Alzheimer's disease spectrum (aMCI-AD; n=265) across levels of factorial invariance.

Indicator	Configural invariance		Metric invariance		Scalar invariance		Strict invariance	
	Controls	aMCI-AD	Controls	aMCI-AD	Controls	aMCI-AD	Controls	aMCI-AD
Memory, Language								
Logical Memory Ia	0.426	0.663	0.589	0.560	0.634	0.599	0.618	0.618
Logical Memory IIa	0.370	0.672	0.549	0.556	0.637	0.651	0.638	0.638
Free & Cued Free Recall	0.425	0.846	0.583	0.733	0.670	0.823	0.734	0.734
Free & Cued Total Recall	0.176	0.761	0.550	0.476	0.647	0.582	0.624	0.624
Boston Naming Test	0.524	0.633	0.655	0.531	0.601	0.475	0.539	0.539
Semantic Fluency	0.629	0.869	0.740	0.856	0.675	0.763	0.708	0.708
Attention								
Digit span-FWD	0.558	0.735	0.679	0.665	0.668	0.653	0.651	0.651
Digit span-BWD	0.815	0.883	0.796	0.888	0.807	0.901	0.850	0.850
Speed, Executive Function								
Trails A	0.768	0.802	0.787	0.788	0.765	0.764	0.768	0.768
Trails B	0.802	0.860	0.804	0.855	0.815	0.868	0.838	0.838
Digit Symbol Coding	0.741	0.816	0.787	0.783	0.790	0.783	0.788	0.788
Cognitive Reserve								
AMNART	0.870	1.0	0.958	1.0	0.960	1.0	0.992	1.0
Years Education	0.570	0.479	0.561	0.463	0.556	0.460	0.494	0.494
Model fit statistics								
RMSEA	0.067		0.076		0.080		0.089	
CFI	0.944		0.923		0.909		0.879	
χ^2 , df	262.510, 117		341.845, 130		385.764, 139		471.145, 151	
S-B χ^2 p value, df	--		<0.001, 13		<0.001, 9		<0.001, 12	
ACFI	--		0.026		0.013		0.028	

AMNART = American National Adult Reading Test; S-B χ^2 = Satorra-Bentler scaled χ^2 test (Satorra & Bentler, 2001).