MEASURING DEMENTIA OF THE ALZHEIMER TYPE MORE PRECISELY

A Thesis

by

DEBORAH ANNE LOWE

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Steve Balsis
Committee Members,	Rick Peterson
	Darrell Worthy
Head of Department,	Paul Wellman

December 2012

Major Subject: Psychology

ABSTRACT

Alzheimer's disease (AD) progressively impairs cognitive and functional abilities. Research on pharmacological treatment of AD is shifting to earlier forms of the disease, including preclinical stages. However, assessment methods traditionally used in clinical research may be inappropriate for these populations. The Alzheimer Disease Assessment Scale-cognitive (ADAS-cog), a commonly used cognitive battery in AD research, is most sensitive in the moderate range of cognitive impairment. It focuses on immediate recall and recognition aspects of memory rather than retention and delayed recall. As clinical trials for dementia continue to focus on prodromal stages of AD, instruments need to be retooled to focus on cognitive abilities more prone to change in the earliest stages of the disease. One such domain is delayed recall, which is differentially sensitive to decline in the earliest stages of AD. A supplemental delayed recall subtest for the ADAS-cog is commonly implemented, but we do not know precisely *where* along the spectrum of cognitive dysfunction this subtest yields incremental information beyond what is gained from the standard ADAS-cog. An item response theory (IRT) approach can analyze this in a psychometrically rigorous way. This study's aims are twofold: (1) to examine where along the AD spectrum the delayed recall subtest yields optimal information about cognitive dysfunction, and (2) to determine if adding delayed recall to the ADAS-cog can improve prediction of functional outcomes, specifically patients' ability to complete basic and instrumental activities of daily living.

Results revealed differential functioning of ADAS-cog subtests across the dimension of cognitive impairment. The delayed recall subtest provided optimal information and increased the ADAS-cog's measurement precision in the relatively mild range of cognitive dysfunction. Moreover, the addition of delayed recall to the ADAS-cog, consistent with my hypothesis, increased covariation with instrumental but not basic activities of daily living. These findings provide evidence that the delayed recall subtest slightly improves the ADAS-cog's ability to capture information about cognitive impairment in the mild range of severity and thereby improves prediction of instrumental functional deficits.

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Steve Balsis, and my committee members, Dr. Rick Peterson and Dr. Darrell Worthy, for their support and guidance. I would also like to thank Dr. Rachelle Doody for her expertise and generous collaboration with the data. Finally, I would like to express appreciation for Dr. Aaron Taylor's consultation on statistical analyses.

NOMENCLATURE

AD	Alzheimer's Disease
ADAS-cog	Alzheimer's Disease Assessment Scale-cognitive
ADL	Activity of Daily Living
DAT	Dementia of the Alzheimer Type
IADL	Instrumental Activity of Daily Living
IRT	Item Response Theory
MCI	Mild Cognitive Impairment
SMI	Subjective Memory Impairment

TABLE OF CONTENTS

Page

LIST OF FIGURES

1. INTRODUCTION AND LITERATURE REVIEW

Alzheimer's disease (AD) is a progressive, neurodegenerative disease that is estimated to currently afflict 5.4 million Americans and up to 26.6 million people worldwide (Alzheimer's Association, 2012; Chamberlain et al., 2011; National Plan to Address Alzheimer's Disease, 2012). Incidence rates double every five years after age 65, so that AD affects about six percent of those between the ages of 65 and 74, but approximately 46 percent of adults 85 and older (Alzheimer's Association, 2012; Mendez & Cummings, 2003). The prevalence of AD worldwide is estimated to quadruple by 2050 (Chamberlain et al., 2011), and with the rapidly increasing older population (from 40 million in 2010 to a projected 72.1 million in 2030; Alzheimer's Association, 2012), AD is clearly a major societal concern. The U.S. Department of Health and Human Services recently developed the National Plan to Address Alzheimer's Disease, stimulated by President Barack Obama signing into law the National Alzheimer's Project Act on January 4, 2011. The National Plan sets forth the goal of developing effective forms of prevention and treatment of AD by 2025. These federal initiatives have brought AD to the forefront of national attention and mirror the drive within the field to better detect preclinical forms of AD, including mild cognitive impairment (MCI; Morris et al., 2001) and even earlier preclinical stages. Being able to better detect these prodromal stages will aid in earlier treatment of cognitive impairment and strengthen understanding of the AD process, with the eventual goal of prevention.

To ultimately achieve these goals of effective treatment and prevention of AD, it is important to understand that the disease is dimensional and progressive, as it increasingly impairs not only neuroanatomical integrity but also cognitive, behavioral, and functional domains. The course of AD is marked by a long preclinical period, during which a gradual decline in episodic memory typically occurs (Riepe, Janetzky, & Lemming, 2011). The earliest stage is characterized by subjective memory impairment (SMI; i.e., memory complaints), which has recently been shown to be a predictor for future development of a dementing condition. Approximately 10% of individuals with SMI converted to MCI over a three-year period in one study (Jessen et al., 2010). Recent work also has revealed differences in functional and structural brain imaging between individuals in their 60s with or without SMI (Scheef et al., 2012). The individuals with SMI showed hypometabolism in the right precuneus as well as hypermetabolism in and reduced volume of the right hippocampus, two neuroanatomical regions associated with pathophysiological processes in AD. Moreover, these SMI individuals showed differentially greater longitudinal declines in their performances on episodic memory tests relative to participants without SMI (Scheef et al., 2012). This recent work has established support for SMI as the earliest prodromal stage of the AD spectrum.

As memory continues to decline and patients move into the MCI and dementia of the Alzheimer type (DAT) stages, memory impairment is followed by deficits in attention, language, visuospatial orientation, and executive functions (Riepe et al., 2011). Language impairment includes word-finding difficulties, with later problems in spontaneous speech, comprehension, and vocal repetitions of words or sounds (Mendez

& Cummings, 2003). Visuospatial deficits include difficulties in copying figures and orienting oneself to one's environment (e.g., getting lost). Attentional and executive dysfunction in part affects planning, abstract reasoning, insight, working memory, sustained and divided attention, and inhibition of responses. Impairment in praxis typically occurs in the middle to late stages of the disease process and is evident by problems performing overlearned motor behaviors, such as brushing one's teeth or dressing oneself (Mendez & Cummings, 2003). AD also produces behavioral symptoms, including mood changes, agitation, delusions, aberrant motor behaviors, and sleep disturbances in almost all patients at some point in the disease (Gauthier et al., 2010). Moreover, patients with AD demonstrate subtle difficulties in instrumental activities of daily living (IADLs; e.g., driving, finances, medication management) as early as MCI (Griffith et al., 2003; Perneczky et al., 2006), with progressively greater impairments in IADLs and ultimately in more basic activities of daily living (ADLs; e.g., eating, bathing, toileting behaviors) as the disease progresses. In the most severe and latest stages of the disease, AD patients are unable to care for themselves and are entirely dependent upon others (Mendez & Cummings, 2003).

