Predicting Conversion to Dementia of the Alzheimer's Type in a Healthy Control Sample: The Power of Errors in Stroop Color Naming

David A. Balota and Chi-Shing Tse
Washington University

Keith A. Hutchison
Montana State University

Daniel H. Spieler
Georgia Institute of Technology

Janet M. Duchek and John C. Morris
Washington University

In the present study, we investigated which cognitive functions in older adults at Time A are predictive of conversion to dementia of the Alzheimer's type (DAT) at Time B. Forty-seven healthy individuals were initially tested in 1992–1994 on a trial-by-trial computerized Stroop task along with a battery of psychometric measures that tap general knowledge, declarative memory, visual–spatial processing, and processing speed. Twelve of these individuals subsequently developed DAT. The errors on the color incongruent trials (along with the difference between congruent and incongruent trials) and changes in the reaction time distributions were the strongest predictors of conversion to DAT, consistent with recent arguments regarding the sensitivity of these measures. Notably in the psychometric measures, there was little evidence of a difference in declarative memory between converters and nonconverters, but there was some evidence of changes in visual-spatial processing. Discussion focuses on the accumulating evidence suggesting a role of attentional control mechanisms as an early marker for the transition from healthy cognitive aging to DAT.

Keywords: Alzheimer's disease, attentional control, conversion, dementia, Stroop

An important goal in research investigating healthy aging and early stage Alzheimer's disease (AD) is to develop predictors of who in an apparently healthy control sample at Time A is destined to convert to a diagnosis of dementia of the Alzheimer's type (DAT) at Time B. This is important because (a) predictive markers for the disease will afford the earliest intervention for individuals through clinical therapies that target the disease mechanisms and (b) if there is some subset of the population that is developing the disease before diagnosis, then researchers' understanding of what constitutes healthy aging may be contaminated by undetected disease pathology (see, e.g., Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003; Sliwinski, Lipton, Buschke, & Stewart, 1996). Neuropathological studies of nondemented older adults consistently show that ~25% or more meet neuropathological criteria for AD, thus providing presumptive evidence that there is a preclinical stage for AD that begins in the brain years before the appearance of symptoms (Price & Morris, 1999).

There have been a number of important preclinical biomarkers already identified including: (a) the presence of ApolipoproteinE4 allele (ApoE4; e.g., Corder et al., 1993), (b) volumetric measures of medial temporal lobe via imaging techniques (e.g., see review by Twamley, Ropacki, & Bondi, 2006), and (c) the ratio of tau and Aβ42 from the cerebral spinal fluid (e.g., Fagan et al., 2007). Finally, certain personality styles, such as individuals high in neuroticism and low in conscientiousness, are also predictors of conversion from healthy aging to DAT (e.g., Balsis, Carpenter, & Storandt, 2005; Wilson, Schneider, Arnold, Bienias, & Bennett, 2007).

In addition to biomarkers, there has been work investigating the cognitive predictors of conversion to AD. There are two basic approaches used in this research. First, individuals are tracked longitudinally on cognitive measures to determine if the slope of performance on specific cognitive measures predicts conversion (see Albert et al., 2007; Storandt et al., 2006). Second, one can measure predictor variables at Time A and track individuals clinically to determine if a cognitive measure at Time A will eventually predict the clinical manifestation of the disease (e.g., Albert, Moss, Tanzi, & Jones, 2001). Both approaches have advantages. In the current study, we relied primarily on the second, preclinical prediction approach.

In order to increase the likelihood of progression to AD, researchers often study individuals who have a mild memory problem or who are diagnosed with mild cognitive impairment (MCI; see Petersen et al.,

2001). For example, Albert et al. (2001) studied 123 individuals with mild memory difficulty over a 3-year period and found that 23 individuals converted to DAT. Both memory and executive measures were useful in predicting who would or would not progress. Likewise, Sarazin et al. (2007) investigated the cognitive predictors for conversion to DAT in a group of individuals with MCI and found that of the 251 individuals at baseline, 59 converted to DAT. The most sensitive measure in this study was the free and cued selective reminding test. Albert et al. (2007) investigated the predictive power of the rate of decline across a 4-year interval for various cognitive domains in individuals with MCI. The rate of decline in the episodic memory measures predicted conversion to DAT; however, performance in both executive function and general knowledge measures was also significantly lower at baseline for converters, relative to nonconverters.

Twamley et al. (2006) have recently provided a review of 91 studies in which researchers have investigated preclinical AD in individuals without MCI and/or memory impairments. They argued that it is useful to exclude such studies because, by definition, these individuals will already have memory impairment, and indeed there is evidence that these individuals may already have early stage AD (Morris et al., 2001; Storandt, Grant, Miller, & Morris, 2002; Storandt, Grant, Miller, & Morris, 2006). Hence, Twamley et al. argued that it is also important to understand the predictors before any memory impairment is observed. One of the intriguing findings in the Twamley et al.'s meta-analysis is that

> attention, although not as commonly assessed as learning and memory in preclinical AD, is even more consistently associated with later development of AD. Only 10% of the longitudinal case-control studies measured attention, but of those 100% found that attention performance discriminated cases vs. controls. (p. 709)

There has been considerable cross-sectional evidence indicating that individuals with early stage DAT have deficits in attentional control measures (see Balota & Faust, 2001; Perry & Hodges, 1999, for reviews). These deficits have been widely observed in early stage DAT across aspects of basic visual attention (Faust & Balota, 1997), selective attention (Balota & Duchek, 1991; Castel, Balota, Hutchison, Logan, & Yap, 2007; Spieler, Balota, & Faust, 1996), divided attention (Baddeley, Baddeley, Bucks, & Wilcock, 2001; Duchek & Balota, 2005), and task switching (Baddeley, 2002; Belleville, Bherer, Lepage, Chertkow, & Gauthier, 2008). Importantly, because attention is critical to memory performance, these breakdowns likely contribute to the observed problems in memory in these individuals (see Balota, Burgess, Cortese, & Adams, 2002).

Because of the relative paucity of attentional measures in the past conversion studies of healthy individuals (see Twamley et al., 2006), and the accumulating evidence that attentional control measures do discriminate healthy older controls from the earliest detectable forms of DAT (see Faust & Balota, 2007), in the current study we report a retrospective analysis on one of the classic attentional selection tasks, Stroop color naming. We were particularly interested in this task for the following reasons: (a) it is the most well studied attentional selection task available (see MacLeod, 1992); (b) there is evidence that Stroop performance changes in the earliest stages of the disease (Spieler et al., 1996); and (c) recently, there is a task switching version of the Stroop task

that is particularly sensitive to cognitive decline in normal-functioning older individuals (see Fine et al., 2008).

