NEUROPSYCHOLOGIA

Neuropsychologia 42 (2004) 1212–1222

www.elsevier.com/locate/neuropsychologia

Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis

Julie D. Henry*, John R. Crawford, Louise H. Phillips

School of Psychology, King's College, University of Aberdeen, AB24 3HN Scotland, UK

Received 3 October 2003; received in revised form 3 February 2004; accepted 4 February 2004

Abstract

A meta-analysis of 153 studies with 15,990 participants was conducted to compare the magnitude of deficits upon tests of phonemic and semantic fluency for patients with dementia of the Alzheimer's type (DAT) relative to healthy controls. As has been found for patients with focal temporal cortical lesions (but not for patients with focal frontal cortical lesions), DAT patients were significantly more impaired on tests of semantic relative to phonemic fluency ($r = 0.73$ and 0.57, respectively). Thus, since phonemic and semantic fluency are considered to impose comparable demands upon executive control processes such as effortful retrieval, but the latter is relatively more dependent upon the integrity of semantic memory, these results suggest that the semantic memory deficit in DAT reflects a degradation of the semantic store. Also supporting this conclusion, confrontation naming, a measure of semantic memory that imposes only minimal demands upon effortful retrieval, was significantly more impaired than phonemic fluency ($r = 0.60$ versus 0.55, respectively). However, since semantic fluency was also significantly more impaired than confrontation naming ($r = 0.73$ versus 0.61), deficits in semantic memory and effortful retrieval may be additive. Semantic, but not phonemic fluency, was significantly more impaired than measures of verbal intelligence and psychomotor speed. Thus, the semantic memory deficit in DAT qualifies as a differential deficit, but executive dysfunction as indexed by phonemic fluency does not constitute an additional isolated feature of the disorder. Dementia severity was not significantly related to the *relative* magnitude of deficits upon phonemic and semantic fluency.

Keywords: Semantic memory; Executive functioning

1. Introduction

Episodic amnesia is widely regarded as the earliest and most severe deficit associated with dementia of the Alzheimer's type (DAT, see, e.g.; Hodges, Salmon, & Butters, 1990). Although there is a great deal of evidence that DAT is also characterised by a prominent deficit in semantic memory (Hodges, Salmon, & Butters, 1992), there remains widespread disagreement as to what this reflects. The most prominent position is that the deficit is attributable to degradation in the structure or content of semantic knowledge. Whilst some argue that semantic memory is abnormally organised but intact (Grober, Buschke, Kawas, & Fuld, 1985), others suggest that representations are actually lost (Chertkow & Bub, 1990; Hodges et al., 1992). In either case, it is presumed that impoverishment of conceptual knowledge will result, as defining attributes become inaccessible, and the associational strength between related concepts weakened.

However, an alternative position that has been suggested is that the deficit in semantic memory reflects a non-linguistic deficit in controlled cognitive processing that disturbs access to semantic knowledge from a relatively preserved semantic store (Balota & Ferraro, 1996; Bayles, Tomoeda, Kaszniak, & Trosset, 1991; Grande, McGlinchey-Berroth, Milberg, & D'Esposito, 1996; Nebes, 1989, 1994; Nebes, Martin, & Horn, 1984; Ober & Shenaut, 1988). Since executive processes are responsible for the more supervisory, or strategic aspects of cognition (Crawford & Henry, in press; Stuss & Benson, 1986), this latter perspective suggests that the semantic memory deficit is attributable to a breakdown in the executive control mechanisms responsible for effortful retrieval. In the present study we will address this issue by applying meta-analytic techniques to the verbal fluency literature in DAT. As will be discussed, comparison of the relative magnitude of deficits on phonemic and semantic fluency may be used to draw inferences regarding the prominence of effortful retrieval deficits and

* Corresponding author. Tel.: +44-1224-273483; fax: +44-1224-273426.
E-mail address: j.d.henry@abdn.ac.uk (J.D. Henry).

0028-3932/$ – see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2004.02.001

semantic store degradation, and, when in the context of performance on other cognitive measures, whether these deficits qualify as differential deficits (Henry & Crawford, in press).

