The Utility of Stroop Task Switching as a Marker for Early-Stage Alzheimer's Disease

Keith A. Hutchison
Montana State University

David A. Balota and Janet M. Ducheck
Washington University

Past studies have suggested attentional control tasks such as the Stroop task and the task-switching paradigm may be sensitive for the early detection of dementia of the Alzheimer's type (DAT). The authors of the current study combined these tasks to create a Stroop switching task. Performance was compared across young adults, older adults, and individuals diagnosed with very mild dementia. Results indicated that this task strongly discriminated individuals with healthy aging from those with early-stage DAT. In a logistic regression analysis, incongruent error rates from the Stroop switching task discriminated healthy aging from DAT better than any of the other 18 cognitive tasks given in a psychometric battery.

Keywords: Alzheimer's disease, attentional control, task switching, dementia, Stroop

In their seminal chapter on the role of attention on goal-driven behavior, Norman and Shallice (1986) argued that people need to exert attentional control whenever they encounter situations that involve planning, troubleshooting, technical difficulty, or novel sequences of action or when they need to overcome a strong habitual response. As an example of the last situation, imagine a recent retiree driving on the interstate with the goal of visiting a new friend that lives off exit 13. If the previous job she held for 30 years required her to take exit 12, it would not be surprising if she accidentally took this exit and perhaps drove toward her former work place before noticing the error.

Action slips such as these can lead to confusion, embarrassment, and even fears of dementia. In fact, such errors are often called "senior moments" and are interpreted as "memory" problems (i.e., forgot to make the correct turn). However, this memory interpretation is not entirely correct. If asked, it is unlikely that the woman really "forgot" where she was going. Instead, her failure was likely in resisting the strong habit of taking exit 12. As argued by Norman and Shallice (1986), attention is required to keep internal goals active long enough to exert influence over our actions. Because memory is dependent upon attending to goal-relevant information and ignoring goal-irrelevant information during en-

coding (see Hasher, Zacks, & May, 1999, for a review), breakdowns in the control of attention often manifest themselves as memory problems.

The difficulty in separating memory from attention has important implications for the diagnosis and treatment of Alzheimer's disease (AD). In fact, Balota and Faust (2001) have argued that individuals with dementia of the Alzheimer's type (DAT) have specific difficulties in attentional selection of relevant information over irrelevant information and that this difficulty contributes to observed memory problems. According to Balota and Faust's attentional control framework, attentional control underlies our ability to orchestrate thought and action in accord with internal goals, that is, to use a goal to modulate competition between relevant and irrelevant information. This view contrasts with the common assumption that memory alone is the earliest and most detectable sign of AD (Grady et al., 1988; Haxby et al., 1988; Lafleche & Albert, 1995; Reid et al., 1996). However, although it is acknowledged that memory impairment occurs early, recent evidence has demonstrated concurrent deficits in executive function even prior to diagnosis of AD (Albert, Moss, Tanzi, & Jones, 2001; Bäckman, Jones, Berger, Laukka, & Small, 2005). Similarly, while it is widely accepted that medial–temporal brain areas are affected earliest in the progression of AD (Braak & Braak, 1991), frontal areas such as the prefrontal cortex and anterior cingulate also show considerable pathology early on (Killiany et al., 2000; Mintun et al., 2007; van der Flier et al., 2002; Yamaguchi, Sugihara, Ogawa, Oshima, & Ithara, 2001). Because attentional control systems involve complex coordination of multiple sources of information from multiple brain areas, they may be especially vulnerable to early neuropathology accompanying AD (see Balota et al., in press, for more discussion).

The best known paradigm with which to investigate attentional control is the classic Stroop task (Stroop, 1935) in which participants are told to name the ink color in which words are presented while ignoring the words themselves. The irrelevant words can be either congruent (the word *blue* written in blue ink), incongruent (the word *red* written in blue ink), or unrelated to the color (the

word *poor* written in blue ink). Spieler, Balota, and Faust (1996) investigated effects of both healthy aging and DAT on Stroop performance and found that aging and dementia produced distinct patterns of performance. Relative to young adults, healthy older adults produced larger Stroop interference (incongruent–neutral) in reaction times (RTs) but not error rates, suggesting that older adults could successfully overcome the automatic "word" response but at a cost to RT. In contrast, individuals diagnosed with both very mild and mild DAT (according to the criteria of the Clinical Dementia Rating scale, CDR 0.5 and CDR 1.0, respectively, Morris, 1993) produced larger Stroop interference effects in RTs and exaggerated error rates to incongruent stimuli relative to age-matched healthy control individuals. The DAT groups also showed larger facilitation effects (neutral–congruent RT) than young and healthy older adults, indicative of incorrectly reading the word rather than color on congruent trials (see MacLeod, 1991, for discussion).

Kane and Engle (2003) provided evidence for two separable control processes necessary in the Stroop task: goal maintenance and response competition. According to Kane and Engle, goal maintenance reflects the ability to maintain the appropriate task set (e.g., respond "color" and ignore "word") across trials, whereas response competition reflects the ease with which people can select between appropriate and inappropriate competing response tendencies. A loss in goal maintenance results in quick errors in which the person simply responds to the word without any influence from the potentially competing color. Behaviorally, this manifests as both an increase in errors to incongruent stimuli and also as an increase in RT facilitation for congruent stimuli (since accidentally reading the word in this condition does not produce an error). This is exactly the pattern found by Spieler et al. (1996) for DAT individuals, and indeed Spieler et al. used the same attentional control interpretation to accommodate these results. However, even with correct goal maintenance, one still must overcome the competition for incongruent stimuli between the inappropriate word dimension and the appropriate color dimension. Although the healthy older adults showed no deficits in goal maintenance, they did show deficits in the resolution of this response competition, as evidenced by longer RTs to incongruent items relative to neutral items. Castel, Balota, Hutchison, Logan, and Yap (2007) recently found a pattern similar to that found by Spieler et al. (1996) using another response competition paradigm, the Simon task (Simon, 1969), in which participants are told to respond to the direction of an arrow (e.g., → or ←) presented in the left, middle, or right of the computer screen, with the left or right index finger.

In a recent imaging study, De Pisapia and Braver (2006) reported evidence for the involvement of both lateral prefrontal cortex (lPFC) and anterior cingulate (ACC) activation in Stroop performance. Specifically, according to De Pisapia and Braver, the lPFC is involved in active maintenance of task goals, whereas the ACC is necessary for the detection of conflict. If the lPFC fails to maintain the task goal, then the person becomes reliant upon the ACC to detect conflict prior to the individual's committing an error. If an error does occur, communication between the ACC and lPFC serves to refresh and maintain the goal on subsequent trials. Of course, these two neural systems nicely parallel the operations hypothesized to be involved in attentional control.

Promise of Attention Measures in Early Diagnosis of AD

Studies such as those by Spieler et al. (1996) and Castel et al. (2007) suggest that attentional selection tasks such as Stroop and Simon, which require both goal maintenance and competition resolution, may be especially sensitive measures to early disruptions in performance created by the onset of AD. In a longitudinal study designed to test this possibility, Balota et al. (in press) tracked the Spieler et al.'s (1996) healthy older subjects to predict progression to DAT within the next 13 years. Those healthy older adults who eventually converted to DAT ($N = 12$) had incongruent error rates that were 2.2 times higher ($M = 17.4\%$) than those ($N = 35$) who did not eventually convert to DAT ($M = 7.9\%$). Balota et al. then compared the predictive power of Stroop error rates to that of tasks involving episodic memory (e.g., Logical Memory and Associative Recall), simple span (e.g., Forward and Backward Digit Span), spatial abilities (e.g., Benton Copy [Benton, 1963] and Wechsler Adult Intelligence Scale–Revised [WAIS-R; Wechsler, 1955] Block Design), and Processing Speed (e.g., Crossing-Off Task and Trail-Making Tests, Parts A and B). Only the WAIS-R Block Design significantly discriminated between converters and noncovertvers. When entered into a regression equation, incongruent error rates were the strongest predictor of progression to DAT.

The Task Switching Paradigm

Given the potential importance of breakdowns in attentional control systems in early-stage DAT, it is possible that tasks that place a premium on control may be especially useful in early discrimination. It is in this light that we explored the task switching paradigm (Allport, Styles, & Hsieh, 1994; Rogers & Monsell, 1995). In this paradigm, two or more tasks are intermixed within a block of trials, and the participant must attend to cues that designate the appropriate task for the current trial. For instance, a participant may be shown two numbers and have to switch between adding and subtracting the second number from the first. The general finding from such tasks is that participants are slower or less accurate in responding on a "switch" trial than on a nonswitch trial, as long as the stimuli are compatible with both tasks (Spector & Biederman, 1976). Accurate performance in task switch paradigms requires participants to maintain multiple task sets (i.e., rules that govern the mapping between stimuli and their appropriate responses) in working memory while selecting the appropriate task set for the current trial. On switch trials, participants must engage in task set reconfiguration, which involves deactivating the current task set and retrieving and enabling the past task set (Rogers & Monsell, 1995). Although some reconfiguration can be initiated endogenously when an individual is given enough time between the cue and the next stimulus, a "residual switch cost" still remains. Rogers and Monsell interpreted this residual switch cost as reflecting an exogenous trigger necessary to complete the reconfiguration process (but see Allport & Wylie, 1999, 2000; Waszak, Hommel, & Allport, 2003; and Wylie & Allport, 2000, for an alternative explanation for residual switch costs).