Because AD is an ever-evolving entity, accurate and appropriate assessment varies across the disease spectrum. For example, early in the course of the disease, even preclinically, subtle neuropsychological deficits can be detected in full cognitive batteries, but functional impairments are relatively subtle. However, in the latest stages of the disease, patients typically cannot respond meaningfully to cognitive batteries, so appropriate assessment often consists of behavioral ratings, brief cognitive screenings,

and physical self-maintenance measurement (Hermann & Gauthier, 2008). The progressive nature of AD can make evaluating the impact of disease-modifying treatments difficult. Multiple outcome measures are typically used, including assessments of cognitive functioning, daily functional impairments, neuropsychiatric symptoms, clinician ratings, and caregiver burden. In terms of cognitive functioning, the "gold standard" for assessment in clinical research has historically been the Alzheimer's Disease Assessment Scale-cognitive (ADAS-cog; Mohs, Rosen, & Davis, 1983; Rosen, Mohs, & Davis, 1984; Kirk, 2007; Rockwood, Fay, Gorman, Carver, & Graham, 2007).

The ADAS-cog, which assesses basic linguistic, memory, orientation, and praxis functions, was specifically designed for use in moderate AD and has been shown to be most sensitive to impairment in cognitive functioning in the moderate stages of the disease (Benge, Balsis, Geraci, Massman, & Doody, 2009). Benge and colleagues (2009) determined that the ADAS-cog discriminates best at moderate levels of Alzheimer's, specifically between -1.00 and 1.74 standard deviations (*SD*s) of cognitive impairment in a largely demented sample. Similarly, research has shown a curvilinear relationship between dementia severity and ADAS-cog scores (Benge et al., 2009; Ito et al., 2011), so that there must be relatively large changes in cognitive dysfunction at the mild and severe stages of AD before the ADAS-cog scores reflect those changes. However, in the moderate stages of AD, ADAS-cog scores are sensitive to smaller differences in cognitive dysfunction. Other studies have found a nonequivalent rate of longitudinal change in ADAS-cog scores across different stages of the AD spectrum. Irizarry and colleagues (2008) found that a moderate AD group had a 24-week change in ADAS-cog

score of 2.99 points, but a mild AD group had a change score of 0.19 points, indicating minimal change in the mild group but larger change in the moderate AD group. Over a 12-month period, patients with moderate AD at baseline declined more on the total score and most individual subtests than did the patients with mild AD (Sevigny, Peng, Liu, & Lines, 2010). In another 12-month longitudinal study, patients with mild and moderate AD showed a similar change in their ADAS-cog total scores at one-year follow-up, but the patients with severe AD had notably less change in their ADAS-cog score (Suh, Ju, Yeon, & Shah, 2004), which is likely indicative of a floor effect. These results reveal the nonlinearity of the ADAS-cog across the AD spectrum, so that changes in observed scores over time do not represent equivalent changes in underlying cognitive impairment along the dimension of disease severity. The ADAS-cog is most sensitive to smaller changes in cognitive functions in the moderate stages of dementia but is not as good at capturing information about cognitive dysfunction in the mild and severe stages of AD.

Because the ADAS-cog was designed to measure cognitive dysfunction in the moderate stages of AD, it focuses on immediate recall and recognition aspects of memory rather than retention and delayed recall of information. However, as AD clinical trials and basic AD research increasingly focus on preclinical or prodromal stages of AD, instruments need to be retooled to focus on the cognitive abilities that are more prone to change in the earliest stages of the disease. One such domain is delayed recall, or ability to retain and retrieve information that was previously learned after a delay. Numerous studies have demonstrated that delayed recall is differentially sensitive to decline in the earliest stages of AD (cf., Chen et al., 2000; Fleisher et al., 2007;

Gallagher et al., 2010; Pozueta et al., 2011; Rabin et al., 2009; Rami et al., 2012; Tabert et al., 2006). However, the ADAS-cog eschews a measure of delayed recall in favor of more immediate recall. Considering its design for use in moderate AD, this was originally warranted, given the prominent deficits in learning and immediately recalling information later in the disease course. Yet to adequately track the disease across its course, this important domain of delayed recall must be assessed when conducting research with patients in the mild and preclinical stages of AD.

In an attempt to bolster the ADAS-cog's appropriateness for use with patients in the earlier stages of the AD spectrum, additional subtests are sometimes used with this test, including a delayed recall subtest (Mohs et al., 1997). Research shows that the delayed recall subtest strengthens the ability to discriminate between controls, patients with MCI, and patients with DAT (Grundman et al., 2004; Llano, Laforet, & Devanarayan, 2011; Petersen et al., 2010). Addition of the delayed recall subtest also may increase power to predict patient conversion from MCI to DAT (Fleisher et al., 2007; Llano et al., 2011). Sano and colleagues (2011) found that the inclusion of the delayed recall subtest strengthened the ability to detect longitudinal cognitive change in an MCI group, but it was not useful in DAT groups, even among those with mild dementia. Performance on delayed recall changed at a similar rate over 12 months for both the MCI and DAT groups. However, when the researchers examined a standardized index of change (12-month change in score/SD of change), they found that the inclusion of delayed recall increased sensitivity to longitudinal cognitive decline in the MCI group but not in the DAT group. They suggested that the insensitivity to cognitive change over

time in the DAT group was due to floor effects, such that the DAT patients already performed poorly on this subtest at baseline, so there was very little room for decline on the test over the 12-month period. In contrast, delayed recall was able to increase sensitivity to cognitive change in the MCI group because they were not as impaired at baseline and had more potential for decline to occur on the subtest. Considering these findings, then, the delayed recall subtest appears to strengthen the discriminative and predictive power of the ADAS-cog, particularly at the milder stages of the AD spectrum.

Although we know that adding delayed recall to the ADAS-cog can help distinguish between different groups along the AD spectrum and can aid in predicting longitudinal decline, we do not know precisely *where* along the spectrum of cognitive dysfunction (e.g., in the mild or moderate ranges) the delayed recall subtest yields incremental information beyond what is gained from the standard ADAS-cog. Based on findings discussed above that delayed recall improved sensitivity to longitudinal change in MCI but not DAT groups (Sano et al., 2011), it would seem that delayed recall likely provides information about cognitive dysfunction in the mild or preclinical stages of the AD spectrum. However, this speculation—self-evident as it may seem—has not been psychometrically analyzed. To do so, research needs to be conducted to examine how the delayed recall subtest functions across the continuum of AD-related cognitive dysfunction. This would provide psychometrically based information about where along the AD spectrum delayed recall can enhance the functioning of the ADAS-cog, discriminate between degrees of severity, and strengthen power for predicting longitudinal change.

An item response theory (IRT) approach provides a framework to analyze these issues. IRT takes into account the differential ability of each ADAS-cog subtest to relate to the latent construct (in this case, AD-associated cognitive dysfunction) along the spectrum of the disease. At any given point along the AD spectrum, particular ADAS-cog subtests may be more sensitive to and better able to capture information about cognitive dysfunction than other subtests. For example, in the moderate stage of AD, it could be that immediate recall of a list of words is strongly related to overall cognitive dysfunction while expressive speech is weakly related. Expressive speech might be strongly related to cognitive dysfunction in the later stages of AD, so it can provide more information about the latent construct among patients in the severe stages of AD but not as much information in the moderate stages. Similarly, word recall might be weakly related to cognitive dysfunction in the severe stages of the disease because of the floor effects of the test; patients with severe DAT are likely so impaired that they cannot recall any words on the subtest, so it cannot provide any discriminative information about dementia severity because everyone is performing the same. However, by weighting all subtests equally and summing them to gain a total score, we lose this information about cognitive impairment. Therefore, by using an IRT framework, we can determine where along the AD spectrum the delayed recall subtest can most optimally provide information about cognitive dysfunction and how its addition to the ADAS-cog changes the test's discriminative and predictive power. My first hypothesis is that the delayed recall subtest will be most sensitive to cognitive dysfunction in the mild stages of the AD

spectrum and so by adding delayed recall to the ADAS-cog, the test as a whole will better be able to discriminate between gradations of severity in this mild range.