It is important to note that there are two ways of administrating a Stroop task. The most common form is the standard neuropsychological Stroop card-reading task, whereas the second approach is to implement a computerized version of the task. Interestingly, the card-reading version of the Stroop task has already been reported in a number of previous conversion studies, although there appear to be some inconsistencies in the results from these studies. For example, Sarazin et al. (2007) did not find Stroop performance to be a particularly useful discriminator between converters and nonconverters in a group of MCI individuals. In contrast, Albert et al. (2007) found a large baseline difference in their attentional factor score (including Stroop) between converters versus nonconverters, although the specific estimates from Stroop were not reported.

The discrepancy in past studies of Stroop performance may question the utility of this task as a particularly sensitive measure for discriminating between converters from nonconverters. However, in all of these studies, the standard card-reading version of the Stroop task was used. This task involves three conditions on three different cards: (a) reading words printed in black, (b) naming the colors of rows of Xs, and (c) naming the color of incongruent words (e.g., say *blue* to the word *red* printed in blue). Perlstein, Carter, Barch, and Baird (1998) have critically reviewed this version of the Stroop task and pointed out that there is some inconsistency with this measure even in studies of schizophrenia. This runs counter to the common view that schizophrenia involves a deficit in attentional control (e.g., Servan-Schreiber & Cohen, 1998). Perlstein et al. noted that part of the problem with this version of the Stroop task is that researchers often simply use the time taken to complete the incongruent condition or number of items completed within a given time period as the dependent measure. It is not surprising that the card-reading version of the Stroop task is used in these studies because it is easy to use, quick, and straightforward to interpret.

In order to explore the utility of attentional selection tasks in discriminating healthy controls from schizophrenics, Perlstein et al. (1998) reported evidence from both the classic card-reading version of this task and the trial-by-trial computerized version of the task and found a number of noteworthy patterns. Specifically, when measured by an appropriate control condition in the card-reading version of the task, individuals with schizophrenia actually produced slightly smaller Stroop effects than did healthy controls, which is surprising in the context of the considerable evidence of attentional deficits in schizophrenia. In contrast, the trial-by-trial computerized version produced the expected larger Stroop effects for the same schizophrenic individuals compared to the healthy controls. Importantly, the authors noticed a particularly high increase in error rates in the incongruent condition, which they noted is not often measured in the card-reading version of the task. The error rates on the trial-by-trial version are particularly intriguing because Spieler et al. (1996) reported that there was a considerable increase in the error rates for the DAT individuals.

In the present study, we explored the utility of the trial-by-trial computerized version of the Stroop task to discriminate converters from nonconverters. Hence, we retrospectively examined the performance of 47 healthy older individuals who participated in the Stroop task, along with various psychometric tests from a study by

Spieler et al. (1996).[1] Since then, a total of 12 individuals converted to DAT. The question addressed is whether the trial-by-trial version of Stroop performance can be a useful predictor of subsequent conversion in a healthy control sample. In addition to reaction time and error rates, we also examined the utility of changes in the reaction time distributions to discriminate converters from nonconverters. In the original Spieler et al. study, there was a considerable increase in the tail of the reaction time distribution in healthy aging, which may have reflected in part the inclusion of individuals who are likely to convert. Finally, we also report evidence from psychometric tests that are available on these individuals in order to determine the extent to which other more traditional cognitive measures are useful in predicting conversion.

Method

Participants

All 47 healthy older adults were recruited through Washington University Alzheimer's Disease Research Center. Disorders potentially affecting cognitive functioning (e.g., depression) were exclusionary. The inclusion and exclusion criteria for DAT are consistent with the criteria for "probable AD" of National Institute of Neurological and Communications Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984). The presence of dementia was assessed according to the Washington University Clinical Dementia Rating (CDR) scale (Morris, 1993; Morris, McKeel, Fulling, Torack, & Berg, 1988), with CDR 0, 0.5, 1, 2, and 3 representing no dementia, very mild dementia, mild dementia, moderate dementia, and severe dementia, respectively. The CDR is based on a 90-min clinical interview that both assesses the participants and obtains information from their family members. This interview assesses changes in participants' cognitive and functional abilities in the areas of memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care relative to previous behavior. The CDR score for each participant at baseline and at each annual assessment thereafter is made without reference to the psychometric performance of the individual. The recruitment and assessment methods permit the diagnosis of DAT in individuals who elsewhere may be characterized as MCI as previously described (Berg et al., 1998; Morris et al., 2001). Of course, it is important to recognize the difference between the clinical diagnosis of dementia of the Alzheimer's type, which is captured by the CDR, and the neuropathological confirmation of the presence of Alzheimer's disease at autopsy. In this light it is important to note that both the reliability of the CDR (Burke et al., 1988) and the validity of the diagnosis based upon autopsy by this research team have been excellent (93% accuracy), including individuals diagnosed with DAT in the very mild CDR 0.5 stage (Berg et al., 1998; Storandt et al., 2006).

These 47 participants all were diagnosed as CDR 0 (i.e., exhibiting no signs of dementia) when they participated in the Stroop task between 1992 and 1994. Based on their most recent clinical evaluation, 35 remained nondemented (CDR 0, nonconverters), whereas 12 have converted to a CDR of 0.5 or greater at the last CDR (3 ultimately converted to CDR 0.5, 3 to CDR 1.0, 5 to CDR 2.0, and 1 to CDR 3.0). Because of the small number of converters

within each CDR level, we have collapsed across this variable in the analyses presented below.

Table 1 displays the demographic information for the converters and the nonconverters. Of course, it is important to ensure that there are no other variables that differ between converters and nonconverters that might bias the results. Indeed, as shown in Table 1, there are a number of variables that are different between converters and nonconverters. As expected the converters were more likely ApoE4 positive (via blood for genotyping), compared to the nonconverters, and were slightly more likely to be female, but neither of these differences was significant. Because the converters were older at initial entry into the project and also at the last CDR rating, we included initial entry age as a covariate in all analyses reported below. Importantly, there were no differences in the mean number of CDR ratings and the mean number of years that the CDR has been given to the converters and nonconverters. Hence, both converters and nonconverters were tracked quite consistently for changes in cognitive performance. The nonzero CDR rating for the nonconverters simply reflects rare instances where the CDR was determined to be 0.5 at one annual assessment but 0 at subsequent assessments.[2]

Materials and Procedures

The experiment involved four color names (*red*, *blue*, *green*, and *yellow*) and four neutral words (*bad*, *poor*, *deep*, and *legal*) and consisted of a word-reading block and a color-naming block. Each block of trials contained 36 congruent, 36 incongruent, and 32 neutral trials. In the congruent condition, each of the four color names appeared nine times in its corresponding color. In the incongruent condition, each of the four color names appeared three times in each of the three nonmatching colors. In the neutral trials, each of the four neutral words appeared twice in each of the four colors. The task order (color naming or word reading) was counterbalanced across participants within each group.

