Neuropsychology
2009, Vol. 23, No. 4, 411–423

© 2009 American Psychological Association
0894-4105/09/$12.00 DOI: 10.1037/a0015384

Cognitive Functioning in Alzheimer's and Vascular Dementia: A Meta-Analysis

J. L. Mathias and J. Burke
University of Adelaide

Differentiating between Alzheimer's disease (AD) and vascular dementia (VaD) remains difficult but important if existing pharmacological treatments are to provide symptomatic relief in the case of AD or to alter disease progression in the case of VaD. Cognitive assessments play an important role in aiding diagnosis, despite a lack of clear evidence defining the cognitive abilities and tests that best distinguish between the two types of dementia. The current study therefore completed a meta-analysis of research comparing the cognitive abilities of persons diagnosed with AD and VaD. A comprehensive search was undertaken of the PubMed and PsychInfo databases, with 81 studies being eligible for inclusion. Weighted Cohen's *d* effect sizes, percentage overlap statistics, fail-safe *N*s, and confidence intervals were calculated for all cognitive tests. Of the tests that were examined by more than one study, there was one test of perception and one test of verbal memory that showed large and significant group differences. There were an additional 12 tests that may prove useful. However, all cognitive tests were limited in their ability to discriminate between AD and VaD, suggesting that they should be used cautiously and only in conjunction with other information (imaging, medical history) when diagnosing patients.

Keywords: Alzheimer's disease, vascular dementia, meta-analysis, cognitive deficits, differential diagnosis

Supplemental materials: http://dx.doi.org/10.1037/a0015384.supp

Alzheimer's disease (AD) and vascular dementia (VaD) are the most common forms of dementia (Alvarado-Esquivel et al., 2004; Lobo, Saz, Marcos, Dia, & De-la-Camara, 1995). Although still relatively limited, the pharmacological treatments for these dementias differ, highlighting the importance of accurate diagnosis. Moreover, delays in the commencement of treatments for AD can significantly reduce their benefits (Cummings, 2003; Raskind et al., 2000) or limit disease-altering opportunities in the case of VaD (Shi, Perry, Smith, & Friedland, 2000), making early diagnosis important. Despite pathological and definitional differences between AD and VaD, diagnosis remains difficult (Boyle, 2001; Rosenstein, 1998). Cognitive assessments frequently form part of the diagnostic process, albeit in the absence of information regarding the cognitive functions and tests that best differentiate between these forms of dementia. The current study therefore evaluated the utility of cognitive tests in this context.

AD is characterized by an insidious onset of impairments in memory and executive function, together with a range of other cognitive (e.g., aphasia, apraxia, agnosia) and behavioral problems (e.g., agitation, depression, apathy; McKeith & Cummings, 2005; Weiner, Hynan, Bret, & White, 2005). The risk factors for AD include advanced age, low education, low levels of mental and physical activity late in life (Mortimer, Snowdon, & Markesbery, 2003; Scarmeas, Levy, Tang, Manly, & Stern, 2001; Stern et al., 1994), and previous traumatic brain injury (Jellinger, 2004; Szczygielski et al., 2005), as well as a number of vascular-related variables such as hypertension, diabetes, hyperlipidemia, cardiovascular disease, atrial fibrillation, and smoking (Breteler, 2000b). It is not known whether these vascular risk factors directly cause the AD pathology or whether they lead to cerebrovascular pathology, which exacerbate the Alzheimer's dementia (Blennow, de Leon, & Zetterberg, 2006). VaD, on the other hand, is generally characterized by an acute onset with a fluctuating intensity of symptoms and a stepwise decline in cognitive functioning (i.e., memory and two or more other cognitive functions), combined with evidence of cerebrovascular disease (e.g., focal neurological signs; Jagust, 2001; Micieli, 2006; Román et al., 1993; Sachdev, Brodaty, & Looi, 1999). The most common behavioral features of VaD are depression, delusions, and apathy (McKeith & Cummings, 2005). The risk factors for VaD include advanced age, a history of hypertension, stroke (Reichman, Coyne, & Shah, 1993), smoking, diabetes mellitus, atherosclerosis, vasculitis, and/or hyperlipidemia (Hodges & Graham, 2001; Roman, 2003).

AD and VaD have traditionally been thought to result from different etiologies (Roth, 1955; Selkoe, 1991). However, there is increasing evidence to suggest that there may be an association between AD and VaD, possibly as a consequence of the vascular risk factors that are common to both (e.g., hypertension, atherosclerosis, diabetes, smoking) (de la Torre, 2002; Hofman et al., 1997; Kalaria, 2002; Ott et al., 1997, 1999). Moreover, neuropathological data have revealed that over 30% of patients whose diagnosis of AD was confirmed at postmortem also had cerebrovascular disease and that 40% of patients with VaD confirmed at postmortem showed evidence of AD pathology (Kalaria & Ballard, 1999). There are also similarities in the clinical presentation of AD and VaD, including cognitive decline, functional deterioration, and behavioral symptoms (Kalaria, 2002, 2003).

J. L. Mathias and J. Burke, School of Psychology, University of Adelaide, Adelaide, Australia.

Correspondence concerning this article should be addressed to J. L. Mathias, School of Psychology, University of Adelaide, Adelaide, SA 5005, Australia. E-mail: psyj-mat@complex.psych.adelaide.edu.au

While many studies have compared the cognitive functioning of persons with AD and VaD, there are conflicting findings concerning the cognitive functions and tests that best discriminate between the two (Hodges & Graham, 2001). Numerous studies have found that AD is associated with more memory problems than VaD (e.g., Baillon et al., 2003; Golden et al., 2005; Lamar, Swenson, Kaplan, & Libon, 2004; Oosterman & Scherder, 2006) but fewer deficits in executive functioning (e.g., Baillon et al., 2003; Cherrier, Mendez, Dave, & Perryman, 1999; Graham, Emery, & Hodges, 2004; Oosterman & Scherder, 2006; Traykov et al., 2002). Other research has failed to find differences in these and other aspects of cognitive functioning (e.g., Almkvist, Fratiglioni, Aguero-Torres, Viitanen, & Backman, 1999; Fahlander, Wahlin, Almkvist, & Backman, 2002; Taylor, Gilleard, & McGuire, 1996). An early systematic review by Almkvist (1994) found that AD was associated with poorer word finding and more intrusion errors but superior verbal fluency, motor functioning, attention, and executive functioning than VaD. A subsequent systematic review by Looi and Sachdev (1999) found that, despite many similarities between the two groups, persons with AD had poorer long-term memory and better executive functioning than persons with VaD. Thus, there is no clear consensus concerning the cognitive functions or tests that best discriminate between AD and VaD.

A shortcoming of this research is that it relies on statistical significance, which is not only affected by the magnitude of the group differences but also the sample size and, unless effect sizes are calculated, it is not possible to unravel the two. Moreover, it is difficult to directly compare the results of studies that have used different tests to assess the same cognitive ability or studies that have examined samples with different patient characteristics. A meta-analysis offers a solution to this problem by providing an objective and quantitative means by which to directly compare the research findings of existing studies, independently of statistical significance and sample size, thereby providing an important addition to the current literature. The only existing meta-analysis in this area is that of Oosterman and Scherder (2006) whose analysis of Wechsler Adult Intelligence Scales (WAIS) subtest performance by AD and VaD samples was limited in scope.

This study therefore undertook a meta-analysis of research that compared the cognitive performance of samples diagnosed with AD and VaD. If, as has increasingly been suggested (Breteler, 2000; de la Torre, Kalaria, Nakajima, & Nagata, 2002; Hofman et al., 1997; Snowdon et al., 1997), AD and VaD have overlapping aetiologies and neuropathological substrates, it is unlikely that cognitive testing will yield distinct clinical profiles. In contrast, if AD and VaD are the result of different diseases with distinct neuroanatomical consequences, there should be measurable differences. Either way, the findings have important implications, as cognitive testing is an integral part of the diagnostic process.

Method

Literature Search and Inclusion Criteria

The PubMed and PsycINFO electronic databases were searched to identify studies that administered cognitive tests to both AD and VaD patients (refer to Table A, Supplementary materials for search terms). The search was limited to studies published between January 1989 and October 2007 because the term "vascular dementia" was first used in PubMed in 1989.

For a study to be eligible for inclusion into the current meta-analysis, it had to meet the following criteria: (1) it examined a group of participants who were diagnosed with AD *and* one with VaD (excludes studies of mixed dementia, case studies), (2) cognitive tests were administered to both groups (excludes questionnaires), (3) the cognitive tests were *not* used for the diagnosis and classification of participants into the AD and VaD groups (i.e., a test could not be used as both a dependent and independent variable), (4) statistical data enabling the calculation of Cohen's *d* effect sizes were provided (e.g., means and *SD*, results of *t* tests or one-way *F* tests), (5) participants were not known to have any other neurological or psychiatric disorders that may have caused the cognitive deficits, and (5) the study was published in English.