Task Switching, Stroop, and DAT

As with the Stroop task, task switching paradigms have been found to activate regions of PFC (Brass, Ullsperger, Knoeche, von Cramon, & Phillips, 2005; Braver, Reynolds, & Donaldson, 2003). In addition, task switching typically activates posterior parietal regions (Eppinger, Kray, Mecklinger, & John, 2007; Filoteo et al., 1992), which are presumably involved in shifts of attention from one stimulus dimension to another. Because both Stroop and task switch paradigms involve related attentional control processes (such as the necessity to maintain and utilize task goals to suppress the more dominant process), it is not surprising that researchers have combined these tasks in order to investigate group differences in attentional control. Versions of Stroop switch tasks have been used to examine presumed cognitive control deficits among older adults (Eppinger et al., 2007), individuals with traumatic brain injury (Perlstein, Larson, Dotson, & Kelly, 2006), attention-deficit/hyperactivity disorder (Wu, Anderson, & Castiello, 2006), and Parkinson's disease (Woodward, Bub, & Hunter, 2002).

However, to our knowledge, investigators in only one study (Fine et al., 2008) have used a Stroop switch task to discriminate healthy aging from the type of cognitive decline typically seen in early-stage DAT. Fine et al. used a Stroop switching task in which participants were given an incongruent color word presented either inside or outside a small box on a computer screen. Participants were instructed to name the ink color if the stimulus appeared outside the box and name the word if it appeared inside the box. Participants also performed three other tasks: a standard (color naming of words only) version of Stroop, color naming of ink patches, and word reading for stimuli presented in black ink. The standardized difference between Stroop switch performance and average performance on the other three measures, called a Stroop discrepancy score, was then used to predict an individual's degree of decline on the Dementia Rating Scale (DRS; Mathis, 1973) over the following year. Fine et al. compared the predictability of this discrepancy score on cognitive decline to that of apolipoprotein E (ApoE) genotype, a known predictor of AD. Specifically, individuals with the e4 allele of this gene have an increased risk for developing AD (e.g., Blacker et al., 1997; Corder et al., 1993; Henderson et al., 1995). Of interest, Fine et al. found that those older adults who declined in cognitive performance over the year (as measured with the DRS) had larger Stroop switch discrepancy scores than those whose cognitive performance remained stable. In fact, logistic regression showed that the discrepancy score significantly predicted whether or not an individual showed cognitive decline (75% correct classification rate), yet ApoE status did not (67% correct classification rate). It is important to note that when the four tasks were examined separately, the difference between the cognitive declining and stable groups was greater in the Stroop switch task than in any of the other three tasks, suggesting the switch component may increase the sensitivity of the Stroop task for predicting cognitive decline.

Although the Fine et al. (2008) results are provocative, there are some limitations that prevent a complete understanding between task switching and DAT. One limitation is that in their Stroop switch task (the color word interference task; Delis, Kaplan, & Kramer, 2001), Fine et al. reported only an overall time to complete the entire block of trials, which prevents researchers from examining two critical components of performance: error analyses

and trial-specific switching effects. As discussed previously, errors on incongruent trials are often the most sensitive measure for discriminating performance between healthy older adults and DAT individuals (Balota et al., in press; Castel et al., 2007; Spieler et al., 1996). Researchers who use only overall completion times miss this important information (see Balota et al., in press, for further discussion of this issue). Similarly, the overall completion time score renders task switch trials inseparable from task repeat trials. Thus, this measure conflates effects of switching between tasks with overall difficulty of mixing tasks, relative to single task conditions. In fact, several researchers have found that the cost of "mixing" is much greater than the cost of "switching." Under mixed task conditions, performance even on task repetition trials is much slower and less accurate than under single task conditions (Pashler, 2000). This mixing cost is thought to represent a sustained effort to keep multiple tasks active in working memory, whereas the switch cost presumably reflects the more transient updating or reconfiguration of stimulus–response (S–R) mappings (Braver et al., 2003). Past research has demonstrated that older adults show a large increase in mixing cost, but only a subtle (often nonsignificant) increase in switching costs, relative to young adults (Kramer, Hahn, & Gopher, 1999). Thus, it is unknown whether Fine et al.'s "cognitive decliners" were impaired relative to the stable group on simply overall RT (due to mixing costs) or on switch trials specifically. This confound of mixing and switching is especially problematic for the Fine et al. study because their decliner group was older ($M = 78.8$ years) than their stable group ($M = 74.4$ years) and thus may have been more impaired by overall mixing costs.

Two other factors may limit generalization of the Fine et al. (2008) results to the study of dementia and attentional control. First, the samples differed in initial cognitive performance. Not only did the decliner group start out older than the stable group, but they also started out with lower scores on the DRS. In fact, the decliner group started out with lower scores than those with which the stable group ended up. Perhaps the decline in DRS accelerates once performance drops beyond certain levels. Second, neither sample progressed into levels typically used as screening cutoffs for DAT. Even though the decliners had lower DRS scores than the stable group, their final scores ($M = 133.3$) were still within the range of healthy older adults (130–144; Rosser & Hodges, 1994). Therefore, although the Fine et al. study is provocative, further work is needed to explore how well Stroop task switching (a) discriminates early-stage DAT individuals from healthy controls or (b) predicts eventual progression into DAT from an initially healthy sample. In the current study, we primarily emphasized the power of this paradigm for discrimination.

The Current Study

In the present study, we examined the performance of young adults, healthy older adults, and individuals with very mild DAT in a trial-by-trial computerized version of the Stroop switching task. Examining performance on a trial-by-trial basis allows testing for not only overall group differences in the task (reflecting a general difficulty in mixing tasks) but also specific deficits among Stroop interference, switch costs, or their potential interaction (i.e., a task switch asymmetry). We also report evidence from psychometric tests that were available for these individuals to determine the

extent to which more traditional cognitive measures are useful in discriminating between healthy aging and early-stage DAT. It is important to emphasize here that the DAT individuals are at the earliest detectable stage of AD. Indeed, the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) scores for the very mild DAT individuals (MMSE = 28.2) are only 1 point lower than those for the healthy control individuals (MMSE = 29.2) in this sample. The important question was whether the stress placed on the attentional system by the present Stroop switching task would afford better discrimination between these high-functioning DAT individuals and the healthy controls than the standard psychometric measures.

Overall, we predicted that performance on this task should decline as a function of both healthy aging and dementia. However, aging and dementia should produce separate patterns of deficits relative to young adults. Specifically, although both healthy older adults and DAT individuals should show overall worse performance than young adults, indicative of difficulty mixing multiple tasks, incongruent error rates should be especially sensitive for discriminating DAT from healthy aging (similar to Spieler et al., 1996, and Castel et al., 2007). In addition, if trial-specific switch costs are due to an automatic retrieval of previous competing S–R mappings (Allport & Wylie, 2000), then we should observe a similar switch cost pattern across all groups, because age-related changes in automatic processes are relatively preserved across the life span (see Balota, Dolan, & Duchek, 2000; Hasher & Zacks, 1979). However, if trial-specific switch costs are due instead to persisting inhibition of the current task set from previous trials, then we might expect reduced costs from the DAT group in whom inhibitory functions have been shown to be compromised (Balota & Faust, 2001).

Method

Participants

Older adult participants were recruited from the Washington University Medical School Alzheimer's Disease Research Center (ADRC) and consisted of 64 healthy older adults and 32 individuals with very mild DAT. There was no significant difference in age ($p > .41$) between healthy older adults ($M = 77.24$ years, $SD = 9.80$) and DAT individuals ($M = 78.78$ years, $SD = 5.89$). There was also no significant difference in education level ($p > .91$) between healthy older adults ($M = 14.66$ years, $SD = 2.610$) and DAT individuals ($M = 14.72$ years, $SD = 2.88$). In addition, 30 young adults (age 25 years or younger) were recruited from the Washington University student community and participated for course credit or were paid $10. The young adults had a mean age of 20.8 years ($SD = 1.5$).[1]

The healthy older adults and the individuals with DAT were seen by a physician and completed a battery of psychometric tests approximately once a year. All participants were screened for neurological, psychiatric, or medical disorders with the potential to cause dementia. The inclusion and exclusion criteria for diagnosis of DAT have been described in detail elsewhere (e.g., Morris, 1993; Morris, McKeel, Fulling, Torack, & Berg, 1988) and conform to those outlined in the criteria of the National Institute of Neurological and Communications Disorders and Stroke–Alzheimer's Disease and Related Disorders Association (McKhann et

al., 1984). Dementia severity for each individual was staged in accordance with the Washington University Clinical Dementia Rating (CDR) Scale (Hughes, Berg, Danziger, Coben, & Martin, 1982; Morris, 1993). According to this scale, a CDR of 0 indicates no cognitive impairment, a CDR of 0.5 indicates very mild dementia, a CDR of 1.0 indicates mild dementia, and a CDR of 2.0 indicates moderate dementia. At the Washington University Medical School ADRC, a CDR of 0.5 has been found to accurately indicate the earliest stages of AD (Morris, McKeel, & Storandt, 1991). All of the current individuals with very mild dementia had a 0.5 CDR score. Both the reliability of the CDR and the validation of the diagnosis (based upon autopsy) by the research team have been excellent (93% diagnostic accuracy) and well documented (e.g., Berg et al., 1998; Burke et al., 1988).