The goals of increasing measurement precision and psychometrically determining where along the AD spectrum delayed recall provides information are critical to better enable research focused on detecting and treating AD-related cognitive impairment in the earliest stages. Being able to more precisely and accurately measure severity of cognitive impairment across the entire spectrum of the disease will likely strengthen ability to predict other important outcomes, such as predicting conversion from MCI to dementia and even from SMI to MCI. In addition, adding delayed recall to the ADAS-cog may strengthen its ability to predict functional impairment. Impairment in IADLs and ADLs progressively disables AD patients and is associated with global cognitive dysfunction. Deficits in IADLs include difficulties with more complex daily activities, such as driving, managing medications, handling finances, operating the telephone, and taking care of household needs (e.g., shopping, cleaning, cooking). Difficulties in these IADLs develop relatively early in the disease process, such that at least subtle problems in IADLs are often present in the MCI stage (Griffith et al., 2003; Perneczky et al., 2006). Assessment of these impairments in IADLs can improve prediction of future conversion to DAT (Gold, 2012; Luck et al., 2012; Rozzini et al., 2007). Impairments in basic ADLs occur later in the disease process, typically in the moderate to late stages (Hermann & Gauthier, 2008). ADLs include more basic functions, including bathing, dressing, eating, and toileting behaviors. This study will examine whether adding delayed recall to the ADAS-cog will strengthen its ability to

detect the association between cognitive dysfunction and these two classifications of functional impairments (i.e., IADLs and ADLs), which in turn can better inform clinical research on effective treatments for the range of deficits in AD, including cognitive and functional variables. My second hypothesis is that adding delayed recall to the ADAS-cog will improve its ability to detect an association between cognitive dysfunction and IADLs, but the inclusion of delayed recall will have a negligible impact on ability to detect the association between cognitive dysfunction and ADLs.

In sum, this study's aims are twofold: (1) to examine where along the AD spectrum the delayed recall subtest yields optimal information about cognitive dysfunction, and (2) to determine if adding delayed recall to the ADAS-cog can improve prediction of functional outcomes, specifically patients' ability to complete ADLs and IADLs. My first hypothesis is that the delayed recall subtest will add incremental information about cognitive dysfunction in the mildest stages of the AD spectrum and thus will improve the ability of the ADAS-cog as a whole to capture information about cognitive dysfunction in these mild ranges. Because IADLs are subject to decline in early AD but ADLs do not typically decline until later in the disease process, my second hypothesis is that adding delayed recall to the ADAS-cog will increase the ability to detect the association between cognitive dysfunction and IADLs but not basic ADLs.

2. METHOD

2.1 Participants

Participants were 788 patients who presented with memory complaints to the Baylor College of Medicine Alzheimer's Disease and Memory Disorders Center (ADMDC) in Houston, Texas. Ages at baseline ranged from 39 to 94 (M = 74.36, SD = 8.26 years) with mean level of education of 14.47 years (SD = 5.46 years). Sixty-one percent (n = 483) were female and 93.4% (n = 736) were White, with other races represented including African-American and Asian-American. Patient diagnoses included probable AD with DAT (n = 593, 75.3%), possible AD with DAT (n = 61, 7.8%), mixed AD with DAT (n = 53, 6.7%), amnestic MCI (n = 58, 7.4%), and SMI (n = 22, 2.8%). These diagnoses were included because they represent the full AD spectrum, from prodromal stages (SMI, MCI) to DAT.

2.2 Materials and Procedures

Patients completed a comprehensive evaluation at baseline admission and annual follow-up examinations, including neurological, neuropsychological, and medical testing. As part of this longitudinal protocol, patients completed the ADAS-cog. As described earlier, this instrument assesses AD-associated cognitive dysfunction across 11 subtests. Subtests can be grouped into three domains of functions: memory (immediate word recall, recognition, remembering test instructions, orientation), language (commands, naming, expressive speech, language comprehension, word finding), and praxis (construction, ideational praxis). Patients also completed the delayed

recall subtest. Raw scores for subtests are summed to create a total score, with higher scores indicating more errors and greater cognitive dysfunction. Total scores for the standard, 11-item ADAS-cog (ADAS-cog11) range from 0 (minimal impairment) to 70 (severe impairment). The delayed recall subtest has a total possible score of 10, so total ADAS-cog scores including the 11 standard subtests and the delayed recall subtest (ADAS-cog12) range from 0 to 80. Patients' most recent ADAS-cog11 and ADAS-cog12 scores were used to avoid oversampling from the preclinical and mild stages of AD by using baseline data.

For each patient, a collateral source such as a spouse or caregiver completed two questionnaires regarding the patient's ability to complete ADLs and IADLs. The Physical Self-Maintenance Scale (Lawton & Brody, 1969) measures six domains of basic ADLs, including toileting, dressing, eating, grooming, ambulation, and bathing behaviors. Patient's ability to complete these activities are scored on a 1-to-5 point scale, where a score of 1 indicates no problems completing these activities and 5 indicates severe impairment in completing these functions. Possible total scores range from 6 (no impairment) to 30 (severe impairment). In this sample, total scores ranged from 6 to 28 (M = 11.03, SD = 5.01). The Instrumental Activities of Daily Living Scale (Lawton & Brody, 1969) assesses patients' ability to independently operate a telephone, take care of shopping needs, prepare meals, perform housekeeping duties, complete laundry, manage transportation needs, take medications properly, and handle finances. For each domain, the patient's ability to complete these activities was scored on a 0-to-3, 0-to-4, or 0-to-5 Likert-type scale, where a low score indicates no impairment and a high score indicates

severe impairment in completing that activity. For select domains (i.e., food preparation, housekeeping, laundry, ability to handle finances, medications), a score of 0 could be selected if the domain was not applicable, such as if the patient was never responsible for managing finances or is not currently taking medications. Possible total scores range from 3 (no impairment) to 31 (severe impairment). In this sample, total scores ranged from 5 to 31 ($M = 21.12$, $SD = 6.84$).

Medical diagnoses were based on a consensus conference review of medical records, neuropsychological test scores, and medical evaluations. Additional information on the ADMDC protocol can be found elsewhere (Doody et al., 2005).

2.3 Data Analyses

For my first hypothesis, an IRT framework provides the best way to examine how the ADAS-cog and its subtests function along the continuum of cognitive dysfunction. IRT analyses assume unidimensionality of the latent construct (in this case, AD-associated cognitive dysfunction). Although the ADAS-cog measures different conceptual constructs (e.g., memory, language, praxis), unidimensionality reflects the tendency of those constructs to covary enough that they represent just one statistical factor. To test for unidimensionality, I conducted exploratory and confirmatory factor analyses in SPSS v. 13.0 and MPLUS (Muthen & Muthen, 2007) and examined goodness-of-fit indices.