To familiarize the participants with the task and screen out any potential color-blind individuals, we first showed participants examples of the colors and words to be used in the experiment. Each color appeared as a patch, with the color name printed underneath the patch, and the eight words used in the experiment were displayed. Participants were presented with a block of 16 practice trials at the beginning of each color or word-reading block.

Each trial began with a fixation stimulus, $+++$, displayed for 700 ms. The screen was then blank for 50 ms, and then the

---

[1] It is noteworthy that there were actually 50 healthy control individuals in the original Spieler et al. (1996) study. However, we only report analyses on 47 of these individuals because two of these individuals were later relabeled as CDR 0.5 or 1, and one individual did not have psychometric testing completed. Hence, to be conservative we only report the results from the 47 healthy control individuals with psychometric testing in the original study.

[2] It is possible that vascular disease may have also contributed to the conversion rates in this sample. Hence, we examined whether there were any reliable differences between converters and nonconverters in the percentage of stroke, high blood pressure, circulation problems, diabetes, or general heart problems. None of these measures were reliable based on a Fisher's exact test. Of course, one needs to be cautious here because of the relatively small number of participants in each group.

Table 1
Demographic Characteristics as a Function of Participant Group

Characteristic	Nonconverter	Converter	Significant difference	d
N	35	12		
% female	51%	67%	No, $\chi^2(1) = 0.84$, *ns*	
% with at least one ApoE4 allele	23%	50%	No, $\chi^2(1) = 3.08$, $p = .08$	
Mean age at 1993	77.30	81.69	No, $t(45) = 1.44$, *ns*	0.21
Mean years of education	14.97	14.83	No, $t(45) = 0.11$, *ns*	0.02
% dropout	74% (26/35)	75% (9/12)	No, $\chi^2(1) = 0.002$, *ns*	
Due to attrition	12% (3/26)	11% (1/9)		
Due to death	88% (23/26)	89% (8/9)		
Mean death date	1999.26	2002.88	Yes, $t(29) = 2.40$, $p < .05$	0.44
Mean ages at the last CDR rating	85.00	91.75	Yes, $t(45) = 2.44$, $p < .05$	0.36
Mean last CDR rating	0	1.46	Yes, $t(45) = 10.88$, $p < .01$	1.6
Mean number of CDR ratings since 1993	6.69	7.42	No, $t(45) = 0.46$, *ns*	0.07
Mean CDR rating since 1993	0.017	0.639	Yes, $t(45) = 6.27$, $p < .01$	0.92
Mean number of years of follow-up CDR ratings since 1993	7.63	10.08	No, $t(45) = 1.50$, *ns*	0.22

Note. The data for "% with at least one ApoE4 allele" were based on 12 converters and 31 nonconverters due to the missing information for the remaining 4 participants. The *d* column indicates the Cohen's *d* for the analyses. CDR = Washington University Clinical Dementia Rating.

stimulus appeared. The stimulus was displayed on a black background and remained on the screen until the participant responded. Once the voice-operated relay, which measured voice onset latency with millisecond resolution, was triggered, the experimenter pressed one of three keys to code the response. Responses were coded as (a) correct, (b) nonintrusion errors (i.e., stutters, false starts, or other noises that triggered the voice key), or (c) an intrusion error (i.e., participant named the word or other color name in the color-naming block).

Psychometric Tests

In addition to the Stroop task, all participants also received a series of psychometric tests designed to assess various aspects of cognitive functioning, including language, memory, psychomotor performance, and intelligence (see Hill, Storandt, & LaBarge, 1992, for a full description). Memory was assessed with logical memory, mental control, forward and backward digit span, and associate memory subtests from the Wechsler Memory Scale (Wechsler, 1987; Wechsler & Stone, 1973). General intelligence was assessed with Block Design, Digit Symbol, and Information subtests of the Wechsler Adult Intelligence Scale—Revised (Wechsler, 1955, 1997). Visual perceptual–motor ability was assessed with Crossing Off (Botwinick & Storandt, 1973) and Trail Making A and B (Armitage, 1945). The spatial ability was assessed with Benton Copy (Form C and D, Benton, 1963). Semantic–lexical retrieval was assessed with Word Fluency S-P (Thurstone & Thurstone, 1949) and Boston Naming tests (Kaplan, Goodglass, & Weintraub, 1983). All psychometric tests are scored such that higher scores indicate better performance except Trail Making A and B tasks, where higher scores indicate poorer performance.

Results

Stroop Performance

Reaction times from trials with incorrect responses were first excluded. Correct reaction times that were faster than 200 ms or three standard deviations below the mean (due to anticipations) and reaction times that were slower than 4,000 ms or three standard deviations above the mean (possibly due to lapses of attention or voice-key failures) for each condition were also excluded. This screening procedure eliminated 1.5% of correct responses.

Because there is little Stroop effect in the word-reading block, and indeed there were no differences between converters and nonconverters in the word-reading trials (all *F*s < 2.43, *p*s > .12, with all Cohen's, 1988, effect size *d* measure < 0.05), we focus only on the color-naming block. (Unless otherwise indicated, all significant effects have *p*s < .05, two-tailed.)

Table 2 presents the mean color-naming reaction times, *z* scored (standardized) reaction times, and errors for converters and nonconverters. Although in the predicted direction, the reaction times were not significantly different between the converters and nonconverters. This was also the case for the *z* score measure, which controls for overall baseline reaction time differences by transforming each reaction time to a *z* score based on each participant's mean and standard deviation (see Faust, Balota, Spieler, & Ferraro, 1999). More importantly, as shown in Figure 1, the error rates appear to be especially sensitive to DAT conversion, consistent with the arguments by Perlstein et al. (1998) and the original results from Spieler et al. (1996). Importantly, the converters showed significantly higher error rates in incongruent color trials (by 9.42%, *p* = .038) and larger Stroop effects in errors (by 9.72%, *p* = .027) than the nonconverters.

We also addressed the role of ApoE4 status for modulating the effects of conversion. Specifically, it is possible that the present effects were simply due to the presence of more ApoE4+ individuals within our conversion group compared to the nonconversion group (see Table 1). Hence, we ran the same analyses as described above with ApoE4 status as a covariate. The results did not change: The converters produced both higher error rates in the incongruent trials and larger Stroop effects in errors (both *p*s < .05, *d*s > 0.10), with ApoE4 status controlled.