A total of 2,349 potentially relevant studies were identified. The initial search was deliberately kept broad to ensure that all relevant studies were captured. The titles and abstracts of these studies were examined to determine their relevance, with 2,191 papers being deemed ineligible at this stage. Full-text versions of the remaining 158 studies were then screened against the inclusion criteria. The reference lists of these studies were also examined for additional studies, with one extra study being identified using this method. Of the 159 studies retrieved, 78 did not meet one or more of the inclusion criteria and were therefore excluded (10 studies did not administer cognitive tests, 39 did not provide adequate statistical data for the calculation of effect sizes, 23 did not have separate AD and VaD groups, 1 used cognitive tests for the purposes of diagnosis, 8 were reviews). A total of 81 studies were included in the final analysis. Together, these studies used 118 different cognitive tests, some of which yielded multiple test scores.

Data Collection and Preparation

The data extracted from each study included the *study characteristics* (e.g., participant source, matching of samples, diagnostic criteria), AD and VaD *sample characteristics* (e.g., age, education, gender, time since diagnosis), information relating to the c*ognitive tests* (e.g., test name, cognitive function, unit of measurement), and *test data* for each group (means, *SD*s).

Consistent with the most commonly used diagnostic criteria for AD (McKhann et al., 1984), the Mini Mental State Examination (MMSE) was often used when diagnosing dementia. MMSE scores are therefore only reported as descriptive data (refer to Table 1).

Where scores were obtained from different editions of a test (e.g., WAIS), these were combined for the purposes of calculating mean effect sizes. Where participants were readministered tests (to evaluate treatments), only the baseline scores were used (Saletu et al., 1992). Four studies provided data for different subgroups of AD and VaD (Crossley, D'Arcy, & Rawson, 1997; Kitabayashi et al., 2001; Marterer, Danielczyk, Simanyi, & Fischer, 1996; Saletu et al., 1992), necessitating the calculation of effect sizes for each subgroup (weighted by sample size) which were then averaged to provide a single score.

Care was taken to ensure that the data used to calculate an effect size were independent ("one-study, one-vote" principal) and that no single study made a disproportionate contribution to the calculation of an effect size and, therefore, the conclusions that were drawn (Rosenthal & DiMatteo, 2001; Lipsey & Wilson, 2001;

Table 1
Demographic Details for the Participants with Alzheimer's Disease and Vascular Dementia

	AD						VaD					
	$N_{studies}$	$N_{participants}$	*M*	*SD*	Minimum	Maximum	$N_{studies}$	$N_{participants}$	*M*	*SD*	Minimum	Maximum
Sample size	81	4,867	60.1	147.5	9.0	1,309	81	2,263	27.7	23.2	6.0	168
Age (years)	77	4,657	75.0	6.0	60.9	92.2	77	2,126	75.6	5.0	64.6	92.0
Education (years)	52	3,418	10.5	2.3	4.6	15.6	53	1,378	10.4	2.2	4.5	13.8
Premorbid IQ score	8	264	106.2	4.8	99.9	112.4	8	245	105.2	4.7	99.2	111.1
Time since diagnosis (months)	10	1,678	29.4	16.0	4.0	51.6	10	252	35.3	24.2	1.7	80.4
MMSE	59	3,852	19.4	3.9	2.6	26.7	59	1,548	20.0	4.0	2.0	27.5
HIS	5	156	3.2	3.0	1.3	8.5	4	100	7.3	.9	6.4	8.3
GDS score	10	366	5.5	0.9	4.2	6.9	10	255	5.8	.9	4.5	7.5

Note. MMSE = Mini-Mental State Examination; HIS = Hachinski Ischemic Scale; GDS = Geriatric Depression Scale.

Sharpe, 1997). Thus, subtest scores and total scores from the same test could not both be used in the calculation of an effect size.

Each cognitive test was categorized into one of 10 cognitive functions/domains (orientation, attention, perception, memory, verbal ability, construction, concept formation and reasoning, motor ability, executive functioning, and general functioning) to simplify data presentation. This does not imply that there is a general consensus about the specific aspects of cognition that are assessed by these tests or that the tests provide pure measures of a cognitive ability. The test descriptions provided by Lezak, Howieson, and Loring (2004); Strauss, Sherman, and Spreen (2006) and the study authors were used to guide this process.

Effect Size Calculations and Analyses

Cohen's *d* effect sizes (Cohen, 1988), which measure the standardized difference between two group means (Zakzanis, Leach, & Kaplan, 1999), were calculated to measure differences in the cognitive functioning of the AD and VaD groups. A small effect size is defined as $d = .2$, a medium effect as $d = .5$, and a large effect as $d = .8$ (Cohen, 1992), with an effect size of .5 indicating that the average scores for the two groups differ by half of a pooled standard deviation. The percentage overlap (%OL) between the scores of the two groups was also calculated for each effect size (Zakzanis, 2001): where $d = 0$ equates to 100% overlap (i.e., the groups are indistinguishable), $d = 1.0$ equates to 45% overlap, and $d = 3.0$ equates to less than 5% overlap in scores.

Effect sizes were calculated in a multistage process. The first stage involved calculating an effect size for each of the cognitive tests that were used by a study (Zakzanis, 2001). All effect sizes were calculated in such a way that *positive effects* indicate that the AD group performed better than the VaD group and a *negative effect* indicated that the AD group performed more poorly. In most cases, higher scores indicated better performance (e.g., accuracy). However, in cases where a higher score reflected greater impairment (e.g., errors), the direction of the effect sizes was transformed so that a positive effect size still indicated better performance by the AD group, thereby ensuring that all effect sizes could be interpreted consistently.

The effect sizes for all studies that reported a particular test score were then aggregated to calculate a mean effect size (and *SD*) for that test. Before doing so, each effect size was weighted by the inverse of the variance (d_w) to take into account the fact that the reliability of an effect size is affected by sample size. According to Lipsey and Wilson (2001), the inverse of the variance provides a better measure of the precision of an effect size than sample size.

Effect sizes are reported for all tests that were used by two or more studies because they are thought to provide a more reliable measure of group differences than effect sizes calculated from a single study (Rosenthal, 1995). Ninety-five percent confidence intervals (CIs) were additionally calculated to evaluate the statistical significance of an effect size. If a CI does not span zero, this indicates that the true population effect size differs from zero, indicating that there is a significant difference between the performance of the AD and VaD groups.

One problem facing all meta-analyses is that studies with statistically significant results are more likely to be published and, therefore, analyzed. The failure to include unpublished studies with nonsignificant results may lead to effect sizes being overestimated (Wolf, 1986; Zakzanis et al., 1999). A fail-safe *N* (N_{fs}) was therefore calculated, using the method described by Rosenthal (1995) and Orwin (1983), to address this problem. This statistic estimates the number of unpublished studies, which used a specific test but had nonsignificant results, that are needed in order to reduce the Cohen's *d* to a small effect ($d = .2$). Thus, the larger the N_{fs}, the more confident we are about a finding.

Another potential problem with meta-analyses is that they combine data from studies that examined different samples, using a range of cognitive tests, which may threaten the validity of the conclusions that are drawn (Sharpe, 1997; Zakzanis, 2001). This problem was addressed first by clearly defining the study inclusion criteria, such that only studies that used well-defined samples and objective cognitive tests were included in the analysis. Second, separate effect sizes were calculated for the different cognitive tests to ensure that only "like" data were combined. Finally, a number of methodological variables were treated as *moderator variables* in order to evaluate whether between-study differences in these variables affected the study findings (Hunter, Schmidt, & Jackson, 1982; Sharpe, 1997; Zakzanis et al., 1999). Data permitting, between-study variations in age, education, MMSE, time since diagnosis, depression, and premorbid IQ were to be examined for this purpose. In order to analyze these moderator variables, it was necessary to calculate a mean age, educational level, MMSE score, time since diagnosis, depression, and pre-

morbid IQ score for each study, which was achieved by averaging the data from the AD and VaD samples for each study (e.g., $M_{\text{DAT + VaD age}}$).