IRT analyses were run in Multilog v. 6.3 (Thissen, 1991). Using Samejima's (1969) graded response model, I estimated parameters for each ADAS-cog item (i.e., subtest). Each item has an *a* (discrimination) parameter, which indicates how related the

item is to the latent construct (i.e., cognitive dysfunction, denoted as theta 'θ'). Theta is represented by standardized *z*-scores from -4.00 to 4.00, with a mean of 0 and a *SD* of 1, where negative scores indicate milder degrees and positive scores represent more severe degrees of cognitive impairment. Values of *a* parameters typically range from 0 to 3 (Fraley, Waller, & Brennan, 2000; Gray-Little, Williams, & Hancock, 1997; Hambleton, Swaminathan, & Rogers, 1991), with higher values indicating greater strength of the relationship between the item and theta. Each item has multiple *b* (difficulty) parameters, which indicate the severity of cognitive dysfunction that is required for certain response options on that item to be endorsed. The *a* and *b* parameters define how each item functions across the continuum of cognitive dysfunction.

To visually inspect how the items and the test function across the dimension of cognitive impairment, I created item and test characteristic curves for the ADAS-cog12. These curves can be used to predict a patient's score on a subtest or the entire ADAS-cog based on underlying severity of cognitive impairment. Alternatively, given a patient's test score, these curves can be utilized to estimate the patient's degree of cognitive impairment. The slope of the curves represents sensitivity of the test or subtest to differences in underlying cognitive dysfunction. Steeper slopes indicate a close relationship between test scores and small degrees of change in cognitive impairment, whereas flatter slopes indicate that somewhat larger degrees of change in cognitive dysfunction are necessary before the test score reflects these differences. Examination of these curves provides information about where along the dimension of the latent

construct each subtest and the ADAS-cog as a whole are most sensitive to gradations of cognitive impairment.

I also created test information curves for the ADAS-cog11 and ADAS-cog12. Examination of these curves reveals that maximum information about the latent construct (i.e., cognitive dysfunction) is provided at the peak of the curve. These two curves can be contrasted to determine how adding delayed recall to the ADAS-cog impacts its measurement precision along the spectrum of cognitive dysfunction.

For my second hypothesis, hierarchical regressions were performed to examine unique variance accounted for by the ADAS-cog delayed recall subtest in predicting functional outcomes. I first entered ADAS-cog11 then delayed recall as independent predictors of ADLs. The same method was repeated with IADLs as the dependent variable. I also performed a regression with ADAS-cog12 as the predictor of IADLs. These regressions permit examination of unique variance in functional outcomes accounted for by the delayed recall subtest, ADAS-cog11, and ADAS-cog12.

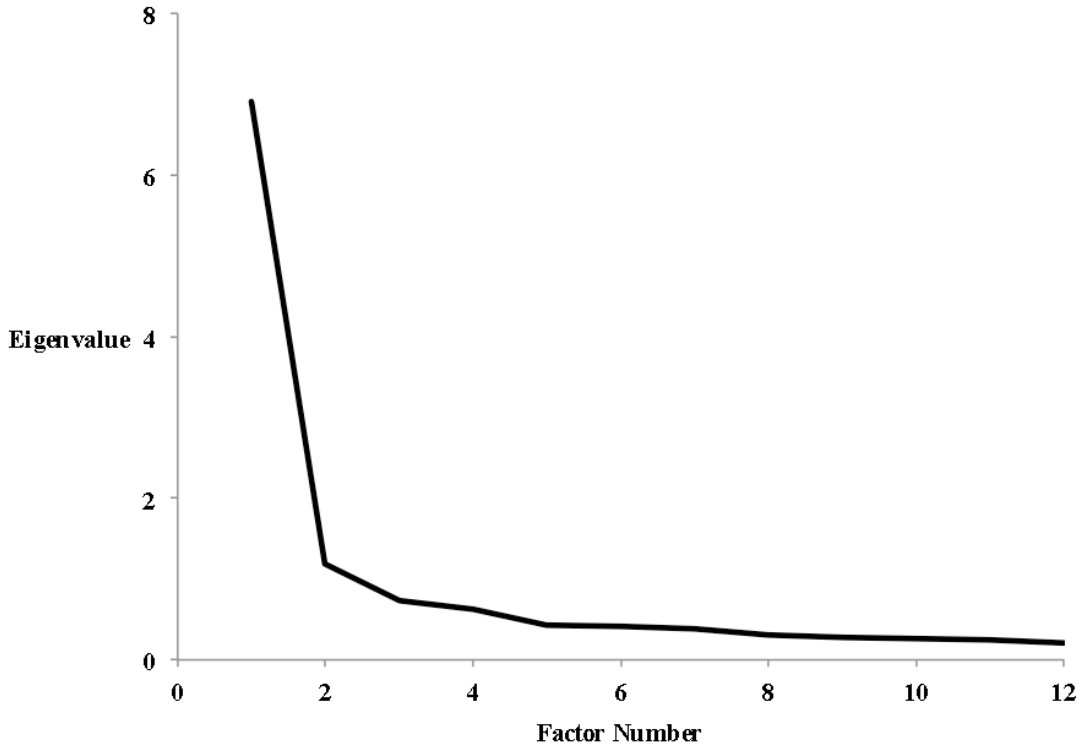
3. RESULTS

3.1 IRT Assumptions

Exploratory and confirmatory factor analyses supported the unidimensionality of the ADAS-cog12. An exploratory factor analysis using maximum likelihood estimation revealed a first factor with an eigenvalue of 6.91 and a second factor with an eigenvalue of 1.19, accounting for 54.60% and 5.95% of the variance, respectively, with all other eigenvalues less than 1.00. The ratio between the first and second eigenvalues was 5.81:1. Visual examination of the scree plot (see Figure 1) reveals a clear break between the first and second factors. Items 1 through 11 (i.e., standard ADAS-cog items) had loadings between 0.68 and 0.81 onto the first factor, whereas item 12 (i.e., delayed recall) had a weaker loading onto the first factor (0.28). A meta-analysis of factor analyses on psychological and behavioral data produced an average factor loading of 0.32 (Peterson, 2000), so the loading of item 12, although notably smaller than the factor loadings for the other items, is approximately average for behavioral data. The delayed recall subtest may be tapping into a dimension of dementia severity that is slightly different than the dimension captured by the standard ADAS-cog items, but it loads strongly enough onto this first factor to justify statistical treatment of the ADAS-cog12 as a one-factor model and to proceed with IRT analyses.

A confirmatory factor analysis using a robust weighted least squares estimation also revealed an adequate fit of the data to a one-factor model. The Tucker-Lewis index was 0.97 and the comparative fit index was 0.87. The values of these indices reach a

Figure 1. Scree plot for the ADAS-cog12



sufficient threshold to confirm the proposed model (i.e., a one-factor model). Although the comparative fix index is somewhat less than Hu and Bentler's (1999) recommended value of 0.95, this suggested cut-off point has not been well explored in data like these. Moreover, recent work has shown particular susceptibility of this fit index to being influenced by smaller factor loadings on exploratory factor analysis (Heene, Hilbert, Draxler, Ziegler, & Buhner, 2011). Therefore, these fit indices are sufficient to demonstrate an adequate fit between the data and the proposed one-factor model.