Psychometric Test Information

In addition to participating in the experimental task, all of the older adults participated in a 2-hr battery of psychometric tests as part of a larger longitudinal study of cognitive performance in healthy aging adults and DAT individuals. The results from the psychometric tests are displayed in Table 1. The Wechsler Memory Scale included the Logical Memory–Immediate subtest (recall of scoring units 0–23), Logical Memory–Delayed subtest (recall of scoring units 0–25), Forward and Backward Digit Span (number of correct digits, 0–8 or 0–7, respectively), Associate Memory Recognition (0–7), and Associate Recall (0–21; Wechsler & Stone, 1973), and the Selective Reminding Task (Grober, Buschke, Crystal, Bang, & Dresner, 1988). The WAIS–R included the Information (scoring range 0–29), Block Design (scoring range 0–48), and Digit Symbol (scoring range 0–90) subtests and were scored according to the manual (Wechsler, 1955). Participants also received the Benton Visual Retention Test and the Benton Copy Test (number correct and number of errors, respectively; Benton, 1963), and Part A and Part B of the Trail-Making Test (number of seconds to complete; Armitage, 1945). Part B of the Trail-Making Test is used to assess not only visual perceptual–motor performance but also the ability to alternate between well-learned sequences (alphabet and numbers). Tests of semantic–lexical retrieval and word fluency included the Boston Naming Test (Goodglass & Kaplan, 1983b) and the Animal (Goodglass & Kaplan, 1983a) and Word S–P (Thurstone & Thurstone, 1949) Fluency Tests (number correct of 60; number named in 1 min., respectively). The National Adult Reading Test (NART; Nelson, 1982) was used as an assessment of word comprehension. Finally, the MMSE was given as an assessment of overall cognitive functioning.

Psychometric tests are scored such that greater scores indicate better performance with the exception of Trail-Making Parts A and B and Benton Copy errors, for which higher scores indicate poorer

---

[1] There were also five individuals with mild DAT in this study. However, data from these individuals are not included due to the small sample. In general, the data from the mild group resembled that from the very mild group. However, there was one mild AD participant with 100% errors in color naming and 99% accuracy in word naming, suggesting either an inability to understand the directions or, perhaps of more interest, a complete inability to overcome the dominant word response in favor of the nondominant color response.

Table 1
Psychometric Measures as a Function of Participant Group

Measures	Healthy older	Very mild AD	df	t value
Logical Memory[a]	10.86	9.09	1, 94	2.42*
Delayed Logical Memory	9.33	6.84	1, 89	2.92*
Associative Recognition	7.00	6.83	1, 90	2.56*
Associative Recall[a]	15.57	12.33	1, 90	3.97*
Selective Reminding Task	27.29	20.61	1, 85	4.31*
Trail-Making (Part A)[a]	39.45	42.47	1, 94	0.69
Trail-Making (Part B)[a]	88.00	110.59	1, 94	2.90*
WAIS-R Digit Symbol[a]	46.86	39.81	1, 94	2.84*
Forward Digit Span[a]	6.59	6.53	1, 94	0.21
Backward Digit Span[a]	5.02	4.47	1, 94	1.99*
WAIS-R Block Design[a]	32.50	27.10	1, 94	2.93*
Benton Copy (Form D)[a]	9.81	9.66	1, 90	1.15
WAIS-R Information[a]	21.84	20.13	1, 94	1.91[†]
Boston Naming Test[a]	55.45	52.09	1, 94	2.63*
Animal Fluency	19.08	16.47	1, 94	2.24*
Word Fluency (Letters S & P)[a]	32.91	27.09	1, 94	2.72*
Mini Mental State Examination	29.19	28.22	1, 94	3.07*
National Adult Reading Test	34.63	31.79	1, 90	1.76[†]

Note. WAIS-R = Wechsler Adult Intelligence Test–Revised.
[a] Denotes tasks used by Balota et al. (in press).
[†] $p < .10$.   * $p < .05$.

performance. Psychometric testing always occurred within a 2-month window of the Stroop switch task testing session. As shown in Table 1, as expected, the very mild DAT group performed more poorly than the healthy older group on most tests. Since young adults were not recruited by the ADRC, they did not receive the psychometric battery.

Apparatus and materials. We performed the experiment on a Pentium II IBM-compatible computer (Intel Corp., Santa Clara, CA), with a standard 38.10-cm (15-in.) monitor using E-prime software (Psychology Software Tools, Pittsburgh, PA). Participants viewed the display from an approximate distance of 50 cm. The response cues (color or word) and Stroop stimuli were presented in Courier New 16-point font in the center of the screen against a black background. The Stroop stimuli were taken from Spieler et al. (1996) and consisted of either color words (*red*, *green*, *blue*, and *yellow*) or neutral words (*bad*, *deep*, *poor*, and *legal*) matched to the color words in phoneme characteristics and printed word frequency.

Participants completed 144 experimental trials consisting of 68 neutral trials and 76 incongruent trials. The four color words were presented 19 times and appeared in each of the three incongruent colors (e.g., the word *red* presented in green, blue, or yellow). Similarly, the matched neutral words were presented 17 times each and appeared in each of the same three colors as their matched color word (e.g., the word *bad* presented in green, blue, or yellow). There were no congruent trials in the experiment.

Procedure

The Stroop task was embedded in a battery of other tasks involving memory and attention performance that lasted approximately 2 hr. At the beginning of the Stroop task, participants were instructed that they would be cued prior to each trial as to whether they should name the color in which a word was presented or the word itself. Either the cue color or the cue word was presented in the center of the screen for 1,400 ms and was immediately followed by the Stroop stimulus. The color and word cues were presented in an alternating-runs (e.g., AABBAABB) fashion (Rogers & Monsell, 1995) such that every two trials, the participants would switch their responding from one dimension (color or word) to the next. This paradigm allows for the comparison of performance on switch trials (i.e., AB, BA) with performance on nonswitch trials (i.e., AA, BB). As noted by Wu et al. (2006), however, this sequence does not provide a "pure" measure of switch costs because costs can persist beyond the first trial of a sequence (Allport & Wylie, 2000). Nonetheless, previous researchers have consistently found larger costs on the first trial (the current switch trial) than on the next trial (the current nonswitch trial), suggesting that this difference can still capture some component of switching difficulty.[2]

Participants were given eight sample trials in the alternate-runs format with the microphone turned off to establish that they understood the instructions. If the participant responded correctly on the sample trials, he or she was presented with 16 practice trials with the microphone turned on and was asked to respond as quickly and accurately as possible into the microphone. Following the practice trials, participants received the 144 experimental trials. The words and colors were presented in a fixed-random order with the constraint that neither the word nor its color was contained in the immediately preceding trial, preventing immediate positive or negative priming from one trial to the next. Self-paced rest breaks were given approximately every 40 trials. Experimenters recorded participants' RTs with millisecond accuracy using E-prime's PST serial response box and coded each response as (a) a correct response, (b) a response error, or (c) a microphone error. Response errors consisted of either responding with the wrong word (e.g., responding "green" to the word *green* written in blue) or responding with a blended word (e.g., "gre–blue").

Results

Only correct responses were considered for the RT analyses. A separate mean and standard deviation were computed for neutral and incongruent stimuli and for color and word cues. We removed outliers using the modified nonrecursive procedure suggested by Van Selst and Jolicoeur (1994). This procedure removed 2.9% of the correct RTs. We computed task switching effects by subtract-

---

[2] Allport and Wylie (2000) found that performance on no-switch trials of an alternating-runs paradigm was impaired relative to performance under pure task blocks. In an experiment in which tasks were blocked separately, Allport et al. (1994, Experiment 4) found evidence of impairment from competing S–R mappings even after 100 trials in the new task. In other experiments in which tasks were blocked, Allport and Wylie (2000; see also Wylie and Allport, 2000) found a "restart" effect for the later task (longer RT to the first trial in a sequence following a break in trials) even though the task itself did not switch. Allport and Wylie used both the persistent cost on no-switch trials and the restart effects to question whether the alternating-runs paradigm could accurately measure switch costs. It should be emphasized, however, that the impairment was greater for initial trials following a task switch than for initial trials that simply followed a rest break, suggesting that the initial trial following a switch reflects more than simply a restart cost.

ing the mean RT in the nonswitch condition from the switch condition, response cue effects by subtracting the mean RT in the word cue condition from the color cue condition, and the Stroop interference effects by subtracting the mean RT in the neutral item condition from the incongruent item condition.