Data Interpretation

Weighted mean effect sizes (*SD*, min *d*, max *d*) were calculated for each individual cognitive test, as were N_{fs}, %OL, and 95% CIs. The conclusions drawn from this study are based on the combined interpretation of these statistics. Clinically, a test could only be used to assist in differentiating between AD and VaD if there were large group differences (Cohen's d_w) and, consequently, limited overlap in the performance of the two groups (%OL), if the 95% CI did not span zero (the true effect size differed significantly from zero), and it was unlikely that unpublished findings could draw the findings into question (large N_{fs}). Therefore, for the purpose of this meta-analysis, it was decided that in order for a test to be useful in differentiating between AD and VaD, it had to meet the following criteria: (1) have a large effect size (i.e., $d \geq .8$) and, consequently, a limited degree of overlap in the scores of the AD and VaD samples, (2) have a 95% CI that did not span zero, and (3) have a N_{fs} score that was large enough to make it unlikely that there were that number of unpublished studies with nonsignificant findings in existence. As different tests were used with varying frequency, it was decided that the N_{fs} should at least be greater than the number of published studies that had used a particular test (N_{studies}).

Results

Participants

Summary demographic data for the 81 studies included in this meta-analysis are given in Table 1. In total, there were 4,867 persons with AD and 2,240 with VaD. Of the 69 studies that specified the recruitment source, 23 (33%) were from hospital inpatient or outpatient services, 13 (19%) from memory clinics, 10 (14%) from neuropsychological services, 4 (6%) from parishes, 4 (6%) from the general community, 3 (4%) from psychiatric services, 3 (4%) from geriatric assessment programs, 4 (6%) from research institutes or universities, 2 (3%) from long term studies, 2 (3%) from medical practices, and 1 (1%) from nursing homes. Gender was only reported in 60 studies (AD: 55% female, 45% male; VaD: 38% female, 62% male). Many studies ($N = 42$) reported matching samples on demographic variables, such as age and education.

None of the studies had postmortem confirmation of the AD or VaD diagnosis for the entire sample, although three studies (Hargrave, Stoeklin, Haan, & Reed, 2000; Johanson et al., 1990; Marterer et al., 1996) had a AD diagnosis that was either confirmed at postmortem or made on the basis of published diagnostic criteria. Two studies used this method for diagnosing their VaD samples (Johanson et al., 1990; Marterer et al., 1996). The majority of studies ($N_{\text{DAT}} = 76$, $N_{\text{VaD}} = 74$) used published criteria to diagnose AD and VaD. An additional two studies used similar but unpublished diagnostic criteria to diagnose AD (Golden et al., 2005; Rai, Scott, & Beston, 1989), and six used unpublished criteria to diagnose VaD (Almkvist et al., 1999; Carlesimo, Fadda, Marfia, & Caltagirone, 1995; Carlesimo, Sabbadini, Fadda, & Caltagirone, 1995; Gainotti, Marra, Villa, Parlato, & Chiarotti, 1998; Golden et al., 2005; Graham et al., 2004; Rai et al., 1989). The National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorder Association (NINCDS-ADRDA) (1984) was the most widely used published criteria for diagnosing AD ($N = 64$), with other criteria being: the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM*; $N = 10$; American Psychiatric Association [APA], 1980, 1987, 1994), Hachinski Ichaemic Scale ($N = 1$), and Cambridge Mental Disorders of the Elderly Examination (CAMDEX) ($N = 1$) (Roth et al., 1986). The National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherché el Í Enseignement en Neurosciences (NINDS-AIREN) (1993) was the most frequently used diagnostic criteria for VaD ($N = 30$), followed by *DSM* ($N = 19$) (APA, 1980, 1987, 1994), the Alzheimer's disease Diagnostic and Treatment Centers criteria (ADDTC; Chui et al., 1992, 2000) ($N = 13$), Hachinski Ichaemic Scale ($N = 11$), and CAMDEX (Roth et al., 1986) ($N = 1$).

When the demographic characteristics of the AD and VaD samples were compared for the studies that reported this data (see Table 1), the groups were found to be comparable in terms of education, $t(51) = 1.03$, $p = .308$, time since diagnosis, $t(9) = -1.19$, $p = .265$, and depression, $t(9) = -2.05$, $p = .071$. Notably, however, the latter two comparisons were based on very limited data and should be treated cautiously. In contrast, those with AD were significantly younger, $t(76) = -2.15$, $p = .035$, and had lower MMSE scores, $t(58) = -3.27$, $p = .002$, than the VaD participants. However, both of these differences equate to small effects ($d_{\text{age}} = .11$, $d_{\text{MMSE}} = .15$), suggesting that the differences were very modest (i.e., average difference: age = 7 months, MMSE <1 point) and were statistically significant because of the large sample sizes. It is therefore unlikely that these are confounding variables. Premorbid IQ and Hachinski Ischemia Scale (HIS) scores were rarely reported, precluding the possibility of group comparisons on these variables. The HIS measures vascular risk factors, with a HIS score <4 needed for a diagnosis of AD and a score >7 required for a diagnosis of VaD (Hachinski et al., 1975). Although not providing specific data, an additional 10 studies reported that their AD samples met this criterion, and five indicated that their VaD samples met this criterion.

Cognitive Tests

The weighted effect sizes (d_w) for all measures (mean, *SD*, min, max, 95% CIs), grouped by cognitive domain and rank ordered by size, are provided in Tables 2 to 7. Fail-safe *N*s (N_{fs}) and the overlap statistics (%OL) are also provided, as are the number of studies (N_{studies}) that used each test, the number of participants that were assessed ($N_{\text{participants}}$), and the study references.

The cognitive tests were used by between two and 35 studies. The effect sizes for these tests ranged from a minimum of zero (test of orientation, WAIS Vocabulary, word-picture matching, copying task, WAIS Similarities, Go/No Go task, IQ) to maximum of 1 (Emotional Recognition test). Thus, there was considerable variation in the extent to which the performance of the AD and VaD groups differed on these cognitive tests. This is reflected in the overlap statistics (%OL), which indicate that there was 100% overlap between the AD and VaD groups on the least discriminating measures and almost 45% overlap between the two groups for

Table 2
Weighted Mean Effect Sizes for the Tests of Orientation and Attention

	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Orientation											
Orientation	2	113	0	0.3	−0.1	0.4	−0.4	0.4	0	100	32, 58
Attention											
Mental control	6	605	0.5	0.4	0.1	0.9	0.3	0.8	9	67	40, 42, 58, 69, 70, 72
Sustained attention	3	219	0.5	0.7	−0.5	0.8	0.1	0.8	4	67	21, 28, 76
Trail making test A	10	573	0.2	0.4	−0.4	0.6	0.03	0.38	0	85	4, 22, 26, 46, 52, 56, 70, 71, 72, 81
Digit span backwards	13	724	0.2	0.4	−0.4	0.8	0.1	0.4	0	85	2, 14, 15, 16, 23, 24, 25, 28, 41, 46, 56, 66, 76
Digit span	11	1,092	0.2	0.4	−0.3	0.8	0.1	0.3	0	85	1, 11, 26, 27, 32, 51, 58, 60, 62, 73, 79
Corsi cube/spatial span	7	545	−0.2	0.4	−1.2	0	−0.4	0	0	85	1, 11, 24, 51, 56, 72, 76
Trail making test B	11	630	0.1	0.3	−0.2	0.6	−0.06	0.27	0	92	4, 8, 22, 26, 46, 51, 52, 70, 71, 72, 81
Digit span forwards	16	967	0.1	0.4	−0.5	0.6	0	0.2	0	92	1, 2, 14, 15, 16, 23, 24, 25, 28, 41, 46, 56, 66, 71, 72, 76
Digit symbol	9	997	0.1	0.3	−0.3	0.7	−0.1	0.2	0	92	1, 26, 27, 38, 52, 56, 62, 79, 80
Sequential operations	2	101	0.1	1.1	−0.7	0.9	−0.5	0.6	0	92	56, 65

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap.
[a] Study references provided in Appendix 1 (supplementary materials).

the most discriminating measure. Indeed, the latter figure suggests that, at best, the tests are limited in their ability to discriminate between AD and VaD.

Orientation and attention. Only one test of orientation was used by more than one study but there were no differences in the performance of the AD and VaD patients on this measure (see Table 2). Similarly, none of the 10 measures of attention successfully discriminated between the two groups. At best, there were only moderate, albeit significant, differences in the performance of the AD and VaD groups on the Wechsler Memory Scale (WMS) test of Mental Control and a sustained attention task, with patients with AD performing better on both of these tests. Moreover, a number of commonly used clinical tests (e.g., Digit Span, Trail Making Test, visual span) were all associated with small differences and over 85% overlap in the performance of the groups.