3.2 Item and Test Functioning

IRT analyses were conducted on the ADAS-cog12 items. The *a* parameters ranged from 1.65 to 2.86, with *b* parameters ranging from -3.29 to 2.59. For item 12 (delayed recall), the *a* parameter was 1.68 and *b* parameters ranged from -3.16 to -0.98.

Item characteristic curves (ICCs) enable visual analysis of how the items function across the dimension of cognitive impairment. Items have been grouped by cognitive domain (i.e., memory, language, and praxis), with curves depicting how each domain functions across the spectrum of cognitive dysfunction (see Figure 2). In a visual analysis of the curves, the part of the curve with the steepest slope indicates that the domain has the greatest discriminative power at that given range of theta. To examine the memory items in more detail, ICCs for the memory subtests are shown in Figure 3. Examination of the delayed recall ICC (see Figure 3) reveals that it is most sensitive in the mild range of the disease spectrum.

Figure 2. ADAS-cog domain curves

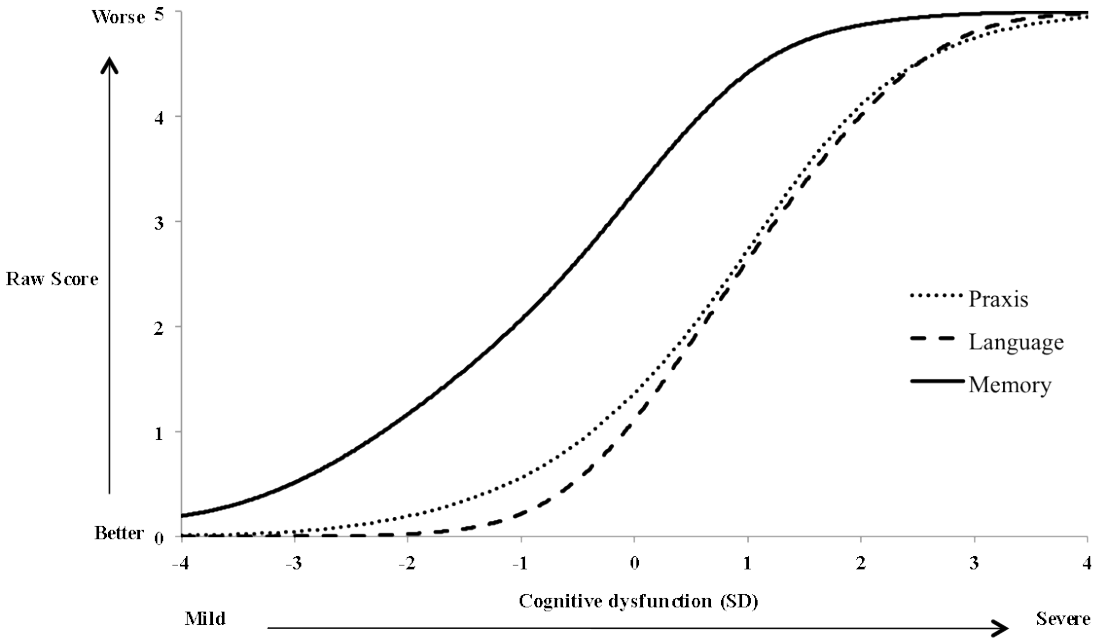


Figure 3. Item characteristic curves for ADAS-cog12 memory items

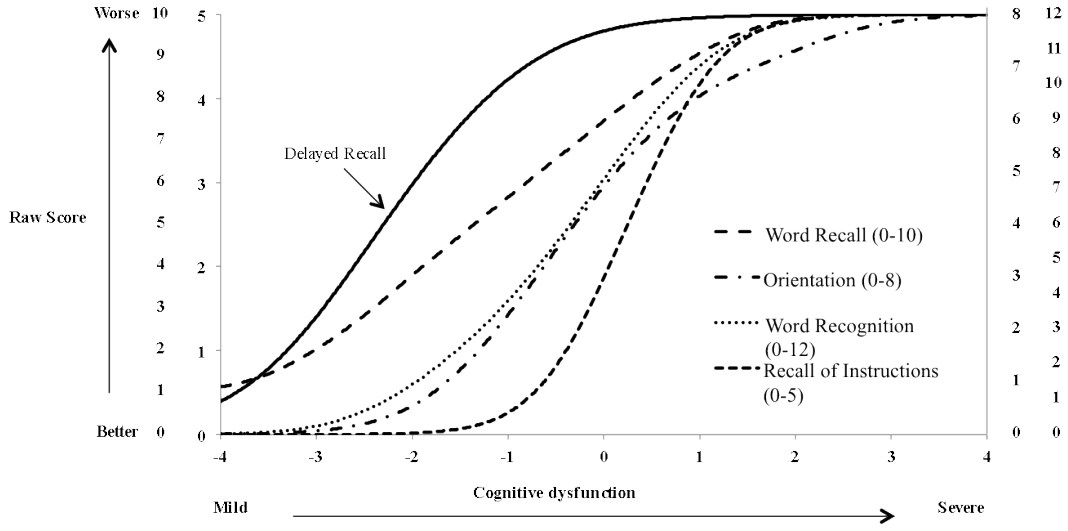


Figure 4. Test characteristic curve for the ADAS-cog12

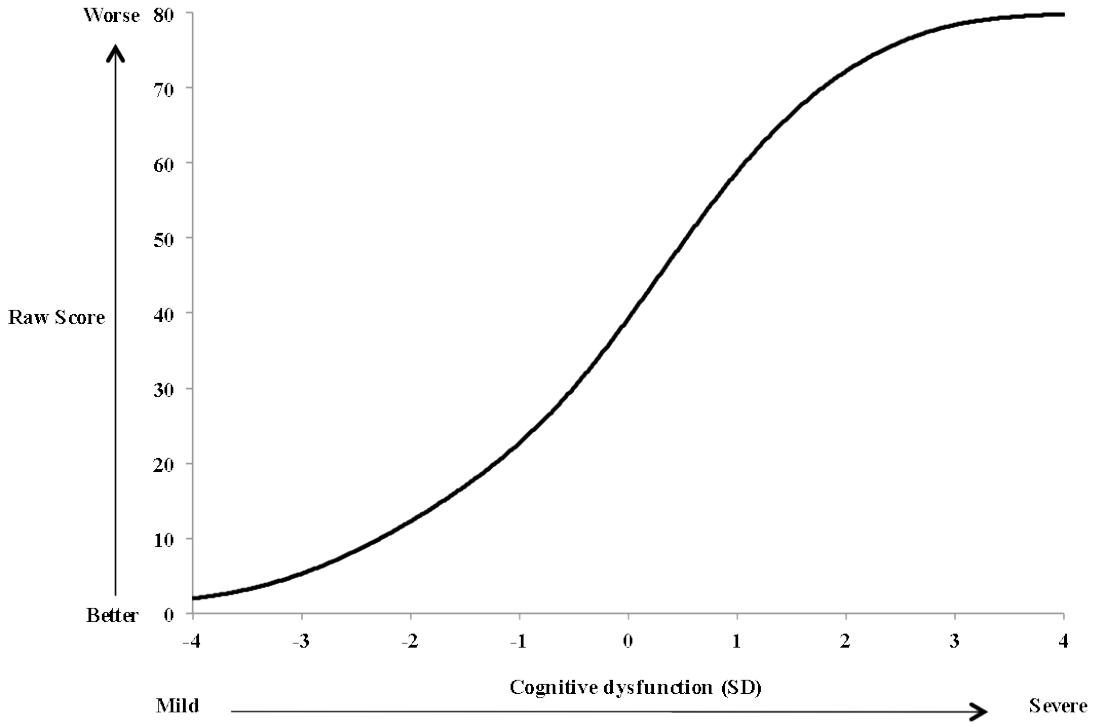
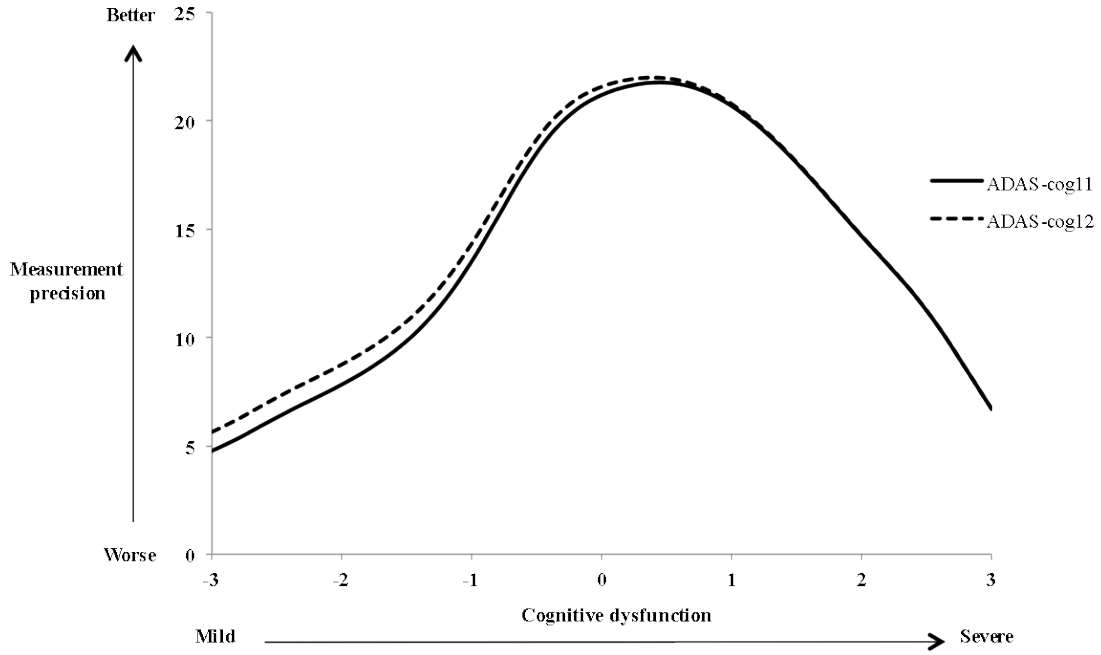


Figure 5. Test information curves for the ADAS-cog11 and ADAS-cog12



These domain curves were combined to produce a test characteristic curve for the ADAS-cog12 (see Figure 4), which shows how the test as a whole functions across the range of cognitive impairment. A test information curve for the ADAS-cog12 was derived from this test characteristic curve (see Figure 5). The information curve visually depicts where along the dimension of AD-associated cognitive impairment the test yields optimal information. Maximum information about the latent construct (i.e., cognitive dysfunction) is provided at the peak of the curve. A test information curve for the ADAS-cog11 was also plotted (see Figure 5) to contrast information gained from the standard ADAS-cog and the ADAS-cog12. This permits evaluation of how the addition of delayed recall changes the ADAS-cog's functioning across the dimension of cognitive impairment. Visual inspection indicates that the ADAS-cog12 provides more information in the relatively mild range of the AD spectrum than does the ADAS-cog11. These results confirm my first hypothesis that adding delayed recall to the ADAS-cog increases sensitivity to cognitive impairment in the mild range of AD, but in a less robust manner than expected.

3.3. ADAS-cog and Functional Impairment

Hierarchical regressions were conducted to examine the second hypothesis that adding delayed recall to the ADAS-cog will increase the ability to detect the association between cognitive dysfunction and IADLs but not basic ADLs. Hierarchical regression permits examination of unique variance accounted for by the delayed recall subtest. The ADAS-cog11 was entered first, followed by item 12 (i.e., delayed recall), with the ADL measure as the dependent variable. There was a non-significant change in R^2 with the