The trimmed RTs for each participant were also transformed into z scores on the basis of each participant's overall trimmed mean RT and standard deviation. This transformation accounts for group differences in overall RT and variability, which can either produce spurious Group × Difference Score interactions or mask true Group × Difference Score interactions (see Faust, Balota, Spieler, & Ferraro, 1999; Hutchison, Balota, Cortese, & Watson, 2008, for more discussion). Rather than report redundant RT analyses using both measures, we only report RT effects that were confirmed by the z score analyses.

Arithmetic means for young, older, and very mild DAT participants based on individual participants' trimmed mean RTs, z-score-transformed means, and errors are presented in Table 2. We first compared young with healthy older participants and DAT participants in order to examine effects of aging and dementia. We then examined the usefulness of incongruent error rates in the Stroop switch task for predicting DAT status in the older adults (CDR 0 vs. CDR 0.5) and compared this to the 18 tests in the psychometric battery. Unless otherwise noted, each effect referred to as statistically significant is associated with a two-tailed $p < .05$.

Effects of Age and Dementia

Reaction time. We analyzed RTs using the general linear model with response cue, switch, and Stroop interference conditions varied within-subjects and group as a between-subjects variable. As anticipated, participants were faster to respond to words than to colors and faster to respond to neutral items than to incongruent items; young adults responded faster than healthy older adults, who, in turn, responded faster than DAT individuals. These observations were confirmed by main effects of response

cue, $F(1, 121) = 97.21$, $MSE = 49,001$; Stroop interference, $F(1, 121) = 99.92$, $MSE = 14,901$; and group, $F(2, 93) = 16.42$, $MSE = 449,880$. In addition to these main effects, there were two significant interactions in trimmed RTs and z scores. First, as shown in the top half of Figure 1, the predicted Switch × Response Cue interaction was significant, $F(1, 121) = 16.86$, $MSE = 10,871$, indicating costs when switching to a word response (34 ± 19 ms) but a marginal switch benefit when switching to a color response (−23 ± 24 ms). (Hereafter, when an $X ± Y$ ms effect is reported, Y refers to the 95 % confidence interval.) Although the significant cost when switching to word responses replicates the findings of Allport et al. (1994) and others, the benefit when switching to a color response does not. As can be seen from Figure 1, this benefit when switching to a color response occurred primarily for the very mild DAT group (switch benefit of 50 ± 45 ms), whereas the young and older adults showed no such benefit (switch benefit of 6 ± 5 ms and 15 ± 31 ms, respectively), consistent with the literature. Second, Stroop interference was three times larger when participants were cued to respond to the color dimension (122 ± 29 ms) than when they were cued to respond to the word dimension (42 ± 22 ms), $F(1, 121) = 15.78$, $MSE = 22,047$. This Stroop Interference × Response Cue interaction is shown in the top half of Figure 2. As can be seen, all three groups showed larger Stroop interference (incongruent–neutral) during color naming than word naming. Although less in word naming than color naming, the significant Stroop interference in word naming is of interest, since incongruent color names do not typically produce such interference (MacLeod, 1991). This finding is common, however, on switch trials in the Stroop switching task (Allport et al., 1994) and has been taken as evidence of incomplete reconfiguration of the new task set (i.e., name the word, not the color), allowing for competition from the color dimension.

Errors. Errors were analyzed in the same manner as RTs. The same main effects found for RTs were also obtained in error rates. As can be seen from Table 2, participants made fewer errors to words than to colors and fewer errors to neutral items than to

Table 2
Mean Reaction Time (RT), Z Score (Z), and Percentage of Errors (PE) for Young, Older, and Very Mild DAT Participants as a Function of Task Switch (NSw vs. SW), Response Cue (Color vs. Word), and Stroop Interference (Incongruent vs. Neutral) Conditions

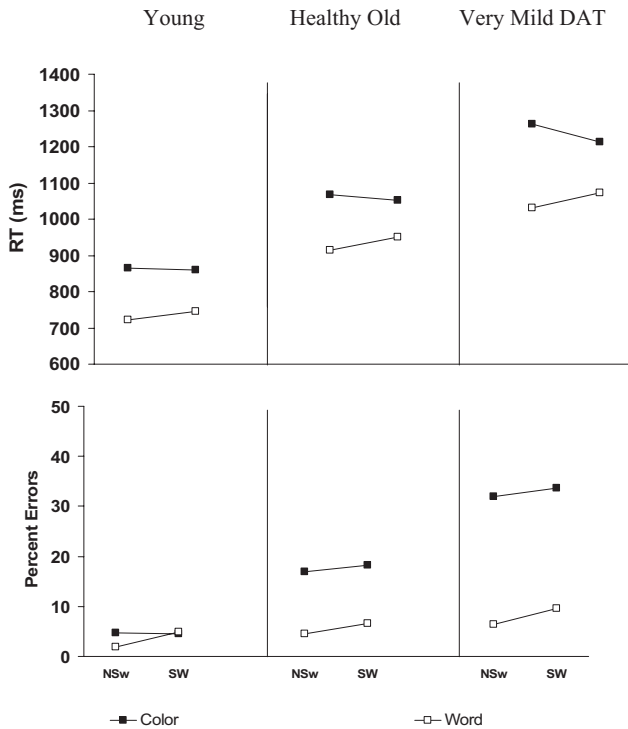| Variable	Young			Healthy older			Very mild DAT		
	RT	Z	PE	RT	Z	PE	RT	Z	PE
Color									
Incongruent									
Switch	905	0.52	5.4	1128	0.48	26.3	1291	0.39	47.9
NonSwitch	896	0.50	6.2	1143	0.50	23.0	1317	0.50	44.4
Neutral									
Switch	812	0.09	3.6	975	−0.02	10.1	1133	−0.02	19.3
NonSwitch	832	0.15	3.2	990	0.03	10.8	1206	0.16	19.3
Word									
Incongruent									
Switch	766	−0.17	4.3	963	−0.09	5.2	1107	−0.01	7.2
NonSwitch	727	−0.30	0.9	929	−0.19	4.3	1069	−0.13	5.2
Neutral									
Switch	725	−0.33	5.2	936	−0.15	7.8	1039	−0.17	11.9
NonSwitch	717	−0.41	2.6	896	−0.28	4.8	992	−0.29	7.3

Figure 1. Mean reaction time (RT) and error rate for young, older, and very mild DAT participants as a function of response cue given on the current trial (color vs. word) and task switch (no-switch [NSw] vs. switch [SW]) conditions. Dark squares indicate trials in which the participant was cued to name the color, whereas white squares indicate trials in which the participant was cued to name the word. A switch trial is one in which the participant responded to the alternate dimension (color vs. word) on the previous trial. Thus, on a switch word trial, the participant must name the word on the current trial but named the color on the previous trial.

incongruent items; young adults made fewer errors (3.9%) than older adults (11.5%), who, in turn, made fewer errors than very mild DAT individuals (20.3%). These observations were confirmed by main effects of response cue, $F(1, 123) = 80.40$, $MSE = 453$; Stroop interference, $F(1, 123) = 82.14$, $MSE = 105$; and group, $F(2, 123) = 26.28$, $MSE = 637$. In addition, there was a main effect of task switching, $F(1, 123) = 12.39$, $MSE = 61$, in which participants made more errors on switch trials than on nonswitch trials.

The RT interactions were replicated in error rates. There was a marginal Switch × Response Cue interaction, $F(1, 123) = 3.75$, $MSE = 48$, $p < .06$, indicating a cost when switching to a word response ($2.7 \pm 1.3\%$) but not when switching to a color response ($0.9 \pm 1.5\%$). These data are shown in the bottom half of Figure 1. There are three interesting aspects of this finding. First, all three groups showed significant costs when switching to a word response (3.0%, 2.0%, and 3.3% for young, healthy older, and very mild DAT groups, respectively). As argued by Allport et al., (1994), the nondominant S–R mapping for color naming requires a more strongly imposed task set. As a result, switch costs (which are created when items associated with the previous task set emerge in a different task) are stronger during word-reading trials. This explains why one typically finds more interference switching

from color naming to word reading than vice versa. Of current debate, however, was whether such interference (typically found in young adults) is due to persisting inhibition or automatic retrieval of the previously irrelevant task set. The lack of any group difference in this pattern supports the hypothesis by Allport and Wylie (2000) that the cost reflects an automatic retrieval of incompatible S–R mappings that must be overcome, rather than a persisting inhibition from control exerted over an incompatible task set on the previous trial. If the latter were true, one would expect greater costs when switching to a word response for young adults and healthy older adults than for DAT individuals.