Perception. Eight measures of perception were used by a total of 14 studies (see Table 3). Of note, the Emotional Recognition Test was the most discriminating of all of the cognitive tests in this study, with large and significant group differences on this test. Specifically, persons with AD were better able to identify the emotions depicted in photographs of facial expressions than were people with VaD. Moreover, the large N_{fs} suggests that considerable confidence can be placed in this result, because it is unlikely that there would be 8 unpublished studies in existence that have used this test but found nonsignificant (small) group differences. However, there was a large degree of overlap in the scores of the

Table 3
Weighted Mean Effect Sizes for the Tests of Perception

Perception	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Emotional recognition	2	143	1.0	0.0	1.0	1.0	0.6	1.5	8	45	63, 64
Visual search	2	110	0.6	0.4	0.2	0.7	0.2	1.0	4	62	21, 73
Line orientation	2	84	−0.6	0.3	−0.8	−0.4	−1.0	−0.1	4	62	11, 61
Object decision	2	73	0.5	0.3	0.3	0.6	0.0	0.9	3	67	16, 28
Dot counting	2	73	0.3	0.5	−0.1	0.6	−0.2	0.7	1	79	16, 28
Cancellation task	3	734	−0.3	0.2	−0.5	−0.1	−0.5	0.0	1	79	15, 23, 51
Incomplete letters	2	73	−0.1	0.4	−0.4	0.1	−0.6	0.4	0	92	16, 28
Visual organization	3	154	−0.1	0.1	0.1	0.1	−0.4	0.3	0	92	56, 61, 76

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap.
[a] Study references provided in Appendix 1 (supplementary materials).

AD and VaD groups on this test (i.e., ≈ 45%). In addition, there were two tests that showed moderate and significant group differences (visual search, line orientation) but the %OL statistics confirm the need for large effects in this context.

Memory was one of the most commonly assessed cognitive domains, with the individuals in the AD group consistently performing more poorly on all tests of memory (see Table 4). One test, Delayed Story Recall, detected large and significant group differences with a large N_{fs} statistic, such that persons with AD showed more deficits in their ability to freely recall a short story after a delay. Persons with AD and VaD also showed high to moderate differences on a further two tests (total recall on Word List tasks [total words recalled with free and cued procedures over all learning trials], delayed Visual Memory). Other immediate recall and recognition scores from these same tests only showed small to moderate, albeit significant, group differences.

Verbal ability. None of the nine tests of verbal ability, which were used by 2 to 29 studies, successfully discriminated between AD and VaD, with at least 73% overlap in the scores of these groups (see Table 5). Moreover, commonly used clinical tests (e.g., Boston Naming Test, WAIS Vocabulary, Token Test) only revealed modest or no group differences.

Construction. Similarly, tests of construction did not adequately discriminate between AD and VaD (refer to Table 6). The WAIS Block Design subtest was frequently used to assess constructional abilities ($N_{studies} = 15$), as were copying tasks ($N_{studies} = 16$), however neither successfully discriminated between AD and VaD.

Concept formation and reasoning. Of the tests that assessed these constructs, Word Category Association (choosing a picture that provides a conceptual match to a target picture) and Temporal Rule Induction (recognizing temporal sequences of increasing complexity) were best able to differentiate between persons with AD and VaD (see Table 6). Persons with AD performed better on the Word Category Association task but were poorer at Temporal Rule induction. Although significant, the effect sizes were medium in size, with a large overlap in the scores of the two groups. The

Table 4
Weighted Mean Effect Sizes for the Tests of Memory

Memory	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Story recall–delayed	7	712	−0.9	1.1	−2.6	1.0	−1.1	−0.7	24	48	14, 16, 28, 56, 72, 76, 77
Word list–total recall[b]	2	126	−0.7	0.4	−0.9	−0.4	−1.1	−0.3	5	57	41, 71
Visual memory–delayed recall	7	706	−0.7	0.4	−1.5	−0.3	−0.8	−0.5	17	57	14, 16, 22, 28, 56, 76, 77
Word list–cued recall	6	428	−0.6	0.3	−0.9	0.0	−0.8	−0.3	12	62	13, 39, 42, 43, 45, 51
Word list–savings score	2	124	−0.6	0.3	−0.7	−0.3	−1.0	−0.3	4	62	45, 46
Story recall–immediate	9	906	−0.6	0.6	−1.2	0.9	−0.7	−0.4	18	62	16, 28, 32, 56, 58, 70, 73, 76, 77
Word list–delayed recall	20	1,776	−0.6	0.5	−1.8	0.4	−0.66	−0.46	40	62	3, 15, 22, 23, 24, 25, 28, 39, 42, 45, 48, 51, 52, 55, 56, 66, 70, 71, 73, 77
Famous faces–recognition	3	167	−0.5	2.8	−4.8	0.2	−0.9	−0.1	4	67	41, 61
Word list–immediate recall	20	2,296	−0.5	0.5	−1.7	0.1	−0.6	−0.4	30	67	3, 4, 22, 23, 24, 25, 28, 37, 38, 41, 42, 43, 45, 48, 51, 55, 69, 72, 73, 77
Word list–errors	2	281	−0.5	0.3	−0.8	−0.3	−0.8	−0.1	3	67	45, 56
Word list–registration	2	110	−0.5	0.3	−0.7	−0.3	−0.8	−0.1	3	67	15, 52
Word list–recognition	16	1,328	−0.4	0.5	−2.0	0.1	−0.6	−0.3	16	73	15, 22, 23, 24, 28, 41, 42, 45, 51, 52, 68, 69, 70, 71, 72
Visual memory–delayed recognition	2	475	−0.4	0.0	−0.4	−0.4	−0.6	−0.2	2	73	28, 77
Visual memory–immediate recognition	7	876	−0.3	0.4	−0.8	0.2	−0.5	−0.2	3	79	3, 23, 24, 28, 30, 61, 77
Visual memory–immediate recall	11	1,048	−0.2	0.5	−0.8	0.7	−0.4	−0.1	0	85	1, 9, 22, 25, 28, 56, 58, 59, 73, 76, 77
WMS memory quotient	2	50	−0.2	1.3	−0.8	1.0	−0.8	0.4	0	85	7, 58
Memory for designs	3	121	−0.2	0.7	−1.0	0.3	−0.5	0.2	0	85	34, 56, 66
Paired associates	5	328	−0.1	0.6	−1.1	0.6	−0.4	0.1	0	92	1, 34, 56, 58, 70

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap; WMS = Wechsler Memory Scale.
[a] Study references provided in Appendix 1 (supplementary materials). [b] Total words recalled with free and cued procedures over all learning trials. Other scores (e.g., delayed recall refer to scores for individual trials).

Table 5
Weighted Mean Effect Sizes for the Tests of Verbal Ability

Verbal abilities	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Information	7	817	−0.4	0.4	−0.9	0.2	−0.6	−0.2	7	73	1, 27, 32, 52, 56, 58, 80
Picture naming	29	2,699	−0.4	0.4	−1.4	0.4	−0.48	−0.31	29	73	1, 4, 5, 6, 8, 10, 16, 22, 27, 28, 32, 39, 42, 46, 47, 50, 51, 52, 56, 66, 70, 71, 72, 73, 74, 75, 76, 77, 78
Reading	3	262	0.3	0.6	−0.3	0.9	0.0	0.6	1	79	38, 57, 78
Phrase construction	3	188	0.2	0.7	−1.0	0.4	−0.1	0.5	0	85	14, 24, 25
Aphasia battery (BDAE, WAB)	3	214	0.1	1.1	−0.5	1.5	−0.3	0.4	0	92	35, 51, 78
Comprehension	5	574	−0.1	0.4	−0.6	0.3	−0.3	0.1	0	92	18, 26, 32, 78, 80
Token test	8	704	−0.1	0.3	−0.8	0.3	−0.2	0.1	0	92	11, 20, 56, 66, 72, 73, 76, 77
Vocabulary	8	841	0.0	0.4	−0.5	0.9	−0.1	0.2	0	100	6, 13, 26, 27, 32, 34, 56, 80
Word-picture matching	2	58	0.0	0.1	0.0	0.1	−0.5	0.6	0	100	6, 28

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap; WAIS = Wechsler Adult Intelligence Scale (all editions); BDAE = Boston Diagnostic Aphasia Examination; WAB = Western Aphasia Battery.
[a] Study references provided in Appendix 1 (supplementary materials).

most commonly used measures of reasoning (Raven's Progressive Matrices, Card Sorting, WAIS Similarities) also showed poor discrimination.

Motor performance, executive, and general functioning. As seen in Table 7, there were two tests of motor performance, neither of which successfully discriminated between AD and VaD. Likewise, there was no measure of executive functioning or general functioning that would be useful for the purposes of differentiating between groups (see Table 7). While the results for individual studies varied considerably (refer to min and max), when considered together they do not indicate a substantial or significant difference in the performance of the two groups. Moreover, there were no clear outliers that could be removed to reduce this variability.