addition of delayed recall to the model, $p = 0.88$. The same procedures were performed again with the IADL measure as the dependent variable. There was a significant R^2 change from the first step ($R^2 = 0.39$) to the second step ($R^2 = 0.40$), $p < .01$. In this model, delayed recall accounts for an additional 1% of the variance in predicting IADLs, which is statistically significant. This confirms my hypothesis that delayed recall improves detection of the relationship between cognitive impairment and IADLs but not ADLs. Delayed recall accounts for an additional 1% of the variance, which is not a large increase but is a statistically significant gain.

Treating the delayed recall subtest as an independent predictor in addition to the ADAS-cog11 yielded a statistically significant gain in prediction of IADLs. However, in clinical and research use, the delayed recall subtest score is generally combined with the other ADAS-cog subtest scores to generate a total score on the ADAS-cog12 rather than being examined independently. To be consistent with current clinical and research use of the ADAS-cog, a regression was conducted to examine variance accounted for by the ADAS-cog12 in predicting IADLs. This regression revealed an R^2 of 0.40. Interestingly, the same amount of variance accounted for in predicting IADLs (40%) is gained whether adding delayed recall to the ADAS-cog total score (i.e., ADAS-cog12) or treating delayed recall as an independent factor and entering it into a hierarchical regression model after the standard ADAS-cog (i.e., ADAS-cog11). Collectively, these results provide evidence not only that the delayed recall subtest provides optimal information about AD-associated cognitive impairment in the mildest ranges of the disease, but it also improves the variance in IADLs for which the ADAS-cog can account.

4. DISCUSSION AND CONCLUSIONS

The results supported my expectations that delayed recall would best detect cognitive impairment in the mild ranges and that it would provide incremental information for predicting IADLs but not ADLs. The ADAS-cog12 fit a one-factor model, indicating that the subtests covary in a way that they measure one statistical factor or latent construct—in this case, a dimension of AD-associated cognitive impairment. IRT analyses revealed moderate-to-strong relationships between each ADAS-cog item and the latent construct (i.e., AD-associated cognitive impairment). Typically, *a* parameters range from 0 to 3 (Fraley et al., 2000; Gray-Little et al., 1997; Hambleton et al., 1991). In this study, the *a* parameters ranged from 1.65 to 2.86. Item 12 (delayed recall) had an *a* parameter of 1.68, which indicates a moderately strong relationship with the latent construct.

An examination of the item characteristic curves for the memory items indicate that these items have somewhat different patterns of functioning across the latent continuum (see Figure 3). Word recall provides optimal information from approximately -3.00 to 2.25 *SD*s, with the orientation and word recognition items capturing information best from approximately -1.50 to 0.50 and -2.25 to 1.00 *SD*s, respectively. Recall for test instructions provides the most information from approximately -0.75 to 1.00 *SD*s. Delayed recall functions optimally at the mildest end of the spectrum, capturing optimal information from approximately -3.00 to -1.00 *SD*s. Of all the ADAS-cog items, the delayed recall subtest is most sensitive to mild degrees of cognitive impairment. This is

consistent with prior research showing differential sensitivity of delayed recall measures in the earliest stages of AD (cf., Chen et al., 2000; Fleisher et al., 2007; Gallagher et al., 2010; Pozueta et al., 2011; Rabin et al., 2009; Rami et al., 2012; Tabert et al., 2006).