Second, unlike the RTs, no switch benefit was observed for DAT individuals in switching to a color response. Thus, if the switch benefit for DAT individuals switching to a color response observed in the RT analyses is a real, rather than spurious, effect, then it likely influences response competition resolution rather than goal maintenance. Third, the asymmetrical switch cost effect did not significantly interact with stimulus type ($p > .09$). In fact, for our healthy older adults and DAT participants, there was more of a switch asymmetry shown for the neutral items than for the incongruent items. This is perhaps surprising in that neutral word names should only mildly interfere with color naming (MacLeod, 1991). However, it should be stressed that these items were not truly neutral because both the ink color and the word itself were valid responses, whereas in most Stroop studies the neutral word name is never a valid response. Thus, the neutral words produced
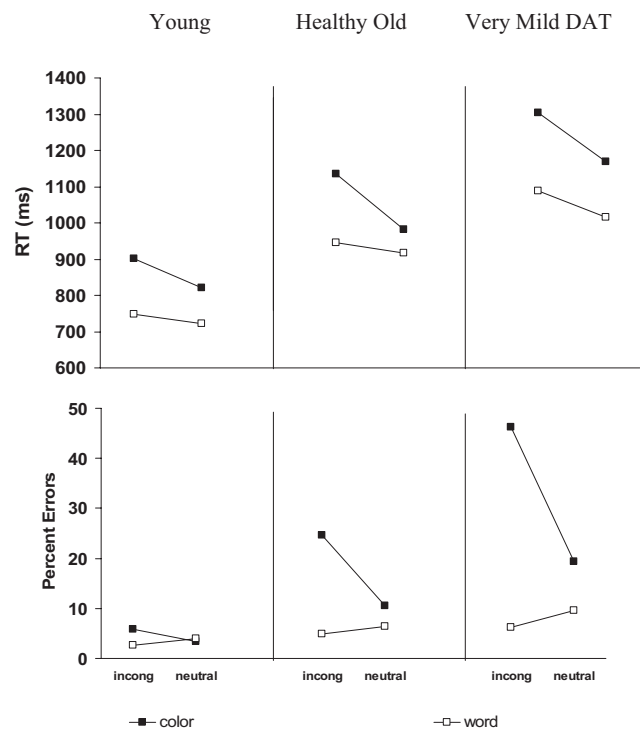

Figure 2. Mean reaction time (RT) and error rate for young, older, and very mild DAT participants as a function of response cue (color vs. word) and Stroop interference (incongruent vs. neutral) conditions. Dark squares indicate trials in which the participant was cued to name the color, whereas white squares indicate trials in which the participant was cued to name the word.

more interference than in most studies, particularly for our older and DAT participants.

There was also a significant interaction between response cue and group in error rates, $F(2, 123) = 19.06$, $MSE = 453$. As can be seen in Figure 1, young adults showed no difference in responding to colors versus words ($1.3 \pm 5.5\%$), whereas older adults did show a difference ($12.0 \pm 3.8\%$), and the very mild DAT group showed a dramatic difference ($24.8 \pm 5.3\%$). Follow-up analyses confirmed that the effect of response cue significantly differed across all three groups. This pattern suggests that young adults can successfully inhibit the more dominant word response in favor of a color response. By contrast, older adults and DAT individuals have much less success inhibiting the dominant word response, likely making fast word responses before a conflict is even detected.

The three interactions involving Stroop interference in error rates can be seen in the bottom half of Figure 2. First, overall Stroop interference (collapsed across response cue) was not significant for young adults ($0.6 \pm 2.6\%$) but increased for older adults ($6.3 \pm 1.8\%$) and for DAT individuals ($11.8 \pm 2.5\%$), $F(1, 123) = 18.39$, $MSE = 105$. Second, as expected, Stroop interference interacted with response cue, $F(1, 123) = 114.36$, $MSE = 134$. Large positive Stroop interference was obtained when participants were cued to respond to the color dimension ($14.5 \pm 2.6\%$), but small negative Stroop interference occurred when participants were cued to respond to the word dimension ($-2.1 \pm 1.2\%$). Finally, both of these two-way interactions were qualified by a significant three-way Response Cue $\times$ Stroop Interference $\times$ Group interaction, $F(2, 123) = 20.43$, $MSE = 134$. Both the positive Stroop interference during color naming and the negative Stroop interference during word naming increased across age and dementia status, though only the interference increase for color naming was significant. For participants' responding to words, the small negative Stroop interference was not significant for either young adults ($-1.3 \pm 2.4\%$) or healthy older adults ($-1.6 \pm 1.7\%$) but was significant for DAT individuals ($-3.4 \pm 2.3\%$). For participants' responding to color, the Stroop interference increased 11.8% from young adults ($2.4 \pm 4.9\%$) to older adults ($14.2 \pm 3.3\%$) and an additional 12.7% from older adults to DAT individuals ($26.9 \pm 4.9\%$). We performed separate post hoc analyses of variance, comparing two groups at a time, and found that both of these increases in Stroop interference for color naming across group were significant. As with the Response Cue $\times$ Group interaction described earlier, this Stroop Interference $\times$ Group interaction replicates earlier findings by Spieler et al. (1996) and Castel et al. (2007) that error rates in the Stroop task can be a particularly useful discriminator between healthy aging and early-stage DAT. Healthy older adults produced fewer errors than DAT individuals in both Stroop interference and response cue effects. However, the findings from this paradigm differ from the earlier studies in that errors also discriminated young from older adults. It is possible that adding the switching component to this congruency task disrupted older adults' ability to maintain the appropriate task set, yet did not cause such problems for young adults. Indeed, the very low overall error rate for young adults ($3.9 \pm 3.3\%$) is impressive given the complexity of this task.

Comparing Incongruent Errors to Psychometric Measures

As mentioned in the introduction, Balota et al. (in press) found that incongruent errors in a Stroop task was the best predictor of which healthy individuals without dementia at the time of testing would become diagnosed with DAT over the next 13 years. The other psychometric measures given to individuals in that study included 12 of the 18 measures used in the current study (shown in Table 1) and included tasks involving episodic memory (e.g., Logical Memory, Associative Recall), simple span (e.g., Forward and Backward Digit Span), spatial abilities (e.g., Benton Copy & WAIR-R Block Design), and processing speed (e.g., Trail-Making Parts A and B). As with Balota et al., we chose to directly compare the usefulness of incongruent error rates for discriminating very mild DAT from healthy aging, $t(94) = 4.35$, $p < .001$, to that of the other psychometric tests (t values shown in Table 1). On the basis of previous work by Balota et al. (in press), Castel et al. (2007), and Spieler et al. (1996), we predicted that errors to incongruent items during the color-naming portion of the current Stroop switch task would be especially sensitive for discriminating healthy aging from the onset of dementia and would perhaps even outperform the currently available psychometric tasks.

The intercorrelations among the 18 psychometric tasks (as well as incongruent error rates in the current study) are presented in Table 3. It is of interest that incongruent error rates produced significant correlations with all 18 psychometric tests. Thus, it is possible that incongruent error rates provide a general assessment of overall cognitive functioning, such as an individual's degree of attentional control. It has been argued elsewhere (Balota et al., 1999; Balota & Faust, 2001) that a single construct such as attentional control may in fact underlie performance deficits in DAT across a wide range of cognitive tasks. Also of interest in Table 3 is the degree of intercorrelations among the psychometric tests themselves. Such intercorrelation among predictors can create the problem of multicollinearity (Cohen, Cohen, West, & Aiken, 2002; Hosmer & Lemeshow, 2000) in interpretation of regression coefficients. Namely, if predicted variance in the dependent measure is highly shared among multiple predictor variables, the contribution of each variable could be underestimated.

In order to diminish the problem of multicollinearity, we used scores on 16 of the 18 tasks (excluding MMSE and NART) to create unit-weighted composite scores that, according to previous research, represent abilities across six cognitive domains. The composite score for each domain was the average z score per participant across each task. These scores (and their component tasks) included a *memory* score (Logical Memory–Immediate, Logical Memory–Delayed, Associative Recognition, Associative Recall, Selective Reminding), a *speed/switching* score (Trail-Making Part A, Trail-Making Part B, Digit–Symbol), a *spatial ability* score (WAIS-R Block Design, Benton Copy), a *digit span* score (Digits Forward, Digits Backward), a *knowledge* score (WAIS-R Information, Boston Naming), and a *verbal fluency* score (Word Fluency, Animal Fluency).

We conducted a logistic regression predicting DAT status (very mild DAT vs. healthy older adult) using incongruent errors (transformed into z scores) plus the composite scores for each of the six cognitive domains (memory, speed/switching, verbal fluency, spa-

Table 3

Intercorrelations Among Predictor Variables Used to Predict Dementia of the Alzheimer's Type

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Incongruent errors	—	-.313*	-.315*	-.471*	-.283*	-.484*	.381*	.579*	-.515*	-.474*	-.313*	-.196*	-.287*	-.317*	-.401*	-.424*	-.394*	-.418*	-.380*
2. Logical Memory—Immediate		—	.938*	.672*	.299*	.381*	-.196*	-.287*	.287*	.381*	.205*	.211*	.252*	.489*	.454*	.262*	.513*	.240*	.302*
3. Logical Memory—Delayed			—	.717*	.319*	.387*	-.149	-.271*	.275*	.392*	.205†	.144	.211*	.488*	.485*	.218*	.487*	.283*	.257*
4. Associative Recognition				—	.254*	.548*	-.232*	-.403*	.411*	.400*	.289*	.200*	.319*	.364*	.482*	.375*	.487*	.389*	.290*
5. Associative Recall					—	.235*	-.112	-.343*	.223*	.384*	.165	.097	.092	.286*	.223*	.171	.153	.652*	.271*
6. Selective Reminding Task						—	-.313*	-.462*	.478*	.297*	.289*	.094	.221*	.213*	.318*	.416*	.501*	.462*	.150
7. Trail-Making (Part A)							—	.632*	-.624*	-.311*	-.355*	-.169	-.157	-.241*	-.189†	-.426*	-.448*	-.266*	-.348*
8. Trail-Making (Part B)								—	-.794*	-.588*	-.380*	-.295*	-.317*	-.264*	-.230*	-.493*	-.559*	-.436*	-.412*
9. Digit/Symbol									—	.482*	.350*	.163	.258*	.310*	.311*	.475*	.652*	.279*	.393*
10. WAIS-R Block Design										—	.363*	.228*	.252*	.468*	.451*	.284*	.409*	.377*	.255*
11. Benton Copy (Form D)											—	.161	.214*	.205†	.393*	.327*	.224*	.277*	.272*
12. Digits Forward												—	.454*	.054	.050	.269*	.275*	.291*	.290*
13. Digits Backward													—	.150	.175†	.366*	.307*	.209*	.291*
14. WAIS-R Information														—	.603*	.217*	.433*	.168	.424*
15. Boston Naming															—	.152	.433*	.273*	.176*
16. Word Fluency																—	.417*	.247*	.435*
17. Animal Fluency																	—	.184*	.331*
18. MMSE																		—	.293*
19. NART																			—