Moderator Variables

Seventy-seven studies provided information relating to participant's age, 52 provided educational data, 59 provided MMSE scores, 10 provided time-since-diagnosis, 15 provided depression scores ($N = 10$ used the Geriatric Depression Scale (GDS)), and 8 provided estimated premorbid IQ (refer to Table 2). Unfortunately, the limited availability of data for the latter three variables precluded them from further analysis. The relationship between the remaining three moderator variables and the mean effect sizes for the studies that reported these data were assessed using Pearson *r* correlations. Small and nonsignificant correlations were observed for age ($r = .07$, $n = 76$, $p > .05$), education ($r = -.2$, $n = 53$, $p > .05$), and MMSE ($r = -.01$, $n = 58$, $p > .05$) indicating that there was no systematic variation in the effect sizes obtained by the different studies as a consequence of differences in these variables.

Tests Used by Only One Study

The data for cognitive tests that were only used by one study ($N_{studies} = 1$) was additionally examined to determine whether there were any less frequently investigated measures that, if supported by additional research, might prove useful for the purpose of differentiating between AD and VaD. Effect sizes for tests that met the study criteria (Cohen's $d \geq .8$, CIs $\neq 0$, $N_{fs} > N_{studies}$) but were based on the findings of a single study are provided in Table B of the Supplementary materials. In total, there were an additional 12 cognitive tests that showed large and significant differences between the performance of the AD and VaD samples. Four measures (e.g., Picture absurdities, the Blessed Information-Memory-Concentration index, Supraspan spatial sequence learning, Visual pattern completion) showed very large group differences ($d \geq 1.5$) and small %OL statistics (2%–25%). Moreover, the large N_{fs} suggest that it is unlikely that these findings are merely a consequence of publication bias. In the case of the Picture Absurdities test (recognition of absurdities in pictures), Blessed Information-Memory-Concentration index, and Supraspan spatial sequence learning (where patients are required to learn a spatial sequence), persons with AD performed much more poorly than patients diagnosed with VaD. This pattern of performance contrasts with that seen for the test of Visual pattern completion where persons with AD performed better than those with VaD. The AD group also performed significantly more poorly on pursuit motor learning (pursuit tracking task), Graphical sequences (drawing from verbal commands), and Visual association memory (visual mnemonics task), but significantly better on a Maze task (Porteus maze and Gibson spiral maze), WMS working memory index, Tinker Toy (unstructured construction of tinker toys), writing (subtest of the Western Aphasia Battery), and Silhouette Naming (naming silhouettes of common objects and animals).

Discussion

Overall, the data for this meta-analysis were obtained from 81 studies that examined the cognitive functioning of 4,867 persons diagnosed with AD and 2,263 with VaD using a wide variety of

Table 6
Weighted Mean Effect Sizes for the Tests of Construction, Concept Formation, and Reasoning

	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Construction											
Clock drawing–copy	3	186	0.6	1.0	−0.2	1.8	0.3	0.9	6	62	4, 42, 44
WAIS object assembly	4	535	0.5	0.3	0.2	0.7	0.3	0.7	6	67	1, 26, 35, 80
WAIS block design	15	1,617	−0.2	0.5	−0.8	0.6	−0.3	−0.1	0	85	1, 10, 26, 27, 32, 34, 35, 38, 41, 50, 56, 61, 66, 76, 80
Copying drawings	16	1,825	−0.2	0.7	−2.4	0.5	−0.3	−0.1	0	85	3, 4, 14, 16, 23, 24, 25, 27, 28, 48, 50, 53, 56, 73, 76, 77
Clock drawing–free	9	629	0.2	0.6	−0.6	1.1	0	0.3	0	85	4, 15, 31, 33, 36, 41, 42, 44, 53
Copying drawings with landmarks	5	354	0.0	0.2	−0.3	0.1	−0.2	0.2	0	100	3, 14, 23, 24, 25
Concept formation and reasoning											
Word category association	2	58	0.6	0.1	0.5	0.6	0.0	1.1	4	62	13, 28
Temporal rule induction	2	150	−0.5	0.2	−0.6	−0.4	−0.8	−0.1	3	67	3, 24
Card sorting	10	810	0.3	0.3	−0.2	0.6	0.1	0.5	5	79	6, 11, 16, 28, 37, 56, 66, 71, 72, 76
WAIS picture completion	3	550	−0.3	0.2	−0.5	0.0	−0.5	−0.1	1	79	26, 27, 80
Arithmetic	4	634	−0.2	0.3	−0.5	0.2	−0.4	0.0	0	85	26, 27, 49, 80
WAIS picture arrangement	4	495	0.2	0.4	−0.2	0.8	0.0	0.5	0	85	18, 26, 35, 80
Raven's progressive matrices	12	843	−0.1	0.4	−0.9	0.8	−0.2	0.1	0	92	3, 13, 14, 23, 24, 25, 48, 55, 56, 66, 73, 76
WAIS similarities	9	987	0	0.3	−0.3	0.5	−0.1	0.1	0	100	1, 26, 27, 37, 54, 66, 71, 72, 80

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap; WAIS = Wechsler Adult Intelligence Scale (all editions).
[a] Study references provided in Appendix 1 (supplementary materials).

tests. The available data suggested that the AD and VaD samples were well matched in terms of years of education, time since diagnosis, and level of depression but that the AD group was significantly younger and had lower MMSE scores than those diagnosed with VaD. However, the latter differences were very modest (Cohen's $d < .2$) and are unlikely to reflect clinically meaningful group differences in age (7 months) and MMSE performance (<1 point). Important to note, this meta-analysis found that although a wide variety of tests have been used in an effort to detect differences in the cognitive performance of persons diagnosed with AD versus VaD, only a limited number discriminated between the two groups. Notwithstanding this, there were additional measures that, if supported by other research, may prove useful when assessing patients with AD and VaD.

For a cognitive test to be considered useful in this context, it was argued that the test must be able to adequately distinguish between AD and VaD ($d \geq .8$), we should be confident that the true population differences in cognitive performance significantly differ from zero (measured by the 95% CIs), and the conclusions drawn from the research literature are not systematically biased by the tendency to publish statistically significant findings (measured by the N_{fs} statistic).

There was only one test of perception (Emotional Recognition) and one test of verbal memory (Delayed Story recall), which was used by more than one study, that met these criteria. The former test was used by 2 studies that assessed a total of 143 patients and the latter test was used by 7 studies that assessed 712 patients. Whereas persons diagnosed with AD performed better than those with VaD on the test of perception, they performed more poorly on the memory test. Specifically, persons with AD were better able to identify facial expressions and appropriately match emotional expressions to situations but were less able to recall verbal material after a time delay, possibly reflecting the damage to the temporal lobe (and limbic system) that is associated with AD. Patients with AD have previously been observed to have problems recognizing the emotional tone in speech and recognizing the emotions conveyed by facial expressions (Allender & Kasniak, 1989). However, it is also noteworthy that even the most discriminating measure (Emotional Recognition) showed approximately 45% overlap in the scores of the two groups, suggesting that the AD and VaD are not easily differentiated using a single cognitive test. If these tests are to be used in clinical settings, they should be used cautiously and in conjunction with other information, such as medical history, behavioral observations, imaging, and information from relatives, when contributing to a diagnosis.