Examining how the items function across the entire dimension of AD-associated cognitive impairment imparts rich information with implications for clinical and research utilization of the ADAS-cog. Currently, though, item scores are typically combined to generate a total score, which is then examined in research or clinical practice. Therefore, it is also critical to examine how the test as a whole functions across the range of cognitive impairment. Inspecting the test characteristic and information curves for the ADAS-cog12 (see Figures 4 and 5) reveals that it captures maximal information from approximately -1.00 to 1.50 *SD*s. The ADAS-cog11 seems to be most sensitive in the same range of AD-associated cognitive impairment, but on closer contrast of the two test information curves (see Figure 5), it is clear that the ADAS-cog12 captures more information in the mild ranges than does the ADAS-cog11. This increase in information occurs between approximately -3.00 and 0.75 *SD*s, with greater incremental information toward the very mildest end of the dimension. The incremental information yielded by the ADAS-cog12 over that provided by the ADAS-cog11 is consistent with the fact that the delayed recall item provides maximal information in the mild range of cognitive impairment. By adding delayed recall to the standard 11-item ADAS-cog, the test as a whole gains information or sensitivity to cognitive impairment in the milder ranges of the disease.

This additional information about cognitive impairment also translates into improved prediction of IADLs. Delayed recall improves variance in IADLs accounted for by the ADAS-cog by 1%, whether including delayed recall in the total ADAS-cog score (i.e., ADAS-cog12) or treating it as an independent predictor in addition to the standard ADAS-cog total score. Accounting for an additional 1% of the variance in predicting IADLs may seem trivial. However, when predicting important outcomes such as ability to drive, properly handle finances, take medications, or operate an oven, it is critical to account for as much variance as possible. These functional impairments, while less disabling than ADLs, have great potential to inflict harm on oneself and others. Problems with driving may translate into increased traffic violations and motor vehicle accidents, while improper usage of medication can negatively impact one's own health and could even be lethal. Moreover, financial mismanagement could lead to depleting one's financial resources and forgetting to turn off an oven or stove creates a fire hazard. For these reasons, enhancing the ability to account for variance in IADLs is critical. Utilizing the delayed recall subtest increases the amount of variance in IADLs for which the ADAS-cog can account.

This increase in predictive ability was found for IADLs but not ADLs, which was consistent with my expectations. Delayed recall provides maximal information in the mildest range of the disease, just as adding delayed recall to the standard ADAS-cog increases information in the mild range of the dementia spectrum. Patients with preclinical and early forms of AD commonly experience declines in their IADLs (Griffith et al., 2003; Perneczky et al., 2006), but deficits in ADLs do not occur until the

moderate-to-late stages of the disease (Hermann & Gauthier, 2008; Mendez & Cummings, 2003). Therefore, it is understandable that delayed recall, both independently and when added to the ADAS-cog, improves the prediction of only those functional outcomes that decline in the earlier stages of the disease, namely IADLs but not ADLs.

In conclusion, delayed recall yields information about AD-associated cognitive impairment and strengthens the sensitivity of the ADAS-cog to information about cognitive dysfunction in the mild range of the disease. Moreover, this incremental information gained from the delayed recall subtest directly translates into increased prediction of deficits in IADLs, which are prone to declines early in the spectrum of AD. Future research should seek to extend these findings with longitudinal data and continue to explore how to further increase measurement precision and predictive power in the earliest stages of AD. As researchers continue to strive for earlier detection and better treatment of AD, delayed recall should be utilized as a sensitive measure of preclinical declines in memory and functional domains. As our understanding of prodromal symptoms of AD increases, we will move closer toward the ultimate goal of prevention.

REFERENCES

Alzheimer's Association. (2012). Alzheimer's Association report: 2012 Alzheimer's disease facts and figures. *Alzheimer's & Dementia, 8*, 131-168.

Benge, J. F., Balsis, S., Geraci, L., Massman, P. J., & Doody, R. S. (2009). How well do the ADAS-cog and its subscales measure cognitive dysfunction in Alzheimer's disease? *Dementia and Geriatric Cognitive Disorders, 28,* 63-69.

Chamberlain, S. R., Blackwell, A. D., Nathan, P. J., Hammond, G., Robbins, T. W., Hodges, J. R., … Sahakian, B. J. (2011). Differential cognitive deterioration in dementia: A two year longitudinal study. *Journal of Alzheimer's Disease, 24,* 125-136.

Chen, P., Ratcliff, G., Belle, S. H., Cauley, J. A., DeKosky, S. T., & Ganguli, M. (2000). Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented. *Neurology, 55*, 1847-1853.

Doody, R. S., Pavlik, V., Massman, P., Kenan, M., Yeh, S., Powell, S., … Chan, W. (2005). Changing patient characteristics and survival experience in an Alzheimer's disease center patient cohort. *Dementia and Geriatric Cognitive Disorders, 20,* 198-208.

Fleisher, A. S., Sowell, B. B., Taylor, C., Gamst, A. C., Petersen, R. C., Thal., L. J., & Alzheimer's Disease Cooperative Study. (2007). Clinical predictors of progression to Alzheimer disease in amnestic mild cognitive impairment. *Neurology, 68,* 1588-1595.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78,* 350-365.

Gallagher, D., Mhaolain, A. N., Coen, R., Walsh, C., Kilroy, D., Belinski, K., … Lawlor, B. A. (2010). Detecting prodromal Alzheimer's disease in mild cognitive impairment: Utility of the CAMCOG and other neuropsychological predictors. *International Journal of Geriatric Psychiatry, 25,* 1280-1287.

Gauthier, S., Cummings, J., Ballard, C., Brodaty, H., Grossberg, G., Robert, P., & Lyketsos, C. (2010). Management of behavioral problems in Alzheimer's disease. *International Psychogeriatrics, 22,* 346-372.

Gold, D. A. (2012). An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. Journal of Clinical and *Experimental Neuropsychology, 34,* 11-34.

Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 23,* 443-451.

Griffith, H. R., Belue, K., Sicola, A., Krzywanski, S., Zamrini, E., Harrell, L., & Marson, D. C. (2003). Impaired financial abilities in mild cognitive impairment: A direct assessment approach. *Neurology, 60*, 449-457.

Grundman, M., Petersen, R. C., Ferris, S. H., Thomas, R. G., Aisen, P. S., Bennett, D. A., … Alzheimer's Disease Cooperative Study (2004). Mild cognitive

impairment can be distinguished from Alzheimer disease and normal aging for clinical trials. *Archives of Neurology, 61,* 59-66.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. Newbury Park, CA: Sage Publications.

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Buhner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods, 16*, 319-336.

Herrmann, N., & Gauthier, S. (2008). Diagnosis and treatment of dementia: 6. Management of severe Alzheimer disease. *Canadian Medical Association Journal, 179*, 1279-1287.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Irizarry, M. C., Webb, D. J., Bains, C., Barrett, S. J., Lai, R. Y., Laroche, J. P., … Weil, J. G. (2008). Predictors of placebo group decline in the Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) in 24 week clinical trials of Alzheimer's disease. *Journal of Alzheimer's Disease, 14,* 301-311.