In addition to the standard indicants of Stroop performance, Spieler et al. (1996) also provided evidence that reaction time distribution analyses were particularly compelling in understanding age-related

Table 2
Means for Color-Naming Reaction Time (in Milliseconds) and z Scored (Standardized) Reaction Times and Color-Naming Errors (in Percentages) as a Function of Participant Group and Condition

Measure	Nonconverter		Converter		Difference	$F(1, 44)$	η_p^2
	M	SD	M	SD			
Mean reaction time							
Incongruent	1,134	215	1,224	302	90	0.59	0.01
Neutral	964	169	1,038	242	74	0.47	0.01
Congruent	878	222	901	205	23	0.00	0.00
Stroop effect	256	91	323	138	67	2.92^	0.06
Mean zRT							
Incongruent	1.31	0.20	1.30	0.16	−0.01	0.02	0.00
Neutral	0.60	0.24	0.64	0.31	0.04	0.00	0.00
Congruent	0.17	0.26	0.16	0.31	−0.01	0.01	0.00
Stroop effect	1.14	0.39	1.14	0.27	0.00	0.02	0.00
Mean errors							
Incongruent	7.94	5.92	17.36	20.96	9.42	4.56*	0.09
Neutral	3.57	5.68	4.43	7.59	0.86	0.01	0.00
Congruent	2.14	2.53	1.85	2.47	−0.29	0.03	0.00
Stroop effect	5.79	5.93	15.51	19.80	9.72	5.24*	0.11

Note. The statistical significances labeled in the *F* column were based on an analysis of covariance, with participants' age at 1993 being treated as a covariate. The patterns were qualitatively identical when the mean age at the last Washington University Clinical Dementia Ratings was instead treated as a covariate. The η_p^2 column indicates the effect size η_p^2 for the analyses.
^ $p < .10$.   * $p < .05$.

and DAT related changes. Hence, we investigated the characteristics of the reaction time distributions for the converters and nonconverters in a number of different ways. First, we conducted an ex-Gaussian analysis. Specifically, we fit the empirical reaction time distributions of individual participants to an ex-Gaussian distribution, which provides estimates of three-parameters reflecting a Gaussian component (μ and σ) and an estimate of an exponential component (τ, i.e., the tail). The algebraic sum of μ and τ is the mean of the fitted ex-Gaussian distribution (see Balota, Yap, Cortese, & Watson, 2008, for a recent discussion of this procedure). (Only the data from 2 nonconverters were not fit by the ex-Gaussian function.) Interestingly, the results from this analysis indicated that there was a reliable difference in the exponential–tail of the distribution (i.e., τ) between converters ($M = 245$ ms) and nonconverters ($M = 160$ ms), $t(43) = 2.53$, $p = .016$, $d = 0.38$, but not for μ (795 vs. 799) or σ (118 vs. 129), both ts < 1.00, $d < 0.12$. This suggests that the converters produced a more skewed reaction time distribution compared to the nonconverters. It is noteworthy that the difference in τ was also observed if one
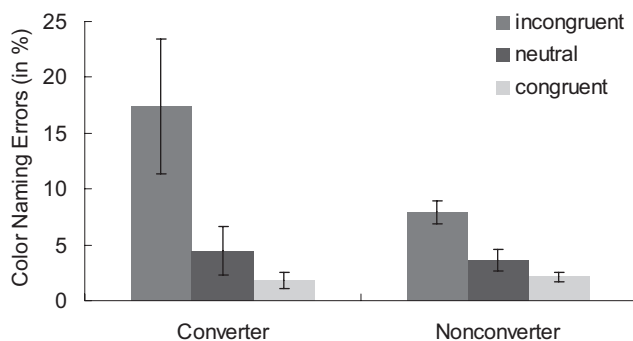
only considers the incongruent condition ($p < .03$, $d = 0.12$), but one must be cautious here because of the number of observations ($N = 36$ possible) is relatively low for fitting the ex-Gaussian within a condition.

In addition to the ex-Gaussian analyses, we also display performance as a function of the Vincentiles. Here for each participant we rank ordered their reaction times within a condition and plot the means of six ordered bins. Figure 2 displays the Stroop effect (incongruent minus congruent) for the converters and nonconverters. As shown in Figure 2, the converters are displaying a larger Stroop effect than the nonconverters, and this difference is largest in the last bin. Indeed this difference in the last bin was reliable in the predicted condition ($p < .05$, one-tailed test; $d = 0.26$), suggesting that the converters produced a larger Stroop effect in the slowest bins, reflecting more skewing in the incongruent con-
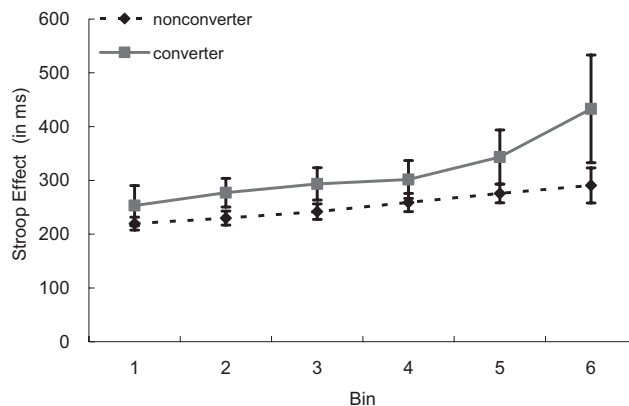

Figure 1. Mean percentage of color-naming errors as a function of condition and participant group.

Figure 2. Vincentile plots of the Stroop effect as a function of participant group.

dition. None of the remaining comparisons approached significance. Thus, the results from the Vincentile analyses converge with the results of the ex-Gaussian analyses.

Cox Proportional Hazard Analysis

Although we partialled out participants' age in the analyses of variance as stated in Table 2, it is important to also take into account (a) the difference in the time of the last follow-up diagnoses for nonconverters and (b) the difference in the time of the DAT conversion for converters. Moreover, some nonconverters would likely have converted to very mild DAT had they not been dropped out of the studies due to death or other reasons. Hence, to statistically control the amount of time that the converters and nonconverters had been followed up, we used the Cox proportional hazard models to predict the time to DAT conversion. Two sets of analyses were performed, one for the Stroop effect in errors and the other for τ. The dependent variable was the DAT conversion status of each individual. The *time* variable was the interval (in days) between the time of the Stroop test and the time of the last follow-up diagnoses for nonconverters and the interval between the time of the Stroop test and the time of first DAT conversion for converters. The Stroop effect in errors and τ significantly predicted the DAT conversion after we partialled out the age and the death status (as a categorical variable) of our participants (for Stroop effect in errors, $\Delta\chi^2[1] = 3.89$, $p < .05$, OR $= 1.043 \pm 0.039$; for τ, $\Delta\chi^2[1] = 6.98$, $p < .01$, OR $= 1.009 \pm 0.006$).