Also important is the fact that there were many tests that are commonly used in clinical practice that did *not* effectively discriminate between AD and VaD. For example, Digit Span (for-

Table 7
Weighted Mean Effect Sizes for the Tests of Motor Ability, Executive Functioning, and General Functioning

	$N_{studies}$	$N_{participants}$	Mean d_w	*SD*	Minimum *d*	Maximum *d*	95% CI	95% CI	N_{fs}	%OL	Study references[a]
Motor performance											
Reaction time	3	288	−0.3	0.5	−0.6	0.3	−0.6	−0.1	1	73	1, 11, 79
Ideomotor apraxia	3	173	−0.1	0.4	−0.6	0.2	−0.4	0.2	0	92	11, 67, 73
Executive functioning											
Letter fluency	35	2,874	0.2	0.6	−0.6	1.5	0.2	0.3	0	85	3, 4, 5, 7, 8, 11, 12, 16, 17, 23, 24, 25, 27, 28, 37, 39, 41, 42, 46, 48, 51, 52, 53, 54, 55, 56, 57, 66, 69, 70, 73, 74, 75, 77, 81
Stroop color–word	4	206	0.1	0.2	−0.2	0.3	−0.2	0.4	0	92	8, 28, 56, 81
Category fluency	29	2,221	−0.1	0.5	−2.1	0.7	−0.2	−0.01	0	92	3, 4, 5, 6, 8, 11, 12, 15, 16, 17, 22, 23, 28, 32, 41, 46, 48, 52, 53, 55, 56, 57, 71, 72, 73, 74, 75, 76, 81
Go/No Go	2	134	0.0	0.0	0.0	0.0	−0.4	0.4	0	100	37, 54
General functioning											
Performance IQ	4	442	0.0	0.13	−0.17	0.12	−0.27	0.24	0	92	7, 18, 26, 80
Verbal IQ	5	500	0.0	0.5	−0.5	0.7	−0.3	0.2	0	100	7, 18, 26, 38, 80
Full-scale IQ	7	786	0.0	0.3	−0.4	0.4	−0.2	0.2	0	100	1, 7, 18, 22, 26, 57, 80

Note. Mean d_w = weighted mean effect size; Minimum d_w = minimum effect size; Maximum d_w = maximum effect size; CI = confidence interval; N_{fs} = approximate fail-safe *N*; %OL = approximate percentage overlap.
[a] Study references provided in Appendix 1 (supplementary materials).

ward, backward, total), the Trail Making Test (Parts A & B), tests of visual/spatial span, the Digit Symbol and Block Design subtests of the WAIS, picture naming tests, verbal fluency (letter & category), drawing tasks, card sorting tests, and Raven's Progressive Matrices all failed to discriminate between these two forms of dementia. While these tests may assist in diagnosing dementia, they do not discriminate between AD and VaD. Moreover, while delayed verbal memory was impaired in Alzheimer patients, these deficits were more evident when assessed using story recall than the more frequently used word list learning tasks (e.g., Rey Auditory Verbal Learning Test). Immediate story recall was also marginally poorer than immediate list recall in AD, although immediate memory was not as severely affected as delayed memory.

Of note, there were an additional 12 tests that were only used by one study, with samples that ranged between 21 and 1,345 participants, which may prove useful. More specifically, there were five tests of executive functioning (visual pattern completion, pursuit motor learning, graphical sequences, mazes, Tinker Toy), two tests of memory (WMS Working Memory Index, Visual Association memory), and one test each of attention (Supraspan spatial sequence learning in which patients learn a sequence that is longer than their spatial span), perception (Silhoutte Naming), verbal ability (writing), reasoning (Picture Absurdities), and general functioning (Blessed Information-Memory-Concentration index), that showed large to very large group differences (i.e., $.d \geq .8$). Persons with AD performed better than those with VaD on half of these tests. Furthermore, in the case of the Picture Absurdities test, where a patient must identify an impossible situation (e.g., sawing a piece of wood with a saw that is upside-down), there was minimal overlap (approximately 2%) in the performance of the AD and VaD groups, with the AD group performing much more poorly. Additional research is needed to confirm the suitability of these measures for assisting in differentiating between AD and VaD.

A previous systematic review by Looi and Sachdev (1999), which examined research published between 1966 and 1997, reported AD and VaD groups did not differ on measures of attention, language, construction or concept formation/reasoning but that persons with AD performed more poorly than those with VaD on verbal memory (poorer immediate and delayed recall of stories and word lists) and better on tests of executive functioning (e.g., Tinker Toy, Porteus mazes, WCST). These findings have largely been confirmed here, with the additional benefit that the current research has updated the search period, combined data from different studies, and quantified the extent of the AD-VaD group differences. The evidence examined by Looi and Sachdev (1999) for orientation, perception, nonverbal memory and general intelligence was inconclusive, largely due to insufficient or conflicting evidence. Our findings suggest that one measure of perception (Emotional recognition) differed between groups. To our knowledge, the only other quantitative analysis of cognitive differences between AD and VaD is that of Oosterman and Schreder (2006) who undertook a very specific analysis of group differences in WAIS subtest

performance. Their analyses, as did ours, revealed only small to moderate, albeit significant, group differences on Digit Span Backwards, Information and Object Assembly.

As indicated, there was more than 50% overlap in the scores of the AD and VaD groups on the majority of cognitive tests. There are a number of factors that may have contributed to these findings. First, our analysis is reliant upon the accuracy of the diagnoses of AD and VaD that were made in the original studies. Similarities between the diagnostic criteria for AD and VaD may result in patients within these samples being misdiagnosed or there may be cases of mixed dementia, both of which may reduce diagnostic accuracy. Unfortunately, no study in the current meta-analysis had pathological confirmation of AD and VaD for all participants and so it was not possible to examine cognitive differences in samples where diagnostic accuracy was certain. An examination of differences in the cognitive performance of AD and VaD samples whose diagnoses have been confirmed at postmortem is needed.

Second, VaD is a heterogeneous disease with limited consensus regarding the diagnostic criteria (Bakchine & Blanchard, 2004), potentially adding to within-group variability in cognitive performance and making it harder to detect between-groups differences (AD-VaD). Indeed, studies that have compared some of the most commonly used diagnostic criteria (e.g., NINDS-AIREN, DSM, ADDTC) have found limited overlap in the numbers and cases that were identified by these different criteria (Chui et al., 2000; Pohjasvaara, Mantyla, Ylikoski, Kaste, & Erkinjuntti, 2000; Wetterling, Kanitz, & Borgis, 1996). This is thought to reflect differences in the criteria used to define dementia (e.g., *DSM–IV* requires $\geq$1 cognitive deficit; NINDS-AIREN requires $\geq$2 cognitive deficits) and cerebrovascular disease. Thus, there needs to be much greater consensus regarding the criteria by which VaD is to be diagnosed. In addition, VaD needs to be further divided into an agreed set of subtypes (e.g., multiinfarct dementia, ischemic VaD, cortical VaD, small vessel disease, Binswanger's disease) to enable a more fine-grained analysis. However, very few studies report data for different subtypes, thereby precluding such an analysis at this point in time.

Third, it has recently been suggested (de la Torre, 2002, 2004a; Kalaria, 2003; Weller, Cohen, & Nicoll, 2004) that AD and VaD have overlapping aetiologies and neuropathological substrates, which would make it less likely that they would have distinct clinical profiles. If, as suggested by de la Torre (2004b), the neurodegeneration that occurs in both AD and VaD is similar and has a vascular origin, there are likely to be overlapping cognitive deficits, which would also explain the limited differentiation shown by many cognitive tests. Further research on the nature and subtypes of VaD, and the contribution of vascular problems to AD, is needed in order improve our understanding of the relationship between AD and VaD and to determine whether any overlap in etiology contributed to the limited success with which the cognitive tests were able to discriminate between these two forms of dementia.

Study Limitations

Although this meta-analysis was based on a comprehensive and systematic search of the literature, it is possible that some relevant studies were not identified. Broad search criteria were used and the reference lists of all included studies were examined to minimize the likelihood of this. However, there were 39 studies that did not provide sufficient data to enable the calculation of effect sizes and were, therefore, excluded. Our experience has been that writing to the study authors does not yield sufficient additional usable data to warrant such a time-consuming exercise, as data are often discarded after 7 years or researchers cannot be located. For this reason, researchers are encouraged to provide summary descriptive data (means, *SD*, *N*) for *all* measures, regardless of statistical significance, so that effect sizes can be calculated. It is also possible that there are other studies that failed to find significant AD-VaD differences but were not published and, therefore, not included in this study. Fail-safe *N* statistics were calculated in an effort to address this problem.

A meta-analysis involves synthesizing the research findings of studies that may differ methodologically (e.g., age, education, premorbid IQ, time since diagnosis). Theoretically, it is possible to examine the impact of these variables by treating them as moderator variables but this can only be done if the original study reports this information. In fact, most studies did not report information relating to time-since-diagnosis, depression, or premorbid IQ, thereby precluding an analysis of these variables. These data are essential to the evaluation of an individual study and to the integration of findings in the form of a meta-analysis and, if consistently provided, will raise the quality of the research in this field with minimal additional effort.

Finally, a meta-analysis cannot assess the discriminative ability of tests when they are used in combination. It is possible that a multivariate approach would improve the clinical usefulness of these tests. Additional research assessing the discriminative ability of a combination of the best tests that were identified by this study is needed to address this question. In doing so, it would more closely reflect current clinical practice where profiles of performance on multiple tests, rather than single test scores, are used in the diagnostic process. However, it must also be remembered that the discriminative ability of any cognitive test, whether it is administered alone or in combination with others, is limited by its psychometric properties (reliability, validity), which are invariably imperfect. Moreover, it may be unreasonable to expect cognitive tests to discriminate effectively between two disorders that, using current diagnostic criteria, are heterogeneous and frequently misdiagnosed.