Note. WAIS-R = Wechsler Adult Intelligence Test–Revised; MMSE = Mini-Mental State Examination; NART = National Adult Reading Test.
† $p < .10$. * $p < .05$.

tial ability, digit span, and knowledge).[3] The results of this analysis are shown in Table 4. The correct classification rate was 81%, which was significantly—$\chi^2(7, N = 91) = 27.20$—above the base rate classification of 68% if one simply predicted that all participants in the sample were CDR 0 (68% of all older adults in the sample were CDR 0). As can be seen, both incongruent errors and memory significantly discriminated between CDR 0 and CDR 0.5 individuals.[4]

In addition to that analysis, we conducted two logistic regression analyses to investigate (a) the ability of incongruent color-naming errors to improve predictability of DAT beyond predictions based upon each of the 18 current psychometric tests and (b) the ability of each of the current psychometric tests to improve predictability of DAT beyond predictions based solely on incongruent color-naming errors. All 18 psychometric tests were entered individually in this analysis because (a) comparing incongruent error rates to a single test at a time reduces the multicollinearity problem (i.e., shared variance among multiple predictor variables) that motivated the earlier analysis and (b) the predicted variance of any individual test might be diluted by its inclusion within a general composite score. The results from these two logistic regressions are summarized in Tables 5a and 5b. As can be seen in Table 5a, 12 of the 18 psychometric tests significantly predicted DAT when entered in the initial step of the logistic regression equations. It is important to note that adding incongruent color-naming errors significantly increased the predictability of DAT regardless of which psychometric test was entered first.

Although incongruent errors increased the predictability of DAT above and beyond that of psychometric tests, this was not true for 17 of the 18 psychometric tests when the order of entry was reversed. As can be seen in Table 5b, the Selective Reminding Task was the only task that significantly increased R^2 when entered after incongruent color-naming errors (with marginal improvements from Delayed Logical Memory and Associative Recall). None of the other remaining variables helped to categorize individuals as a DAT versus healthy control individual once incongruent errors were already entered.

In summary, among all psychometric variables, incongruent error rate was the best discriminator of healthy older adults from very mild DAT individuals, followed by the Selective Reminding Task. Of all 18 psychometric variables, only the Selective Reminding Task significantly boosted R^2 when entered after incongruent Stroop errors, whereas Stroop errors boosted R^2 regardless of which task it followed. These results suggest that the current Stroop switch task and episodic memory measures are best for early discrimination of individuals with healthy aging from those with DAT.

General Discussion

The purpose of the present investigation was to examine how well a switching version of the Stroop task could capture attentional control deficits that presumably accompany aging and DAT. The results indicated the Stroop switch task was sensitive to group differences in age and exceeded the performance of current psychometric tests in discriminating healthy aging from early-stage DAT.

As expected, when cued to respond to color, individuals with very mild DAT produced larger Stroop interference in error rates than healthy older adults. This finding replicates those in past studies by Spieler et al. (1996) and Castel et al. (2007) in which very mild DAT individuals showed increased errors in the incongruent Stroop and Simon conditions, respectively. This pattern of errors on incongruent trials has become a signature of early-stage DAT. In the current study, it is likely that the added requirement of switching between naming words and colors weakened participants' ability to consistently ignore words in favor of color naming. This is especially true for healthy older adults and early-stage DAT individuals, who showed considerable increases in errors during color-naming trials relative to young adults (see Figure 2). In contrast to the Spieler et al. (1996) and Castel et al. (2007) results, we found that Stroop interference errors on color-naming trials were also greater for older adults than for young adults. However, as previously mentioned, it is likely that the addition of the switch component to the Stroop task increased the attentional demands of the task and thus disrupted older adults' ability to maintain the appropriate task set.

In contrast to the robust group differences in Stroop interference observed in the current study, there was little-to-no difference in cross-trial switch costs across groups. All three groups showed larger costs when switching to word naming than when switching to color naming. This pattern replicates the interesting switch cost asymmetry first observed by Allport et al. (1994) and lends support to involuntary retrieval accounts of residual switch costs (Allport & Wylie, 1999, 2000; Waszak et al., 2003; Wylie & Allport, 2000) in which stimuli involuntarily invoke previous S–R mappings, rather than to a persisting controlled inhibition of the previously irrelevant (but currently relevant) task (Gilbert & Shallice, 2002; Mayr & Keele, 2000; Yeung & Monsell, 2003). If the latter were true, then one would have expected diminished switch costs among DAT individuals, in whom controlled inhibitory processes are presumably deficient.

Our regression analyses further bolstered the claim that Stroop interference errors are particularly sensitive to early-stage DAT. We used logistic regression to examine the usefulness of incongruent errors in the Stroop switch task for predicting DAT status, and this discrimination ability was compared with that of the 18 psychometric measures. Of particular interest, incongruent error rates were better for discriminating early-stage DAT individuals from healthy controls than any of the standard 18 psychometric tests available for these participants. Moreover, only one of those tasks (the Selective Reminding Task) significantly predicted DAT status when entered into the logistic regression analysis after incongruent errors.

As discussed previously, frontal brain areas such as the PFC and ACC show considerable pathology early in DAT (Killiany et al., 2000; Mintun et al., 2007; van der Flier et al., 2002; Yamaguchi et

---

[3] This analysis excludes data from five (two healthy older individuals and three with very mild DAT) of the 96 total healthy older and very mild DAT participants for whom one or more of the psychometric tests were missing.

[4] For diagnostics for collinearity, all tolerance values for the predictors used in this analysis were greater than 0.4, and none of the variance inflation factors were greater than 2.5 (See Hosmer & Lemeshow, 2000). Thus, we feel confident that the problem of multicollinearity was reduced in this analysis.

Table 4

Logistic Regression Coefficients (β), Standardized Errors (SE), Wald Values, Significance Levels, Odds Ratios, and Odds Ratio 95% Confidence Intervals for Standardized Incongruent Errors Plus Average Z Score Values Across Six Cognitive Domains Used to Discriminate Healthy Older Adults From Those with Very Mild DAT

Predictor variable	β	SE	Wald (z score)	p	Odds ratio	95% Confidence intervals	
						Lower	Upper
Memory (five tasks)	−1.226	.562	4.754*	.029	0.294	0.098	0.883
Speed/switching (three tasks)	0.167	.467	0.128	.720	1.182	0.473	2.952
Verbal fluency (two tasks)	−0.085	.516	0.027	.869	0.919	0.334	2.524
Spatial ability (two tasks)	0.226	.445	0.259	.611	1.254	0.524	3.001
Digit span (two tasks)	0.125	.369	0.115	.734	1.133	0.550	2.335
Knowledge (two tasks)	0.106	.365	0.085	.771	1.112	0.544	2.275
Incongruent Stroop errors	1.088	.377	8.334*	.004	2.969	1.418	6.215

[†] $p < .10$. [*] $p < .05$.

al., 2001). Given that both the Stroop task and task switch paradigms have been shown to activate such frontal regions (Brass et al., 2005; Braver et al., 2003; De Pisapia and Braver, 2006), it is not surprising that a task combining these paradigms is sensitive to early impairments accompanying DAT. An abbreviated version of this task may therefore be useful as part of the diagnostic tools available to clinicians, as suggested by Fine et al. (2008).

It is of interest that in addition to the Stroop switch task, average performance across the five memory tasks also discriminated individuals with very mild DAT from controls. Of course, this finding is also not surprising, since many studies have found evidence for memory measures as early discriminators between

healthy aging and DAT (see Bäckman, et al., 2005, for a review), and medial temporal brain areas show some of the earliest neuropathology linked to DAT (Braak & Braak, 1991). In a recent meta-analysis, Bäckman et al. (2005) reported that the largest effect sizes in discriminating DAT from healthy aging came from tasks measuring episodic memory (e.g., California Verbal Learning Test, Delis, Kramer, Kaplan, & Ober, 1987; Logical Memory tests), executive functioning (e.g., Stroop task, Trail-Making Part B), and perceptual speed (e.g., Digit–Symbol Task, Letter Cancellation Task). The current evidence for the importance of the Stroop switch task and memory measures for early discrimination is therefore in line with the meta-analysis of Bäckman et al.