Psychometric Test Performance

Table 3 presents standard psychometric means for the converters and nonconverters across a wide range of tasks. Overall the converters performed slightly worse (but not statistically) than the nonconverters on most of the measures. However, there are two points that are particularly noteworthy. First, the standard episodic memory measures, logical memory and associate recall, did not differ between converters and nonconverters ($p = .82$ and $p = .47$, respectively). Second, the converters performed worse than did nonconverters in tasks that measure spatial abilities, Benton Copy Form D and the Wechsler Adult Intelligence Scale—Revised Block Design ($p = .084$ and $p = .041$, respectively). None of the remaining psychometric measures approached statistical significance between converters and nonconverters.

Finally, we addressed whether the targeted measures from the Stroop task (i.e., the Stroop effect in error rates and the tail of the reaction time distributions) predicted status (converter vs. nonconverters) above and beyond each of the psychometric measures in a series of logistic regressions predicting conversion status. These analyses were identical to the Cox proportional hazard analyses described above, with the addition of adding a given psychometric measure as a covariate in the analysis. As shown in Table 4, the error rates consistently produced a reliable increment in discrimination between converters and nonconverters over most of the psychometric measures. Also, as shown in Table 5, the same pattern holds for the tail of the reaction time distribution. (The discrepancy in the tables for the psychometric measures is due to two of the nonconverters not being fit by the ex-Gaussian analyses, as noted above.) Hence, these variables do appear to afford unique predictive power above and beyond standard psychometric test performance.[3]

General Discussion

The present analyses were motivated by (a) recent suggestions that attentional measures may be useful in discriminating preclinical AD from healthy brain aging, (b) the relative paucity of attentional measures in investigating conversion rates, (c) evidence that the trial-by-trial computerized Stroop task is a particularly sensitive measure in discriminating healthy older adults from early stage DAT individuals, and (d) the importance of better understanding the nature of cognitive decline in what appears to be an apparently healthy control group. Hence, we focused on a set of 47 healthy control individuals who were originally tested during 1992–1994 to determine if there was any specific utility of the trial-by-trial Stroop color-naming task in predicting conversion. The results were quite clear. Although reaction times from the Stroop task were in the predicted direction, the strongest differences among all measures, including a large set of psychometric tests, were in the error rates in the Stroop incongruent condition.

It is particularly interesting that the error rates were the strongest discriminator. This is consistent with the Perlstein et al. (1998) study reviewed in the introduction, where they showed that the Stroop error rates were the most sensitive measure of the attentional breakdown between healthy controls and individuals with schizophrenia. Moreover, Spieler et al. (1996) found that the error rates in the incongruent condition in the Stroop task were particularly useful in discriminating the healthy controls from the DAT individuals. Hence, it may indeed be the case that the error rate in the Stroop task, which has often not been measured in the card-reading version of the task (see Perlstein et al., 1998), may be a particularly sensitive measure of cognitive decline (see Balota & Faust, 2001, for further discussion of error rates).

Past researchers of DAT conversion that have employed the Stroop task have to our knowledge always included the card-reading version of this task. As noted, the card-reading version has a number of advantages, including ease of administration and short duration but may not be as sensitive as the trial-by-trial computerized version of the task. As Perlstein et al. (1998) have shown, the card-reading version of the task is relatively less sensitive to differences between two groups of individuals who have attentional control differences. Hence, if one is interested in subtle cognitive changes in attention, the present results converge with Perlstein et al. in suggesting that the error rates in the Stroop computerized task are particularly useful. Given the current ease with which computers can be used to administer such a task, and the increased ease for older adults using computers, this should be considered as a standard measure for future studies. In addition, it is not possible to explore the reaction time distributions with a card-reading version of the task, and this also

---

[3] We also looked at the correlations between the Stroop error rate and the τ estimates with each of the psychometric measures, after partialling out the effects of age. Interestingly, the only two measures that were correlated with both measures were Trailmaking B ($r = .27$, $p = .08$, for error rates and $r = .36$, $p = .02$, for the τ estimates) and the Benton Copy Form D task ($r = -.33$, $p = .03$, for error rate and $r = -.29$, $p = .06$, for the τ estimates). Trailmaking B is probably the best indicant of attention in the psychometric battery and Benton Copy Form D is one of the few psychometric measures that also discriminates between the converters and nonconverters.

Table 3
Psychometric Measures as a Function of Participant Group

Measure	Nonconverter		Converter		df	F	η_p^2
	M	SD	M	SD			
Logical Memory	9.83	2.91	9.00	3.42	(1, 44)	0.05	0.00
Forward Digit Span	6.63	1.37	6.25	1.48	(1, 44)	0.58	0.01
Backward Digit Span	5.23	1.37	4.50	1.24	(1, 44)	1.81	0.04
Trail Making (Form A)	48.54	15.63	47.75	12.94	(1, 44)	0.55	0.01
Trail Making (Form B)	104.43	41.31	126.83	43.50	(1, 44)	1.04	0.02
WAIS-R Information	22.29	3.80	19.75	3.96	(1, 44)	2.77	0.06
WAIS-R Block Design	33.31	8.35	26.00	8.52	(1, 44)	4.43*	0.09
WAIS-R Digit Symbol	45.09	12.86	41.08	7.87	(1, 44)	0.17	0.00
Benton Delay (Form C)	6.21	2.07	5.17	1.59	(1, 43)	0.78	0.02
Benton Copy (Form D)	9.91	0.38	9.58	0.67	(1, 43)	3.12^	0.07
Boston Naming Test	56.29	3.56	53.50	4.78	(1, 44)	2.79	0.06
Crossing Off	159.54	40.82	158.08	29.03	(1, 44)	0.16	0.00
Mental Control	7.41	2.31	6.42	2.19	(1, 43)	0.49	0.01
Associate Recall	14.90	4.45	14.71	4.08	(1, 44)	0.53	0.01
Word Fluency (Letters S&P)	31.23	10.41	28.75	10.04	(1, 44)	0.18	0.00

Note. The statistical significances labeled in the *F* column were based on an analysis of covariance, with participants' age at 1993 being treated as a covariate. Except the Benton Delay Form C, Benton Copy Form D, and Mental Control, which were based on 12 converters and 34 nonconverters, all of the other variables are based on 12 converters and 35 nonconverters. The η_p^2 column indicates the effect size η_p^2 for the analyses. WAIS-R = Wechsler Adult Intelligence Test—Revised.
^ $p < .10$.   * $p < .05$.

appears to be a useful discriminator between the converters and nonconverters.[4]

Of course, one might ask why there should be a breakdown in Stroop performance for those individuals destined to be diagnosed with DAT. As noted, there has been accumulating evidence that attentional systems may break down early in the progression of the disease. One might argue that the attentional systems that are used to control prepotent pathways are especially complex and highly evolved. For example, performance in the Stroop task has been linked to higher level executive control systems (see, e.g., Kane & Engle, 2003). Indeed, a recent study by Fine et al. (2008) showed that estimates from a Stroop switching task (the CWIT) were better predictors of cognitive decline than ApoE4 status at baseline. Thus, it is quite possible that attentional control systems are placed under considerable stress when strong prepotent pathways need to be controlled. Indeed, it is interesting that the predictive power of ApoE4 for conversion did not reach standard levels of reliability (likely due to the small sample), but the Stroop effect in error rates did.