1.1. Verbal fluency deficits in DAT

Tests of verbal fluency are amongst the most widely employed measures used to assess cognitive functioning following neurological damage, and involve associative exploration and retrieval of words based on phonemic or semantic criteria (phonemic and semantic fluency, respectively), usually conducted in the setting of a time constraint. Thus, whilst for phonemic fluency participants are asked to generate as many words as possible beginning with a specified letter (e.g. F), for semantic fluency search is constrained by a specified category (e.g. animals). These measures are considered to impose comparable demands upon executive or supervisory processes because both require efficient organisation of verbal retrieval and recall, as well as self-monitoring aspects of cognition (the participant must keep track of responses already given), effortful self-initiation, and inhibition of responses when appropriate (Crawford & Henry, in press; Rosser & Hodges, 1994). However, whilst phonemic fluency requires the creation of search strategies based primarily on lexical representations, tests of semantic fluency require searching for semantic extensions of a target superordinate, and thus depend intrinsically upon the integrity of semantic associations within the lexicon (Rohrer, Salmon, Wixted & Paulsen, 1999). Deficits on tests of semantic fluency may therefore reflect problems with semantic memory, and not executive dysfunction.

In a recent meta-analytic review that included 31 studies and 1791 participants, Henry and Crawford (in press) investigated the relative magnitude of cognitive deficits upon tests of phonemic and semantic fluency for patients with focal cortical lesions. The pattern of results strongly suggested that whilst the two types of fluency impose comparable demands upon executive functioning, semantic fluency is relatively more dependent upon the integrity of semantic memory. Henry and Crawford (in press) found that focal frontal lobe injuries were associated with equivalent phonemic and semantic fluency deficits ($r = 0.52$ and 0.54, respectively). Whilst it is important not to treat the terms ‘frontal’ and ‘executive’ as interchangeable, there is nevertheless a great deal of evidence that frontal structures are particularly implicated in executive functioning (see; Stuss & Benson, 1986), and thus comparable impairment upon tests of phonemic and semantic fluency for patients with focal frontal cortical injuries suggests that the two measures impose equivalent demands upon executive control processes such as effortful retrieval. However, semantic fluency was more impaired following focal temporal damage ($r = 0.61$), and this deficit was substantially larger than the corresponding phonemic fluency deficit ($r = 0.44$). Since there is a great deal of evidence that temporal structures are the neural substrates particularly responsible for semantic memory this was presumed to reflect the greater reliance of semantic fluency upon the integrity of semantic memory. Comparison of the relative magnitude of deficits on phonemic and semantic fluency may therefore be used to draw inferences regarding the prominence of effortful retrieval deficits and semantic store degradation, respectively (see also; Rohrer et al., 1999).

Thus, of particular importance in the debate regarding whether the semantic memory deficit in DAT reflects a degradation of the semantic store or retrieval difficulties, has been the pattern of disproportionately impaired semantic relative to phonemic fluency often reported (Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Monsch et al., 1992; Monsch et al., 1997; Rohrer et al., 1999). Whilst most studies of DAT have found both phonemic and semantic fluency to be significantly impaired, a deficit for the latter is more consistently reported, and is typically the more prominent of the two. Monsch et al. (1997) for instance, found that whilst both semantic and phonemic fluency were significantly impaired, when these values were converted to standard scores the deficit for the former was nearly twice as large.

However, the opposite deficit pattern, or equivalent performance on both fluency measures has also, although less frequently, been reported (Nebes et al., 1984; Ober, Dronkers, Koss, Delis, & Friedland, 1986; Suhr & Jones, 1998) suggesting that the semantic memory deficit associated with DAT may instead reflect a disruption of the executive control processes responsible for effortful retrieval, and not a degradation of the semantic store (Henry & Crawford, in press). Indeed, there is considerable evidence that executive deficits are also an important feature of the cognitive deterioration in DAT (Amieva, Phillips, Della Sala, & Henry, in press) and that these deficits typically occur early in the disease, and may be the first non-memory deficits to occur (Perry, Watson, & Hodges, 2000).

To the present authors’ knowledge, the current paper is the first to apply meta-analytic techniques to compare performance upon phonemic and semantic fluency in DAT. The results of this study will be particularly useful because they emphasise the magnitude of effects. Although researchers are strongly encouraged to report effect sizes for their individual studies (American Psychological Association, 2001), this is rarely done in practice, yet is far more informative than simply reporting whether a particular effect is significant or not. Moreover, because using meta-analysis it is possible to integrate effects across studies that differ in both the participants sampled and methodology employed, the effects can be considered to be very reliable, robust estimates of the corresponding parameters of interest. Thus, an effect’s generalisability can be subjected to a level of scrutiny not possible in a single study, and with a level of objectivity and methodological consistency that is difficult to achieve in non-quantitative reviews (Stanley, 2001).