Table 5a

Logistic Regression Analyses of the Ability of Incongruent Color-Naming Errors to Improve Discrimination of Healthy Aging from DAT Beyond Psychometric Test Performance

Psychometric test	Initial regression model: Initial $\chi^2(1)$	Nagelkerke R^2: Initial R^2	After incongruent color-naming errors were added		
			$\Delta\chi^2$	ΔR^2	Total R^2
Logical Memory[a]	6.70*	.099	16.39*	.215	.314
Delayed Logical Memory	8.30*	.122	15.87*	.205	.327
Associative Recognition	11.30*	.164	12.20*	.155	.319
Associative Recall[a]	7.03*	.105	17.15*	.222	.327
Selective Reminding Task	16.76*	.245	9.65*	.121	.366
Trail-Making (Part A)[a]	0.36	.006	22.16*	.301	.307
Trail-Making (Part B)[a]	8.58*	.126	12.56*	.164	.290
WAIS-R Digit Symbol[a]	6.72*	.100	14.37*	.190	.290
Forward Digit Span[a]	0.23	.003	20.97*	.288	.291
Backward Digit Span[a]	3.87*	.058	17.57*	.236	.294
WAIS-R Block Design[a]	9.20*	.135	13.06*	.169	.304
Benton Copy (Form D)[a]	1.19	.018	20.97*	.285	.303
WAIS-R Information[a]	3.64[†]	.055	17.65*	.237	.292
Boston Naming Test[a]	5.74*	.086	15.85*	.210	.296
Animal Fluency	3.80[†]	.057	17.21*	.232	.289
Word Fluency[a]	6.49*	.096	15.05*	.199	.295
Mini-Mental State Examination	8.15*	.120	14.10*	.184	.304
National Adult Reading Test	2.96[†]	.045	17.95*	.243	.288

Note. DAT = dementia of the Alzheimer's type; WAIS-R = Wechsler Adult Intelligence Test–Revised. The $\Delta\chi^2$ column indicates the increment in chi-square statistics after incongruent color-naming errors were added in the second step of the model.
[a] Denotes tasks used by Balota et al. (in press).
[†] $p < .10$. [*] $p < .05$.

Table 5b

Logistic Regression Analyses of the Utility of Psychometric Tests for Discriminating Healthy Aging From DAT Beyond Incongruent Color-Naming Errors

Variable	Initial regression model: Initial $\chi^2(1)$	Nagelkerke R^2: Initial R^2	After individual psychometric battery predictors were added		
			$\Delta\chi^2$	ΔR^2	Total R^2
Logical memory[a]	20.90*	.287	2.19	.027	.314
Delayed logical memory	20.90*	.287	3.26[†]	.040	.327
Associative recognition	20.90*	.287	2.56	.032	.319
Associative recall[a]	20.90*	.287	3.33[†]	.040	.327
Selective reminding task	20.90*	.287	6.27*	.079	.366
Trail-Making (Part A)[a]	20.90*	.287	1.61	.020	.307
Trail-Making (Part B)[a]	20.90*	.287	0.24	.003	.290
WAIS-R Digit Symbol[a]	20.90*	.287	0.19	.003	.290
Forward Digit Span[a]	20.90*	.287	0.30	.004	.291
Backward Digit Span[a]	20.90*	.287	0.54	.007	.294
WAIS-R Block Design[a]	20.90*	.287	1.35	.017	.304
Benton Copy (Form D)[a]	20.90*	.287	1.26	.016	.303
WAIS-R Information[a]	20.90*	.287	0.38	.005	.292
Boston Naming Test[a]	20.90*	.287	0.69	.009	.296
Animal Fluency	20.90*	.287	0.10	.002	.289
Word Fluency[a]	20.90*	.287	0.63	.008	.295
Mini-Mental State Examination	20.90*	.287	1.34	.017	.304
National Adult Reading Test	20.90*	.287	0.00	.001	.288

Note. DAT = dementia of the Alzheimer's type; WAIS-R = Wechsler Adult Intelligence Test–Revised. The $\Delta\chi^2$ column indicates the increment in chi-square statistics after psychometric tests were added in the second step of the model.
[a] Denotes tasks used by Balota et al. (in press).
[†] $p < .10$.   * $p < .05$.

Within the memory domain, it is of interest that the Selective Reminding Task was especially effective for discriminating DAT. As with the current study, past research has demonstrated that this task can accurately discriminate healthy aging from DAT (Grober et al., 1988; Grober, Lipton, Hall, & Crystal, 2000). In the Selective Reminding Task (Grober et al., 1988), participants are presented 16 stimuli at study and then are queried for the immediate cued recall of each stimulus item. If the participants fail to retrieve the item, they are then given the cue and item together. After a short delay, the participants are given a free recall test for all 16 items and then a cued recall test for those items not recalled during the free recall period.[5] This entire procedure is then repeated two more times (with 16 new items each time). Each participant receives both a free recall score (how many items recalled during the three free recall attempts, 0–48) and a total recall score (free recall score plus how many items recalled when given the cue, 0–48). Only the free recall score was used in the present study (the total recall score did not aid in discriminating DAT).

One reason why this task is especially effective for discriminating DAT may lie in the requirement to repeat the study–test sequences across three separate blocks. This repetition requires subjects on later blocks to recall currently relevant items while suppressing potentially interfering items from earlier blocks. Indeed, an examination of Grober et al.'s (1988) initial data shows that the free recall of healthy older adults improved over trials, whereas the recall of DAT participants did not, increasing the difference between groups across blocks. It is therefore possible that any practice effects among DAT participants are counteracted by the buildup of such proactive interference across blocks (see Tse, Balota, Moynan, Duchek, & Jacoby, 2010, for recent evi-

dence of strong proactive and retroactive interference effects in early-stage DAT).

Neuroimaging researchers have identified the critical importance of prefrontal cortices to memory-monitoring processes such as those needed to avoid proactive interference from earlier trials (Feredoes, Tononi, Postle, & Smith, 2006; Hedden & Yoon, 2006; Postle, Brush, & Nick, 2004). Hence, part of attentional control involves the ability to exclude no-longer-relevant information from entering into the current search set (Unsworth & Engle, 2007). Indeed, tasks designed to measure working memory (e.g., Operation Span or Reading Span Tasks) are often influenced to a large extent by participants' ability to overcome proactive interference from previous trials (Bunting, 2006; Lustig, May, & Hasher, 2001).

Given that the Selective Reminding Task likely involves attentional control in the form of memory monitoring (presumably located in the PFC), it is not surprising that it surpasses other memory tasks in its discrimination of DAT. It is also not surprising, therefore, that this task correlates more strongly with Stroop interference errors in the current study than any other memory task. Sommers and Huff (2003) also obtained a correlation between Stroop performance and memory in a memory task (the Deese (1959) and Roediger & McDermott (1995) [DRM] false-memory paradigm) that required memory monitoring between studied and highly similar but unstudied items. Indeed, this is precisely the account afforded by Balota et al. (1999) in explaining

---

[5] The selective reminding task also includes an additional recognition test.

the relative increased false memory with the DRM stimuli in early-stage DAT compared with healthy control individuals.

In a similar vein, because attentional control requires the ability to maintain a goal across time to modulate competition between relevant and irrelevant information, this ability should rely partially on memory processes. Thus, like the Selective Reminding Task, the Stroop task is not process pure. However, the fact that the Stroop switch and Selective Reminding tasks function so well in discriminating DAT suggests perhaps that it is the process of coordinating attention and memory that is most sensitive to DAT, rather than any one particular cognitive domain or system.

The current results converge with accumulating evidence that errors on incongruent trials are a particularly useful marker for early-stage DAT. The advantage of incongruent errors in discriminating healthy aging and very mild DAT over memory tasks in the present cross-sectional analysis is somewhat surprising. It was expected that the utility of memory tests for predicting DAT status cross-sectionally should be large, because memory deficits are explicitly used in the diagnosis of DAT. However, even under these conditions that should have strongly biased memory measures, incongruent Stroop errors were still the best discriminator of healthy aging from DAT. In this light, our work converges with that of Fine et al. (2008) who emphasized the power of a Stroop switching task to detect early cognitive decline. One could perhaps argue that the Stroop switch task was more difficult than the memory measures and, if difficulty had been equated, memory tasks would have prevailed. However, the nature of what makes a task "difficult" becomes important. We would argue that the most difficult memory tasks are those that require some of the same PFC regions involved in Stroop to monitor memory between currently relevant items and highly similar or recent distractors (Jacoby, 1999; Jennings & Jacoby, 1997; McCabe, Roediger, McDaniel, Balota, & Hambrick, in press; Roediger & McDermott, 1995; Tse et al., 2010). Thus, we believe it is important to consider the coordination of attentional control systems and memory systems in developing a better understanding of the cognitive changes in the earliest detectable forms of DAT.

References

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

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp.421–452). Cambridge, MA: MIT Press.

Allport, D. A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. Treisman (Eds.), *Attention, space, and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford, England: Oxford University Press.