Conclusions

In summary, the findings of this meta-analysis suggest that the neuropsychological tests that best discriminate between AD and VaD are the Emotional Recognition task (Shimokawa et al., 2000; Shimokawa et al., 2003), and Delayed Story recall, as assessed by WMS Logical Memory (Wechsler, 1987) Adult Memory and Information Processing Battery (Coughlan & Hollows, 1985), and Babcock Story Recall (Spinnler & Tognoni, 1987). However, it is important to note that none of the tests showed acceptably low overlap between the scores of the two groups to confidently discriminate between the two types of dementia. These cognitive tests must therefore be used cautiously and in conjunction with other diagnostic information, such as medical history, behavioral observations, imaging, and information from relatives, when making a diagnosis. In addition, there are a number of other tests that

may prove suitable for assisting with diagnosis if additional research supports their discriminative ability.

References

Allender, J., & Kasniak, A. W. (1989). Processing of emotional cues in patients with dementia of the Alzheimer's type. *International Journal of Neuroscience, 46,* 146–155.

Almkvist, O. (1994). Neuropsychological deficits in vascular dementia in relation to Alzheimer's disease: Reviewing evidence for functional similarity or divergence. *Dementia, 5,* 203–209.

Almkvist, O., Fratiglioni, L., Aguero-Torres, H., Viitanen, M., & Backman, L. (1999). Cognitive support at episodic encoding and retrieval: Similar patterns of utilization in community-based samples of Alzheimer's disease and vascular dementia patients. *Journal of Clinical and Experimental Neuropsychology, 21,* 816–830.

Alvarado-Esquivel, C., Hernandez-Alvarado, A. B., Tapia-Rodriguez, R. O., Guerrero-Iturbe, A., Rodriguez-Corral, K., & Martinez, S. E. (2004). Prevalence of dementia and Alzheimer's disease in elders of nursing homes and a senior center of Durango City, Mexico. *BMC Psychiatry, 4,* 3.

American Psychiatric Association (APA). (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

APA. (1987). *Diagnostic and statistical manual of mental disorders: DSM–III–R* (3rd ed., rev ed.). Washington, DC: Author.

APA. (1994). *Diagnostic and statistical manual of mental disorders: DSM–IV* (4th ed.). Washington, DC: Author.

Baillon, S., Muhommad, S., Marudkar, M., Suribhatla, S., Dennis, M., Spreadbury, C., et al. (2003). Neuropsychological performance in Alzheimer's disease and vascular dementia: Comparisons in a memory clinic population. *International Journal of Geriatric Psychiatry, 18,* 602–608.

Bakchine, S., & Blanchard, F. (2004). Diagnostic criteria for vascular dementia: A step towards new ones? *Psychogeriatrics, 4,* 127–129.

Blennow, K., de Leon, M. J., & Zetterberg, H. (2006). Alzheimer's disease. *Lancet, 368,* 387–403.

Boyle, P. A. (2001). Decision-making strategies used by neuropsychologists in the differential diagnosis of dementia. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 61*(7-B), 3832.

Breteler, M. M. B. (2000). Vascular risk factors for Alzheimer's disease: An epidemiologic perspective. *Neurobiology of Aging, 21,* 153–160.

Carlesimo, G. A., Fadda, L., Marfia, G. A., & Caltagirone, C. (1995). Explicit memory and repetition priming in dementia: Evidence for a common basic mechanism underlying conscious and unconscious retrieval deficits. *Journal of Clinical and Experimental Neuropsychology, 17,* 44–57.

Carlesimo, G. A., Sabbadini, M., Fadda, L., & Caltagirone, C. (1995). Forgetting from long-term memory in dementia and pure amnesia: Role of task, delay of assessment and aetiology of cerebral damage. *Cortex, 31,* 285–300.

Cherrier, M. M., Mendez, M. F., Dave, M., & Perryman, K. M. (1999). Performance on the Rey-Osterrieth Complex Figure Test in Alzheimer disease and vascular dementia. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology, 12,* 95–101.

Chui, H. C., Mack, W., Jackson, E., Mungas, D., Reed, B. R., Tinkenberg, J., et al. (2000). Clinical criteria for the diagnosis of vascular dementia. *Archives of Neurology, 57,* 191–196.

Chui, H. C., Victoroff, J. I., Margolin, D., Jagust, W., Shankle, R., & Katzman, R. (1992). Criteria for the diagnosis of ischemic vascular dementia proposed by the State of California Alzheimer's Disease Diagnostic and Treatment Centers. *Neurology, 42,* 473–480.

Cohen, J. (1988). *Statistical power analyses for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155–159.

Coughlan, A. K., & Hollows, S. E. (1985). *The adult memory and information processing battery (AMIPB*). Leeds: St. James's University Hospital.

Crossley, M., D'Arcy, C., & Rawson, N. S. B. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology, 19,* 52–62.

Cummings, J. L. (2003). Use of cholinesterase inhibitors in clinical practice: Evidence-based recommendations. *American Journal of Geriatric Psychiatry, 11,* 131–145.

de la Torre, J. C. (2002). Alzheimer disease as a vascular disorder: Nosological evidence. *Stroke, 33,* 1152–1162.

de la Torre, J. C. (2004a). Alzheimer's disease is a vasocognopathy: A new term to describe its nature. *Neurological Research, 26,* 517–524.

de la Torre, J. C. (2004b). Is Alzheimer's disease a neurodegenerative or a vascular disorder? Data, dogma, and dialectics. *Lancet Neurology, 3,* 184–190.

de la Torre, J. C., Kalaria, R., Nakajima, K., & Nagata, K. (Eds.). (2002). Alzheimer's disease: Vascular etiology and pathology. *Annals of the New York Academy of Sciences, 4,* 977.

Fahlander, K., Wahlin, A., Almkvist, O., & Backman, L. (2002). Cognitive functioning in Alzheimer's disease and vascular dementia: Further evidence for similar patterns of deficits. *Journal of Clinical and Experimental Neuropsychology, 24,* 734–744.

Gainotti, G., Marra, C., Villa, G., Parlato, V., & Chiarotti, F. (1998). Sensitivity and specificity of some neuropsychological markers of Alzheimer dementia. *Alzheimer Disease and Associated Disorders, 12,* 152–162.

Golden, Z., Bouvier, M., Selden, J., Mattis, K., Todd, M., & Golden, C. (2005). Differential performance of Alzheimer's and vascular dementia patients on a brief battery of neuropsychological tests. *International Journal of Neuroscience, 115,* 1569–1577.

Graham, N. L., Emery, T., & Hodges, J. R. (2004). Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *Journal of Neurology, Neurosurgery, and Psychiatry, 75,* 61–71.

Hachinski, V., Iliff, L. D., Zilhka, E., Du Boulay, G. H., McAllister, V. L., Marshall, J., et al. (1975). Cerebral blood flow in dementia. *Archives of Neurology, 32,* 632–637.

Hargrave, R., Stoeklin, M., Haan, M., & Reed, B. (2000). Psychopathology and functional impairment in Alzheimer's disease and related dementias. *Journal of Mental Health and Aging, 6,* 201–211.

Hodges, J. R., & Graham, N. L. (2001). Vascular dementia. In J. R. Hodges (Ed.), *Early-onset dementia: A multidisciplinary approach.* New York: Oxford University Press.

Hofman, A., Ott, A., Breteler, M. M. B., Bots, M. L., Slooter, A. J., van Harskamp, F., et al. (1997). Atherosclerosis, apolipoprotein E, and the prevalence of dementia and Alzheimer's disease in the Rotterdam Study. *Lancet, 349,* 151–154.

Hunter, J. E., Schmidt, F. L., & Jackson, G. B. (1982). *Meta-analysis: cumulating research findings across studies.* Beverly Hills, CA: Sage.

Jagust, W. (2001). Untangling vascular dementia. *Lancet, 358*(9299), 2097–2098.

Jellinger, K. A. (2004). Head injury and dementia. *Current opinion in neurology, 17,* 719–723.

Johanson, A., Gustafson, L., Smith, G. J. W., Risberg, J., Hagberg, B., Nilsson, B. (1990). Adaptation in different types of dementia and in normal elderly subjects. *Dementia, 1,* 95–101.

Kalaria, R. (2002). Similarities between Alzheimer's disease and vascular dementia. *Journal of the Neurological Sciences, 203–204,* 29–34.

Kalaria, R. N. (2003). Comparison between Alzheimer's disease and vascular dementia: Implications for treatment. *Neurological Research, 25,* 661–664.

Kalaria, R. N., & Ballard, C. (1999). Overlap between pathology of Alzheimer disease and vascular dementia. *Alzheimer Disease and Associated Disorders, 13,* S115–123.