1.2. Aims of the current meta-analysis

The first aim was to derive effect size estimates for phonemic and semantic fluency for patients with DAT relative to healthy controls. Although there is a trend for semantic fluency to be more impaired than phonemic fluency, inconsistencies have been noted. Thus, it will be extremely informative to derive reliable estimates of the magnitude of the effects associated with each of these measures. As has been discussed, comparison of the relative magnitude of each will permit an assessment of whether the semantic memory deficit associated with DAT predominantly reflects a disrupted/degraded semantic store, or difficulties with effortful retrieval.

The second aim was to estimate effect sizes for other cognitive measures in order to provide comparison standards, and thus assess to what extent fluency deficits in DAT qualify as differential deficits. Of particular importance was to address the possibility that phonemic and semantic fluency deficits reflect a more general deficit in verbal intelligence (see; Miller, 1984). Thus, the pattern of deficits across fluency versus verbal intelligence as measured by the WAIS (Wechsler, 1981) verbal and vocabulary scales (VIQ) will be compared. Vocabulary has been permitted to contribute to the construct of VIQ because it is the scale that correlates most highly with the Total score on the VIQ scale; the average correlation across nine age-ranges was 0.85 (Wechsler, 1981). It will also be investigated whether deficits on semantic and phonemic fluency are in excess of deficits on the Trail Making Test Part A (TMT A; Reitan, 1990), a widely used measure of psychomotor speed; this will address the possibility that deficits on tests of verbal fluency reflect generalised slowing (i.e. a reduction in cognitive speed). Performance on tests of phonemic and semantic fluency will also be compared with scores on the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983), a measure of semantic memory that imposes only minimal demands upon effortful retrieval, and Wechsler and Stone's (1974) Immediate Logical Memory (ILM) test of episodic memory.

Finally, a third issue relates to whether the relative prominence of retrieval deficits and semantic store degradation is comparable at each stage of the disease. Lack of consistency across studies in terms of the relative prominence of deficits upon tests of phonemic and semantic fluency has been attributed to differences in dementia severity between patient groups (Butters et al., 1987; Margolin, Pate, Friedrich, & Elia, 1990). It has, for instance, been suggested that whilst executive deficits are present in the early stages of the disorder, they become more pronounced as the illness progresses (Duke & Kaszniak, 2000), which would predict that deficits on the two measures should become more comparable with increasing disease severity. To address this possibility, correlational analyses will be conducted to test whether the difference in the relative magnitude of semantic and phonemic fluency deficits is related to dementia severity, as measured by the mini mental state exam (MMSE; Folstein, Folstein, & McHugh, 1975).

2. Method

2.1. Sample of studies

A computer-based search involving the *Web of Science*, *Psych Lit CD-ROM*, and *Science Direct* databases was undertaken, using the following terms; 'letter fluency', 'FAS', 'semantic fluency', 'category fluency', 'controlled oral word association', 'COWA(T)', 'word fluency' 'verbal fluency', 'oral fluency', 'phonemic fluency', 'executive test' and 'frontal test'. In addition, a manual search of issues of *Neuropsychologia*, *The Journal of the International Neuropsychological Society*, *Neuropsychology*, *The Clinical Neuropsychologist*, *Neuropsychiatry Neuropsychology and Behavioural Neurology*, *Journal of Neuropsychiatry and Clinical Neurosciences*, *Brain*, and the *Journal of Clinical and Experimental Neuropsychology* was conducted. The search was completed in December 2002.

The inclusion criteria were firstly that the patient group had to consist entirely of adults with DAT. Most studies (84%) employed National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA; McKhann et al., 1984) criteria for possible or probable DAT. For the remaining studies diagnoses were based on other accepted criteria for DAT, the opinions of neurologists, or neuropsychological assessment. In addition, the study had to include a healthy control group, a measure of phonemic and/or semantic fluency, and have presented precise statistics convertible to effect size r.

2.2. Statistical analysis

Meta-analysis is a rigorous alternative to the traditional review process, as it involves statistical integration of results. The basis of this methodology is the effect size, a standardised statistic that quantifies the magnitude of an effect. In the present study the effect size r was used, which corresponds to the degree of correlation between group membership (i.e. presence or absence of DAT), and performance on the cognitive measure of interest.

For each construct, effects were pooled to derive an estimate of the mean, with each effect weighted for sample size to correct for sampling error. To do so, the random effects meta-analytic model (Shadish & Haddock, 1994) was selected in preference to the more commonly employed fixed effects model as it yields more generalisable parameter estimates. The National Research Council (1992) argues that the fixed effects model should be the exception rather than the rule, as it may lead to inappropriately strong conclusions. Thus, although more technically demanding than the fixed effects model, it was considered important to use the random effects model in the present work.