Allport, D. A., & Wylie, G. (2000). Task-switching, stimulus–response bindings, and negative priming. In S. Monsell & J. S. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 35–70). Cambridge, MA: MIT Press.

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.

Bäckman, L., Jones, S., Berger, A.-K., Laukka, E. J., & Small, B. J. (2005). Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology, 19,* 520–531.

Balota, D. A., Tse, C. S., Hutchison, K. A., Spieler, D. H., Duchek, J. M., & Morris, J. C. (in press). Predicting conversion to dementia of the Alzheimer type in a healthy control sample: The power of Stroop performance. *Psychology and Aging.*

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., Dolan, P. O., & Duchek, J. M. (2000). Memory changes in healthy older adults. In E. Tulving & F. I. M. Craik (Eds.). *The Oxford handbook of memory.* (p. 395–409). New York, NY: Oxford University Press.

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.

Benton, A. L. (1963). *The Revised Visual Retention Test: Clinical and experimental applications.* New York, NY: Psychological Corp.

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. *Archives of Neurology, 55,* 326–335.

Blacker, D., Haines, J. L., Rodes, L., Terwedow, H., Go, R. C., Harrell, L. E., . . . Tanzi, R. (1997). APOE-4 and age at onset of Alzheimer's disease: The NIMH genetics initiative. *Neurology, 48,* 139–147.

Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathology, 82,* 239–259.

Brass, M., Ullsperger, M., Knoeche, T. R., von Cramon, D. Y., & Phillips, N. A. (2005). Who comes first? The role of the prefrontal cortex in cognitive control. *Journal of Cognitive Neuroscience, 17,* 1367–1375.

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726.

Bunting, M. F. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 183–196.

Burke, W. J., Miller, J. P., Rubin, E. H., Morris, J. C., Coben, L. A., Duchek, J. M., . . . Berg, L. (1988). The 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, P., Cohen, J., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, 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.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test.* San Antonio, TX: Psychological Corporation.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis–Kaplan Executive Function System (D-KEFS).* San Antonio, TX: Psychological Corp.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate free recall. *Journal of Experimental Psychology, 58,* 17–22.

De Pisapia, N., & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing, 69,* 1322–1326.

Eppinger, B., Kray, J., Mecklinger, A., & John, O. (2007). Age differences in task switching and response monitoring: Evidence from ERPs. *Biological Psychology, 75,* 52–67.

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.

Feredoes, E., Tononi, G., Postle, B. R., & Smith, E. E. (2006). Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 19530–19534.

Filoteo, J. V., Delis, D. C., Massman, P. J., Demadura, T., Butters, N., & Salmon, D. P. (1992). Directed and divided attention in Alzheimer's disease: Impairment in shifting of attention to global and local stimuli. *Journal of Clinical and Experimental Neuropsychology, 14,* 871–883.

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.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44,* 297–337.

Goodglass, H., & Kaplan, E. (1983a). *Boston Diagnostic Aphasia Examination booklet, III: Oral expression. J: Animal naming (Fluency in controlled association).* Philadelphia, PA: Lea & Febiger.

Goodglass, H., & Kaplan, E. (1983b). *Boston Naming Test.* Philadelphia, PA: Lea & Febiger.

Grady, C. L., Haxby, J. V., Horwitz, B., Sundaram, M., Berg, G., & Schapiro, M. (1988). Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer type. *Journal of Clinical and Experimental Neuropsychology, 10, 5,* 576–596.

Grober, E., Buschke, H., Crystal, H., Bang, S., & Dresner, R. (1988). Screening for dementia by memory testing. *Neurology, 38,* 900–903.

Grober, E., Lipton, R. B., Hall, C., & Crystal, H. (2000). Memory impairment on free and cued selective reminding predicts dementia. *Neurology, 54,* 827–832.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance. Interaction of theory and application* (pp. 653–657). Cambridge, MA: MIT Press.

Haxby, J. V., Grady, C. L., Koss, E., Horwitz, B., Schapiro, M., Friedland, R. P., & Rapaport, S. I. (1988). Heterogeneous anterior–posterior metabolic patterns in dementia of the Alzheimer type. *Neurology, 38,* 1853–1863.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20,* 511–528.

Henderson, A. S., Easteal, S., Jorm, A. F., Mackinnon, A. J., Korten, A. E., Christensen, H., . . . Jacomb, P. A. (1995). Apolipoprotein-E allele epsilon-4, dementia, and cognitive decline in a population sample. *Lancet, 346,* 1387–1390.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression.* Hoboken, NJ: Wiley.

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry, 140,* 566–572.

Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item level. *The Quarterly Journal of Experimental Psychology, 61,* 1036–1066.

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

Jennings, J. M., & Jacoby, L. L. (1997). An opposition procedure for detecting age-related deficits in recollection: Telling effects of repetition. *Psychology and Aging, 12,* 352–361.

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.

Killiany, R. J., Gomez-Isla, T., Moss, M., Kikinis, R., Sandor, T., Jolesz, F., . . . Albert, M. S. (2000). *Annals of Neurology, 47,* 430–439.

Kramer, A. F., Hahn, S., & Gopher, D. (1999). Task coordination and aging: Explorations of executive processing in the task switch paradigm. *Acta Psychologia, 101,* 339–378.

Lafleche, G., & Albert, M. S. (1995). Executive function deficits in mild Alzheimer's disease. *Neuropsychology, 9,* 313–320.

Lustig, C., May, C. P., & Hasher, L. (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General, 130,* 199–207.

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

Mathis, S. (1973). *Dementia Rating Scale: Professional manual.* Odessa, FL: Psychological Assessment Resources.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129,* 4–26.

McCabe, D. P., Roediger, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (in press). The relationship between working memory capacity, executive functioning, and general fluid intelligence: Converging evidence for an executive attention construct. *Neuropsychology.*

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., & Storandt, M. (1991). Very mild Alzheimer's disease: Informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology, 41,* 469–478.

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.

Nelson, H. (1982). *The National Adult Reading Test.* Windsor, United Kingdom: NFER-Nelson.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York, NY: Plenum.

Pashler, H. (2000). Task switching and multitask performance. In S. Monsell and J. Driver (Eds.), *Attention and performance Vol. XVIII*: *Control of cognitive processes.* Cambridge, MA: MIT Press.

Perlstein, W. M., Larson, M. J., Dotson, V. M., & Kelly, K. G. (2006). Temporal dissociation of components of cognitive control dysfunction in severe TBI: ERPs and the cued-Stroop task. *Neuropsychologia, 44,* 260–274.

Postle, B. R., Brush, L. N., & Nick, A. M. (2004). Prefrontal cortex and the mediation of proactive interference in working memory. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 600–608.

Reid, W., Broe, G., Creasey, H., Grayson, D., McCusker, E., Bennett, H., . . . & Sulway, M. R. (1996). Age at onset and pattern of neuropsycho-

logical impairment in mild early-stage Alzheimer disease: A study of a community-based population. *Archives of Neurology, 53,* 1056–1061.

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.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124,* 207–231.

Rosser, A., & Hodges, J. R. (1994). Initial letter and semantic category fluency in Alzheimer's disease, Huntington's disease, and progressive supranuclear palsy. *Journal of Neurology, Neurosurgery, and Psychiatry, 57,* 1389–1394.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81,* 174–176.

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.

Spector, A., & Biederman, I. (1976). Mental set and shift revisited. *American Journal of Psychology, 89,* 669–679.

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.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.

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

Tse, C. S., Balota, D. A., Moynan, S. C., Duchek, J. M., & Jacoby, L. L. (2010). The utility of placing recollection in opposition to familiarity in early discrimination of healthy aging and very mild dementia of the Alzheimer's type. *Neuropsychology, 24,* 49–67.

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104–132.

van der Flier, W. M., van den Heuvel, D. M. J., Weverling-Rijnsburger, A. W. E., Bollen, E. L. E., Westendorp, R. G. J., van Buchem, M. A., & Middelkoop, H. A. M. (2002). Magnetization transfer imaging in normal aging, mild cognitive impairment, and Alzheimer's disease. *Annals of Neurology, 52,* 62–67.

Van Selst, M., & Jolicoeur. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology, 47*A, 631–650.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46,* 361–413.

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

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

Woodward, T. S., Bub, D. N., & Hunter, M. A. (2002). Task switching deficits associated with Parkinson's disease reflect depleted attentional resources. *Neuropsychologia, 40,* 1948–1955.

Wu, K. K., Anderson, V., & Castiello, U. (2006). Attention-deficit/hyperactivity disorder and working memory: A task-switching paradigm. *Journal of Clinical and Experimental Neuropsychology, 28,* 1288–1306.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of switch costs. *Psychological Research, 63,* 212–233.

Yamaguchi, H., Sugihara, S., Ogawa, A., Oshima, N., & Ithara, Y. (2001). Alzheimer beta amyloid deposition enhanced by ApoE epsilon 4 gene precede neurofibrillary pathology in the frontal association cortex of nondemented senior subjects. *Journal of Neuropathology and Experimental Neurology, 60,* 731–739.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus–attribute and response–set selection. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 455–469.