Kitabayashi, Y., Ueda, H., Narumoto, J., Nakamura, K., Kita, H., & Fukui, K. (2001). Qualitative analyses of clock drawings in Alzheimer's disease and vascular dementia. *Psychiatry and Clinical Neurosciences, 55,* 485–491.

Lamar, M., Swenson, R., Kaplan, E., & Libon, D. J. (2004). Characterizing alterations in executive functioning across distinct subtypes of cortical and subcortical dementia. *The Clinical Neuropsychologist, 18,* 22–31.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological Assessment* (4th ed.). New York: Oxford University Press.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Lobo, A., Saz, P., Marcos, G., Dia, J. L., & De-la-Camara, C. (1995). The prevalence of dementia and depression in the elderly community in a southern European population. The Zaragoza study. *Archives of General Psychiatry, 52,* 497–506.

Looi, J. C., & Sachdev, P. S. (1999). Differentiation of vascular dementia from AD on neuropsychological tests. *Neurology, 53,* 670–678.

Marterer, A., Danielczyk, W., Simanyi, M., & Fischer, P. (1996). Calculation abilities in dementia of Alzheimer's type and in vascular dementia. *Archives of Gerontology and Geriatrics, 23,* 189–197.

McKeith, I., & Cummings, J. (2005). Behavioural changes and psychological symptoms in dementia disorders. *Lancet Neurology, 4,* 735–742.

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 Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology, 34,* 939–944.

Micieli, G. (2006). Vascular dementia. *Neurological sciences: official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology, 27,* S37–39.

Mortimer, J. A., Snowdon, D. A., & Markesbery, W. R. (2003). Head circumference, education and risk of dementia: Findings from the Nun Study. *Journal of Clinical and Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 25,* 671–679.

Oosterman, J. M., & Scherder, E. J. (2006). Distinguishing between vascular dementia and alzheimer's disease by means of the WAIS: A meta-analysis. *Journal of Clinical and Experimental Neuropsychology, 28,* 1158–1175.

Orwin, R. G. (1983). A fail-safe N for effect size. *Journal of Educational Statistics, 8,* 157–159.

Ott, A., Breteler, M. M. B., de Bruyne, M. C., van Harskamp, F., Grobbee, D. E., & Hofman, A. (1997). Atrial fibrillation and dementia in a population-based study. The Rotterdam Study. *Stroke, 28,* 316–321.

Ott, A., Stolk, R. P., van Harskamp, F., Pols, H. A., Hofman, A., & Breteler, M. M. (1999). Diabetes mellitus and the risk of dementia; The Rotterdam Study. *Neurology, 53,* 1937–1942.

Pohjasvaara, T., Mantyla, R., Ylikoski, M. A., Kaste, M., & Erkinjuntti, T. (2000). Comparison of different clinical criteria (*DSM–III,* ADDTC, ICD-10, NINDS_AIREN, *DSM–IV*) for the diagnosis of vascular dementia. *Stroke, 31,* 2952–2957.

Rai, G. S., Scott, L., & Beston, B. (1989). Assessment of the objective value of Mini-Mental State Examination. *Journal of Clinical and Experimental Gerontology, 11,* 41–46.

Raskind, M. A., Peskind, E. R., Wessel, T., Yuan, W. (2000). Galantamine in AD: A 6-month randomized, placebo-controlled trial with a 6-month extension. The Galantamine USA-1 Study Group. *Neurology, 54*(2261–2268).

Reichman, W., Coyne, A. C., & Shah, A. (1993). Diagnosis of multi-infarct: Predictive value of clinical criteria. *Perceptual and Motor Skills, 76,* 793–794.

Roman, G. (2003). Vascular dementia: Distinguishing characteristics, treatment, and prevention. *Journal of the American Geriatrics Society, 51,* S296–S304.

Román, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H., et al. (1993). Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology, 43,* 250–260.

Rosenstein, L. D. (1998). Differential diagnosis of the major progressive dementias and depression in middle and late adulthood: A summary of the literature of the early 1990s. *Neuropsychology-Review, 8,* 109–167.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183–192.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 59–82.

Roth, M. (1955). The natural history of mental disorder in old age. *British Journal of Psychiatry, 101,* 281–301.

Roth, M., Tym, E., Mountjoy, C. Q., Huppert, F. A., Hendrie, H., Verma, S., et al. (1986). CAMDEX. A standardised instrument for the diagnosis of mental disorder in the elderly with special reference to the early detection of dementia. *The British Journal of Psychiatry, 149,* 698–709.

Sachdev, P. S., Brodaty, H., & Looi, J. C. L. (1999). Vascular dementia: Diagnosis, management and possible prevention. *Medical Journal of Australia, 170,* 81–85.

Saletu, B., Anderer, P., Fischhof, P. K., Lorenz, H., Barousch, R., & Bohmer, F. (1992). EEG mapping and psychopharmacological studies with denbufylline in SDAT and MID. *Biological Psychiatry, 32,* 668–681.

Scarmeas, N., Levy, G., Tang, M. X., Manly, J., & Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer's disease. *Neurology, 57,* 2236–2242.

Selkoe, D. J. (1991). Alzheimer's disease. In the beginning. *Nature, 354*(6353), 432–433.

Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analysis will not go away. *Clinical Psychology Review, 17,* 881–901.

Shi, J., Perry, G., Smith, M. A., & Friedland, R. P. (2000). Vascular abnormalities: The insidious pathogenesis of Alzheimer's disease. *Neurobiology of Aging, 21,* 357–361.

Shimokawa, A., Yatomi, N., Anamizu, S., Ashikari, I., Kohno, M., Maki, Y., et al. (2000). Comprehension of emotions: Comparison between Alzheimer type and vascular type dementias. *Dementia and Geriatric Cognitive Disorders, 11,* 268–274.

Shimokawa, A., Yatomi, N., Anamizu, S., Torii, S., Isono, H., & Sugai, Y. (2003). Recognition of facial expressions and emotional situations in patients with dementia of the Alzheimer and vascular types. *Dementia and Geriatric Cognitive Disorders, 15,* 163–168.

Snowdon, D. A., Greiner, L. H., Mortimer, J. A., Riley, K. P., Greiner, P. A., & Markesbery, W. R. (1997). Brain infarction and the clinical expression of Alzheimer disease. The Nun Study. *Journal of the American Medical Association, 277,* 813–817.

Stern, R. G., Mohs, R. G., Davidson, M. R., Schmeidler, J., Silverman, J., Kramer-Ginsberg, E., Searcey, T. (1994). A longitudinal study of Alzheimer's disease: Measurement, rate, and predictors of cognitive deterioration. *American Journal of Psychiatry, 151,* 390–396.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.

Szczygielski, J., Mautes, A., Steudel, W. I., Falkai, P., Bayer, T. A., & Wirths, O. (2005). Traumatic brain injury: Cause or risk of Alzheimer's disease? A review of experimental studies. *Journal of Neural Transmission, 112,* 1547–1564.

Taylor, R., Gilleard, C. J., & McGuire, R. J. (1996). Patterns of neuropsychological impairment in dementia of the Alzheimer type and multi-infarct dementia. *Archives of Gerontology and Geriatrics, 23,* 13–26.

Traykov, L., Baudic, S., Thibaudet, M. C., Rigaud, A. S., Smagghe, A., & Boller, F. (2002). Neuropsychological deficit in early subcortical vascular dementia: Comparison to Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 14,* 26–32.

Wechsler, D. (1987). *Wechsler memory scale revised manual.* San Antonio, TX: The Psychological Corporation.

Weiner, M. F., Hynan, L. S., Bret, M. E., & White, C. (2005). Early behavioral symptoms and course of Alzheimer's disease. *Acta Psychiatrica Scandinavica, 111,* 367–371.

Weller, R. O., Cohen, N. R., & Nicoll, J. A. (2004). Cerebrovascular disease and the pathophysiology of Alzheimer's disease: Implications for therapy. *Panminerva Medica, 46,* 239–251.

Wetterling, T., Kanitz, R., & Borgis, K. (1996). Comparison of Different Diagnostic Criteria for Vascular Dementia (ADDTC, *DSM–IV,* ICD-10, NINDS-AIREN), *Stroke, 27,* 30–36.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis.* Newbury, London: Sage.

Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology, 16,* 653–667.

Zakzanis, K. K., Leach, L., & Kaplan, E. (1999). *Neuropsychological differential diagnosis.* Lisse: Swets and Zeitlinger Publishers.

Received February 15, 2008
Revision received December 12, 2008
Accepted December 12, 2008 ■

E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at http://notify.apa.org/ and you will be notified by e-mail when issues of interest to you become available!