The statistic Q quantifies within-group heterogeneity (i.e. the degree to which the studies contributing to each respective mean can be regarded as homogenous). If the Q statistic associated with a mean effect is significant, this suggests that there are substantive differences between the studies contributing to that particular mean. In contrast, a non-significant estimate of Q suggests that once sampling error has been removed, no substantive differences between the studies contributing to the respective mean in question remain (i.e. the null hypothesis of homogeneity of effects cannot be rejected).

It was also important to test whether the difference in the magnitude of mean effects between, for instance, phonemic versus semantic fluency, was statistically significant. However, there is no agreed method for statistically comparing mean effects using the random effects meta-analytic model. A particular difficulty is whether the degrees of freedom (d.f.) in such analyses should be based on N (the number of participants) or K (the number of studies). In the present work, a relatively large number of studies were included, and therefore, t-tests were computed using the more conservative K as the d.f.

Since dementia severity will moderate the magnitude of deficits across individual studies, for each statistical comparison, only studies that assessed both types of fluency were included. For example, although in total 143 studies assessed phonemic fluency, and 126 assessed semantic fluency, since only 70 assessed both phonemic and semantic fluency, when conducting inferential statistics to compare phonemic and semantic fluency, only these 70 studies were permitted to contribute to the analyses. This effectively 'controls' for the effect of illness severity. Since the same participants were compared upon each measure, paired t-tests were employed for all statistical comparisons. Mean effects were also calculated for each of the non-fluency variables identified (VIQ, TMT A, BNT and ILM) and compared with the corresponding effects for phonemic and semantic fluency. Again, to ensure that dementia severity was controlled for, only studies that assessed both the fluency and non-fluency variable of interest were included in each comparison.

To interpret how important a particular effect was in practical terms, Cohen's (1977) guidelines were adopted. These suggest that a correlation of 0.1 should be regarded as representing a small effect, 0.3 as medium, and 0.5 as large. In addition, squares of the effect size multiplied by 100 were also presented as these latter quantities represent the percentage of the variance accounted for (PVAF) on a measure of interest by group membership (i.e. the presence of DAT versus being a member of the healthy adult population). For inferential statistics comparisons were made using the PVAF.

3. Results

3.1. Participant characteristics

153 research articles published between 1983 and 2002 contributed to the present study (a complete list of the papers included in this meta-analysis, can be found by following the link at http://www.abdn.ac.uk/~psy299/dept/). In total data from 8356 DAT patients and 7634 controls were included in these research articles. Patients and controls were closely matched for age (M = 71.60, S.D. = 4.77 versus M = 69.66, S.D. = 7.26, respectively) and gender (43.46% versus 42.95% male, respectively), but patients were less educated than controls (M = 12.17, S.D. = 2.49 versus M = 13.11, S.D. = 2.50, respectively). On the MMSE, the mean for patients was substantially lower than for controls (M = 20.05, S.D. = 3.56 versus M = 28.73, S.D. = 0.59).

3.2. Effect sizes for patients with DAT relative to healthy elderly

Stem and leaf displays were constructed to provide a visual illustration of the distribution of effects for phonemic and semantic fluency (Table 1). A positive sign indicates that patients performed worse than controls, a negative sign the reverse. For both types of fluency, there is a broad range of scores. However, one outlying value is apparent in each table; −0.24 for phonemic fluency, and −0.11 for semantic fluency; both were omitted from the following analyses.

Table 2 presents estimates of the mean effects, their variability, and practical importance for phonemic and semantic fluency. To estimate the degree of heterogeneity of the effects contributing to each mean, the homogeneity statistic Q and the random effects variance (σ_θ^2) were also estimated; for interpretative purposes, larger estimates of this latter statistic are indicative of increased heterogeneity. The S.D. of random effects, and the 95% confidence intervals (CI) within which random effects can be expected to fall are also presented in Table 2.