If attentional control systems are a useful early marker for DAT conversion, then one needs to reconcile this perspective with the striking memory decline in early stage DAT and the neuropathology that develops in the medial temporal structure, an area where declarative memory has been a hallmark function. First, with respect to the neuropathology, there is now considerable evidence of widespread breakdown across a number of distinct areas of the brain (e.g., Buckner, 2004). Interestingly, recent Pittsburgh compound B imaging studies indicate that amyloid buildup occurs relatively strongly in frontal areas (e.g., Mintun et al., 2007). There is also evidence of cortical–cortical white matter disconnection (see Bartzokis et al., 2004, for a recent review), which could be particularly important for attentional control systems that need to coordinate multiple sources of information. As Bartzokis et al. (2004) argued, there is evidence that AD selectively affects certain

laminae and cell types within the cortex in layers II, III, and IV and the pyramidal neurons, which presumably participate in cortico-cortical communication. In addition, it is quite likely that the cholinergic deficits observed in AD produce a decrease in attentional control systems because of the innervation to frontal, thalamic, and parietal areas.

Turning to the cognitive manifestation of the disease, there is no doubt that memory is disrupted in DAT. However, it is also clear that there is an intimate relation between attention and the declarative type memory tests tapped by the standard logical memory, associate recall, and selective reminding tasks that are often used in both discriminating healthy controls and early stage DAT individuals and in progression studies (e.g., Sarazin et al., 2007). Memory researchers have long recognized the fact that attention is critical both during encoding in laying down distinct traces and during retrieval when generation processes are necessary to retrieve the earlier encoded traces (see Craik, 2001; Craik & Lockhart, 1972; Mangels, Picton, & Craik, 2001; Naveh-Benjamin, Craik, Guez, & Kreuger, 2005). Interestingly, there has been evidence by Sommers and Huff (2003) showing a direct relationship between Stroop effects (above and beyond overall slowing) and false alarms in the Deese/Roediger–McDermott false memory paradigm (Roediger & McDermott, 1995), a task that has been shown to be particularly sensitive to DAT (see Balota et al., 1999). Moreover, Balota et al. (2002) have demonstrated that healthy young adults produce a memory pattern similar to individuals with

---

[4] One might also ask if a vocal Stroop task is necessary, which demands the use of a voice key, compared to a simpler button press task. Although it may be possible to develop a button press version of the Stroop that produces large interference effects (possibly by increasing the proportion of congruent trials), in his review article, MacLeod (1991) concluded that the vocal version of the Stroop produces larger interference effects than does a manual response.

Table 4

Cox Regression Analyses of Stroop Effect in Errors on Predicting Dementia of the Alzheimer's Type Conversion (Converters vs. Nonconverters)

Variable	Initial regression model: $\chi^2(3)$	After adding Stroop effect in errors	
		OR ± 95% CI	$\Delta\chi^2$
Logical Memory	3.75	1.044 ± 0.038*	4.20*
Forward Digit Span	5.09	1.084 ± 0.058*	8.85*
Backward Digit Span	5.12	1.038 ± 0.038*	3.35^
Trail Making (Form A)	2.61	1.043 ± 0.040*	3.88*
Trail Making (Form B)	4.86	1.036 ± 0.043^	2.58
WAIS-R Information	6.06	1.036 ± 0.038^	3.17^
WAIS-R Block Design	8.80*	1.021 ± 0.045	0.83
WAIS-R Digit Symbol	3.67	1.042 ± 0.040*	3.74^
Benton Delay (Form C)	4.83	1.038 ± 0.039^	3.19^
Benton Copy (Form D)	12.95*	1.023 ± 0.045	1.08
Boston Naming Test	9.43*	1.051 ± 0.044*	4.54*
Crossing Off	2.62	1.043 ± 0.041*	3.81^
Mental Control	3.12	1.042 ± 0.040*	3.67^
Associate Recall	2.59	1.043 ± 0.040*	3.85*
Word Fluency (Letters S&P)	3.04	1.053 ± 0.045*	4.87*

Note. Analyses were based on 12 converters and 35 nonconverters. The $\chi^2(3)$ column indicates χ^2 statistics for the model including age, death, and psychometric measure in the first to third step, respectively. The $\Delta\chi^2$ column indicates the increment in χ^2 statistics after adding Stroop effect in errors in the last step of the model. WAIS-R = Wechsler Adult Intelligence Test—Revised.
^ $p < .10$.   * $p < .05$.

early stage DAT, when placed under deadlines that minimize attentional recollection (see Jacoby, 1999, for evidence of the role of attention in such procedures). Thus, it is clear that there is an intimate tie between attention and declarative memory performance and so a breakdown in declarative memory performance could at least in part be linked to the breakdown in attentional systems.

In this light we are in clear agreement with Twamley et al. (2006) who argued that attentional measures should receive a higher priority in future studies of progression. Indeed, by studying individuals who already are at the MCI stage or are beginning to produce some memory impairment, one may be stacking the deck in favor of finding more of a breakdown in a memory system that

Table 5

Cox Regression Analyses of τ on Predicting Dementia of the Alzheimer's Type Conversion (Converters vs. Nonconverters)

Variable	Initial regression model: $\chi^2(3)$	After adding τ	
		OR ± 95% CI	$\Delta\chi^2$
Logical Memory	4.11	1.008 ± 0.006*	6.28*
Forward Digit Span	6.28^	1.011 ± 0.007*	9.33*
Backward Digit Span	6.69^	1.008 ± 0.006*	6.97*
Trail Making (Form A)	2.93	1.009 ± 0.006*	6.98*
Trail Making (Form B)	5.49	1.008 ± 0.007*	4.97*
WAIS-R Information	7.27^	1.007 ± 0.006*	5.03*
WAIS-R Block Design	9.77*	1.005 ± 0.007	2.72^
WAIS-R Digit Symbol	4.10	1.008 ± 0.006*	6.01*
Benton Delay (Form C)	5.88	1.008 ± 0.007*	5.14*
Benton Copy (Form D)	12.47*	1.006 ± 0.007^	3.87*
Boston Naming Test	11.74*	1.011 ± 0.008*	8.05*
Crossing Off	2.97	1.008 ± 0.006*	6.91*
Mental Control	4.84	1.008 ± 0.007*	5.85*
Associate Recall	2.96	1.009 ± 0.006*	7.08*
Word Fluency (Letters S&P)	3.82	1.010 ± 0.007*	8.13*

Note. Analyses were based on 12 converters and 33 nonconverters (after dropping 2 nonconverters with misfit ex-Gaussian models). The $\chi^2(3)$ column indicates χ^2 statistics for the model including age, death, and psychometric measure in the first to third step, respectively. The $\Delta\chi^2$ column indicates the increment in χ^2 statistics after adding τ in the last step of the model. WAIS-R = Wechsler Adult Intelligence Test—Revised.
^ $p < .10$.   * $p < .05$.

is already in decline. Of course, the difficulty one encounters here is to utilize attentional measures that are sensitive to such subtle changes in attentional control systems. The present results, along with those by Perlstein et al. (1998) and Spieler et al. (1996), suggest that the intrusion errors in a computerized Stroop task may be a useful early marker. Attentional control as measured by switching tasks (see Fine et al., 2008) and possibly working memory measures should also be targeted in such studies (see Rosen, Bergeson, Putnam, Harwell, & Sunderland, 2002).