Ito, K., Corrigan, B., Zhao, Q., French, J., Miller, R., Soares, H., … Fullerton, T. (2011). Disease progression model for cognitive deterioration from Alzheimer's Disease Neuroimaging Initiative database. *Alzheimer's Dementia, 7,* 151-160.

Jessen, F., Wiese, B., Bachmann, C., Eifflaender-Gorfer, S., Haller, F., Kolsch, H., … German Study on Aging, Cognition and Dementia in Primary Care Patients Study Group. (2010). Prediction of dementia by subjective memory impairment: Effects of severity and temporal association with cognitive impairment. *Archives of General Psychiatry, 67,* 414-422.

Kirk, A. (2007). Target symptoms and outcome measures: *Cognition. Canadian Journal of Neurological Sciences, 34*(S1), S42-26.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist, 9,* 179-186.

Llano, D. A., Laforet, G., & Devanarayan, V. (2011). Derivation of a new ADAS-cog composite using tree-based multivariate analysis: Prediction of conversion from mild cognitive impairment to Alzheimer disease. *Alzheimer's Disease and Associated Disorders, 25,* 73-84.

Luck, T., Luppa, M., Wiese, B., Maier, W., van den Bussche, H., Eisele, M., … AgeCoDe Study Group. (2012). Prediction of incident dementia: Impact of impairment in instrumental activities of daily living and mild cognitive impairment- Results from the German Study on Ageing, Cognition, and Dementia in Primary Care Patients. *American Journal of Geriatric Psychiatry,* Advance online publication. doi 10.1097/JGP.0b013e31825c09bc

Mendez, M. F., & Cummings, J. L. (2003). Alzheimer's disease. In Dementia: A clinical approach (3rd ed.). Philadelphia, PA: Butterworth-Heinemann.

Mohs, R. C., Knopman, D., Petersen, R. C., Ferris, S. H., Ernesto, C., Grundman, M., … Thal, L. J. (1997). Development of cognitive instruments for use in clinical trials of antidementia drugs: Additions to the Alzheimer's Disease Assessment Scale that broaden its scope. *Alzheimer's Disease and Associated Disorders, 11*(Supplement 2), S13-S21.

Mohs, R. C., Rosen, W. G., & Davis, K. L. (1983). The Alzheimer's Disease Assessment Scale: An instrument for assessing treatment efficacy. *Psychopharmacology Bulletin, 19*, 448-450.

Morris, J. C., Storandt, M., Miller, J. P., McKeel, D. W., Price, J. L., Rubin, E. H., & Berg, L. (2001). Mild cognitive impairment represents early-stage Alzheimer disease. *Archives of Neurology, 58*, 397-405.

Muthen, L. K., & Muthen, B. O. (2007). MPLUS User's Guide (Ed. 5). Los Angeles, CA: Muthen & Muthen.

National Plan to Address Alzheimer's Disease (2012). aspe.hhs.gov/daltcp/napa/NatlPlan.shtml (Accessed 19 May 2012)

Perneczky, R., Pohl, C., Sorg, C., Hartmann, J., Komossa, K., Alexopoulos, P., … Kurz, A. (2006). Complex activities of daily living in mild cognitive impairment: Conceptual and diagnostic issues. *Age and Ageing, 35,* 240-245.

Petersen, R. C., Aisen, P. S., Beckett, L. A., Donohue, M. C., Gamst, A. C., Harvey, D. J., … Weiner, M. W. (2010). Alzheimer's Disease Neuroimagining Initiative (ADNI): Clinical characterization. *Neurology, 74,* 201-209.

Peterson, R. A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters, 11*, 261-275.

Pozueta, A., Rodriguez-Rodriguez, E., Vazquez-Higuera, J. L., Mateo, I., Sanchez-Juan, P., Gonzalez-Perez, S., … Combarros, O. (2011). Detection of early Alzheimer's disease in MCI patients by the combination of MMSE and an episodic memory test. *BMC Neurology, 11,* doi:10.1186/1471-2377-11-78

Rabin, L. A., Pare, N., Saykin, A. J., Brown, M. J., Wishart, H. A., Flashman, L. A., & Santulli, R. B. (2009). Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. *Neuropsychology, Development, and Cognition. Section B: Aging, Neuropsychology, and Cognition, 16,* 357-376.

Rami, L., Sole-Padulles, C., Fortea, J., Bosch, B., Llado, A., Antonell, A., … Molinuevo, J. L. (2012). Applying the new research diagnostic criteria: MRI finding and neuropsychological correlations of prodromal AD. *International Journal of Geriatric Psychiatry, 27,* 127-134.

Riepe, M. W., Janetzky, W., & Lemming, O. M. (2011). Measuring therapeutic efficacy in patients with Alzheimer's disease: Role of instruments. *Dementia and Geriatric Cognitive Disorders, 31,* 233-238.

Rockwood, K., Fay, S., Gorman, M., Carver, D., & Graham, J. E. (2007). The clinical meaningfulness of ADAS-Cog changes in Alzheimer's disease patients treated with donepezil in an open-label trial. *BMC Neurology, 7,* doi:10.1186/1471-2377-7-26

Rosen, W. G., Mohs, R. C., & Davis, K. L. (1984). A new rating scale for Alzheimer's disease. *American Journal of Psychiatry, 141,* 1356-1364.

Rozzini, L., Vicini Chilovi, B., Conti, M., Bertoletti, E., Delrio, I., Trabucchi, M., & Padovani, A. (2007). Conversion of amnestic mild cognitive impairment to dementia of Alzheimer type is independent to memory deterioration. *International Journal of Geriatric Psychiatry, 22*, 1217-1222.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores (Psychometric Monograph No. 17). Richmond, VA: Psychometric Society.

Sano, M., Raman, R., Emond, J., Thomas, R. G., Petersen, R., Schneider, L. S., & Aisen, P. S. (2011). Adding delayed recall to the Alzheimer Disease Assessment Scale is useful in studies of mild cognitive impairment but not Alzheimer disease. *Alzheimer's Disease and Associated Disorders, 25,* 122-127.

Scheef, L., Spottke, A., Daerr, M., Joe, A., Striepens, N., Kolsch, H., … Jessen, F. (2012). Glucose metabolism, gray matter structure, and memory decline in subjective memory impairment. *Neurology, 79,* advance online publication.

Sevigny, J. J., Peng, Y., Liu, L., & Lines, C. R. (2010). Item analysis of ADAS-Cog: Effect of baseline cognitive impairment in a clinical AD trial. *American Journal of Alzheimer's Disease and Other Dementias, 25,* 119-124.

Suh, G.-H., Ju, Y.-S., Yeon, B. K., & Shah, A. (2004). A longitudinal study of Alzheimer's disease: Rates of cognitive and functional decline. *International Journal of Geriatric Psychiatry, 19,* 817-824.

Tabert, M. H., Manly, J. J., Liu, X., Pelton, G. H., Rosenblum, S., Jacobs, M., … Davanand, D. P. (2006). Neuropsychological prediction of conversion to Alzheimer disease in patients with mild cognitive impairment. *Archives of General Psychiatry, 63*, 916-924.

Thissen, D. (1991). Multilog User's Guide, version 6.3. Mooresville, IN: Scientific Software.