The present results also converge with conclusions from recent investigations that there may be preclinical cognitive changes in a subset of our healthy control individuals who may be lowering overall mean performance of the older adult sample. For example, in a large scale longitudinal project, Sliwinski et al. (2003) showed that virtually all of the memory decline in about 25% of healthy older adult samples could be explained by eventual disease conversion. Thus, healthy older adult samples in cognitive aging research are likely to include individuals who are in the earliest stages of cognitive decline associated with the onset of dementing illnesses. Indeed, as noted earlier approximately one quarter of healthy control individuals who have come to autopsy have sufficient neuropathology to be diagnosed with AD (see Price & Morris, 1999). It is noteworthy that the control sample in the current study is likely to be relatively cleaner than most samples, because one of the major goals of this research program is the detection of the earliest stages of dementing illnesses (see Storandt et al., 2006). The rigorous screening process involved in the CDR scale (see Morris et al., 2001) is clearly beyond the resources typically used to screen individuals in cognitive aging studies.

Of course, it is also important to recognize the limitations in the present study. First, although there were clearly reliable differences between converters and nonconverters, the sample size was relatively small. Larger scaled studies with a greater variety of behavioral and biological markers will be critical in the next step. Second, it will be useful to determine if the slope of the attentional measures changes across longitudinal tests, as advocated by Albert et al. (2007), among others. Both of these issues are being addressed in the ongoing research program. Finally, it is also important to recognize that this is not a population-based study but a study based on a sample of dedicated participants who have been participating in this longitudinal study.

In conclusion, the Stroop effect in errors and the tail of the reaction time distributions appear to be useful markers to identify cognitively healthy older adults who are at a higher risk for developing DAT. Because the Stroop color-naming task has been viewed as the prototype measure of attentional control (e.g., MacLeod, 1992), it is possible that deficits in attentional control are particularly predictive of later DAT. This contrasts with the view that an episodic memory deficit is the critical symptom for predicting the development of DAT. Indeed the declarative memory measures were relatively poor predictors in the present study. Clearly, further exploration of changes in attentional control mechanisms as an early marker for the onset of DAT appears to be a useful avenue in future studies of progression. Finally, it is important to keep in mind that "healthy" older adult samples used in cognitive aging studies are likely to include individuals with preclinical cognitive changes that not only influence memory performance but also influence attentional selection tasks such as Stroop.

References

Albert, M., Moss, M. B., Blacker, D., Tanzi, R., & McArdle, J. J. (2007). Longitudinal change in cognitive performance among individuals with mild cognitive impairment. *Neuropsychology, 21,* 158–169.

Albert, M., Moss, M. B., Tanzi, R., & Jones, K. (2001). Preclinical prediction of AD using neuropsychological tests. *Journal of the International Neuropsychological Society, 7,* 631–639.

Armitage, S. G. (1945). An analysis of certain psychological tests used for the evaluation of brain injury. *Psychological Monographs, 60*(1, Whole No. 177), 1–48.

Baddeley, A. D. (2002). Fractionating the central executive. In R. T. Knight & D. T. Stuss (Eds.), *Principles of frontal lobe function* (pp. 246–260). New York, NY: Oxford University Press.

Baddeley, A. D., Baddeley, H. A., Bucks, R. S., & Wilcock, G. K. (2001). Attentional control in Alzheimer's disease. *Brain, 124,* 1492–1508.

Balota, D. A., Burgess, G. C., Cortese, M. J., & Adams, D. R. (2002). The word-frequency mirror effect in young, old, and early stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language, 46,* 199–226.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memory in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16,* 361–384.

Balota, D. A., & Duchek, J. M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. *Brain and Language, 40,* 181–201.

Balota, D. A., & Faust, M. E. (2001). Attention in dementia of the Alzheimer's type. In F. Boller & S. Cappa (Eds.), *Handbook of neuropsychology* (2nd ed., pp. 51–80). New York, NY: Elsevier Science.

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59,* 495–523.

Balsis, S., Carpenter, B. D., & Storandt, M. J. (2005). Personality change precedes clinical diagnosis of dementia of the Alzheimer type. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 60,* P98–P101.

Bartzokis, G., Sultzer, D., Lu, P. H., Nuechterlein, K. H., Mintz, J., & Cummings, J. L. (2004). Heterogeneous age-related breakdown of white matter structural integrity: Implications for cortical "disconnection" in aging and Alzheimer's disease. *Neurobiology of Aging, 25,* 843–851.

Belleville, S., Bherer, L., Lepage, E., Chertkow, H., & Gauthier, S. (2008). Task switching capacities in persons with Alzheimer's disease and mild cognitive impairment. *Neuropsychologia, 46,* 2225–2233.

Benton, A. L. (1963). *The Revised Visual Retention Test: Clinical and Experimental Applications.* New York, NY: Psychological Corporation.

Berg, L., McKeel, D. W., Jr., Miller, J. P., Storandt, M., Rubin, E. H., Morris, J. C., . . . & Saunders, A. M. (1998). Clinicopathologic studies in cognitively healthy aging and Alzheimer's disease: Relation of histologic markers to dementia severity, age, sex, and apolipoprotein E genotype. *Archives of Neurology, 55,* 326–335.

Botwinick, J., & Storandt, M. (1973). Age differences in reaction time as a function of experience, stimulus intensity, and preparatory interval. *Journal of Genetic Psychology, 123,* 209–217.

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron, 44,* 195–208.

Burke, W. J., Miller, J. P., Rubin, E. H., Morris, J. C., Coben, L. A., Duchek, J. M., . . . & Berg, L. (1988). Reliability of the Washington University Clinical Dementia Rating. *Archives of Neurology, 45,* 31–32.

Castel, A. D., Balota, D. A., Hutchison, K. A., Logan, J. M., & Yap, M. J. (2007). Spatial attention and response control in healthy younger and older adults and individuals with Alzheimer's disease: Evidence for

disproportionate selection breakdowns in the Simon task. *Neuropsychology, 21,* 170–182.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., . . . & Pericak-Vance, M. A. (1993, August 13). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science, 261,* 921–923.

Craik, F. I. M. (2001). Effects of dividing attention on encoding and retrieval processes. In I. Neath, A. M. Surprenant, H. L. Roediger, & J. S. Nairne (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 55–68). Washington DC: American Psychological Association.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Duchek, J. M., & Balota, D. A. (2005). Failure to control prepotent pathways in early stage dementia of the Alzheimer's type: Evidence from dichotic listening. *Neuropsychology, 19,* 687–695.

Fagan, A. M., Roe, C. M., Xiong, C. J., Mintun, M. A., Morris, J. C., & Holtzman, D. M. (2007). Cerebrospinal fluid tau/beta-amyloid(42) ratio as a prediction of cognitive decline in nondemented older adults. *Archives of Neurology, 64,* 343–349.

Faust, M. E., & Balota, D. A. (1997). Inhibition of return and visual–spatial attention in healthy older adults and individuals with dementia of the Alzheimer's type. *Neuropsychology, 11,* 13–29.

Faust, M. E., & Balota, D. A. (2007). Inhibition, facilitation, and attention control in dementia of the Alzheimer type: The role of unifying principles in cognitive theory development. In D. S. Gorfein & C. M. McLeod (Eds.), *The place of inhibition in cognition.* New York, NY: Psychology Press.

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin, 125,* 777–799.

Fine, E. M., Delis, D. C., Wetter, S. R., Jacobson, M. W., Jak, A. J., McDonald, C. R., . . . & Bondi, M. W. (2008). Cognitive discrepancies versus APOE genotype as predictors of cognitive decline in normal-functioning elderly individuals: A longitudinal study. *American Journal of Geriatric Psychiatry, 16,* 366–374.

Hill, R. D., Storandt, M., & LaBarge, E. (1992). Psychometric discrimination of moderate senile dementia of the Alzheimer type. *Archives of Neurology, 49,* 377–380.

Jacoby, L. L. (1999). Ironic effects of repetition: Measuring age-related differences in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 3–22.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test scoring booklet.* Philadelphia, PA: Lea & Febiger.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

MacLeod, C. M. (1992). The Stroop task: The "gold standard" of attentional measures. *Journal of Experimental Psychology: General, 121,* 12–14.

Mangels, J. A., Picton, T. W., & Craik, F. I. M. (2001). Attention and successful episodic encoding: An event-related potential study. *Cognitive Brain Research, 11,* 77–95.

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

Mintun, M. A., LaRossa, G. N., Sheline, Y. I., Dence, C. S., Lee, S. Y., Mach, R. H., . . . & Morris, J. C. (2007). [11C]Pib in a nondemented population: Potential antecedent marker of Alzheimer disease. *Neurology, 67,* 446–452.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology, 43,* 2412–2414.

Morris, J. C., McKeel, D. W., Jr., Fulling, K., Torack, R. M., & Berg, L. (1988). Validation of clinical diagnostic criteria for Alzheimer's disease. *Annals of Neurology, 24,* 17–22.

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's disease. *Archives of Neurology, 58,* 397–405.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Kreuger, S. (2005). Divided attention in younger and older adults: Effects of strategy and relatedness on memory performance and secondary task costs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 520–537.

Perlstein, W. M., Carter, C. S., Barch, D. M., & Baird, J. W. (1998). The Stroop task and attention deficits in schizophrenia: A critical evaluation of card and trial-by-trial Stroop methodologies. *Neuropsychology, 12,* 414–425.

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease: A critical review. *Brain, 122,* 383–404.

Petersen, R. C., Stevens, J. C., Ganguli, M., Tangalos, E. G., Cummings, J. L., & DeKosky, S. T. (2001). Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review). *Neurology, 56,* 1133–1142.

Price, J. L., & Morris, J. C. (1999). Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Annals of Neurology, 45,* 358–368.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Rosen, V. M., Bergeson, J. L., Putnam, K., Harwell, A., & Sunderland, T. (2002). Working memory and apolipoprotein E: What's the connection? *Neuropsychologia, 40,* 2226–2233.

Sarazin, M., Berr, C., De Rotrou, J., Fabrigoule, C., Pasquier, F., Legrain, S., . . . & Dubois, B. (2007). Amnestic syndrome of the medial temporal type identifies prodromal AD: A longitudinal study. *Neurology, 69,* 1859–1867.

Servan-Schreiber, D., & Cohen, J. D. (1998). Stroop task, language, and neuromodulation: Models of cognitive deficits in schizophrenia. In D. L. Long, R. W. Parks, & D. S. Levine (Eds.), *Fundamentals of neural network modeling: Neuropsychology and cognitive neuroscience* (pp. 192–208). Cambridge, MA: MIT Press.

Sliwinski, M. J., Hofer, S. M., Hall, C., Buschke, H., & Lipton, R. B. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age. *Psychology and Aging, 18,* 658–671.

Sliwinski, M. J., Lipton, R. B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 51,* P217–P225.

Sommers, M. S., & Huff, L. M. (2003). The effects of age and dementia of the Alzheimer's type on phonological false memories. *Psychology and Aging, 18,* 791–806.

Spieler, D. H., Balota, D. A., & Faust, M. E. (1996). Stroop performance in healthy younger and older adults and in individuals with dementia of the Alzheimer's type. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 461–479.

Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2002). Rates of

progression in mild cognitive impairment and early Alzheimer's disease. *Neurology, 59,* 1034–1041.

Storandt, M., Grant, E. A., Miller, J. P., & Morris, J. C. (2006). Longitudinal course and neuropathologic outcomes in original vs. revised MCI and in pre-MCI. *Neurology, 67,* 467–473.

Thurstone, L. E., & Thurstone, T. G. (1949). *Examiner manual for the SRA Primary Mental Abilities Test.* Chicago, IL: Science Research Associates.

Twamley, E. W., Ropacki, S. A., & Bondi, M. W. (2006). Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. *Journal of International Neuropsychological Society, 12,* 707–735.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale.* Oxford, England: Psychological Corporation.

Wechsler, D. (1987). *Manual: Wechsler Memory Scale—Revised.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (3rd ed.): Administration and scoring manual.* San Antonio, TX: Psychological Corporation.

Wechsler, D., & Stone, C. P. (1973). *Manual: Wechsler Memory Scale.* New York, NY: Psychological Corporation.

Wilson, R. S., Schneider, J. A., Arnold, S. E., Bienias, J. L., & Bennett, D. A. (2007). Conscientiousness and the incidence of Alzheimer disease and mild cognitive impairment. *Archive of General Psychiatry, 64,* 1204–1212.