Both mean effects were significantly different from zero ($P < 0.001$). However, for semantic fluency the mean effect ($r = 0.73$) was very large and substantially larger than for phonemic fluency ($r = 0.55$). Since the number of DAT patients contributing to each of these mean effects is also very large (N = 6105 and 4914, respectively), these can be regarded as providing very reliable estimates of the corresponding parameters of interest. The PVAF by phonemic and semantic fluency is illustrated in Fig. 1 alongside the corresponding values for patients with frontal versus temporal focal cortical lesions from Henry and Crawford's (in press) meta-analysis. For the focal lesion analyses, K ranged from 5 to 36, and N from 60 to 495. It can be seen that patients with DAT are more impaired on both phonemic and semantic fluency than either focal lesion group. However,

Table 1
Stem and leaf display of effect sizes in DAT

Phonemic fluency		Semantic fluency	
Stem	Leaf	Stem	Leaf
−0.2	4	−0.2	
−0.1		−0.1	1
−0.0		−0.0	
0.0		0.0	
0.1	246	0.1	
0.2	04455889	0.2	458
0.3	0013344566777789	0.3	2
0.4	0112334455667777788999999	0.4	238
0.5	00111112222333444555556666677788889999	0.5	233467799
0.6	001222222344444445556668999999	0.6	0122233333344455678888889999
0.7	00033344455566789	0.7	00011111111222233334444455566666677777788899999
0.8	013789	0.8	000000122233344555556677777888
0.9		0.9	023339

Table 2
Verbal fluency performance for patients with DAT relative to healthy controls

	M	*K*	Patient, *N*	S.E.	95% CI of mean		*Z**	PVAF	*Q**	σ_θ^2	S.D.	95% CI of mean effects	
					Lower	Upper						Lower	Upper
Phonemic fluency	0.55	143	4914	0.016	0.52	0.58	34.4	30.6	1023.5	0.028	0.169	0.22	0.88
Semantic fluency	0.73	126	6105	0.016	0.70	0.76	44.9	53.0	4208.0	0.029	0.171	0.39	0.99

Note—*M*: the mean effect size, *K*: the number of studies, *N*: the number of participants, S.E.: the standard error, CI: the confidence intervals, PVAF: the proportion of variance accounted for, *Q*: the homogeneity statistic, σ_θ^2: the random effects variance component, S.D.: the standard deviation, *Z*: a test of the null hypothesis that the mean effect size is zero; if *Z* exceeds 1.96, the mean effect differs significantly from zero at the 0.05 level.

* All values of *Q* and *Z* are significant (i.e. $P < 0.001$).

patients with DAT and focal temporal lobe lesions are both substantially, and comparably (i.e. the lines for these groups are parallel), more impaired on semantic than on phonemic fluency.

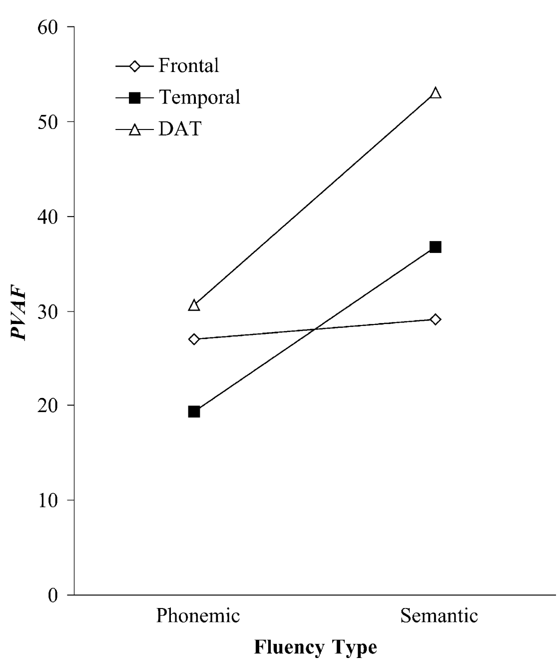

Fig. 1. PVAF in performance upon phonemic vs. semantic fluency by the presence of DAT, focal frontal, or focal temporal cortical lesions (focal lesion data taken from; Henry & Crawford, in press).

3.3. Phonemic and semantic fluency deficits relative to other cognitive deficits

As expected, the magnitude of the deficits for phonemic and semantic fluency in terms of the PVAF were both substantially and significantly related to dementia severity as measured by the MMSE (phonemic; $r = -0.55$, $K = 90$, $P < 0.001$, semantic; $r = -0.45$, $K = 78$, $P < 0.001$). Thus, as noted previously, to control for dementia severity, when applying inferential statistics only studies that included both measures of interest were permitted to contribute to each respective statistical comparison.

Table 3 presents estimates of the mean effects, their variability, and practical importance for phonemic fluency, semantic fluency, VIQ, TMT A, BNT and ILM; these mean effects are those employed for inferential statistics, and were thus calculated using only those studies that included the particular non-fluency measures of interest in addition to phonemic or semantic fluency. Thus, it can be seen in Table 3 that for each non-fluency measure, for instance ILM, two mean effects have been calculated; one for studies that also assess phonemic fluency ($r = 0.75$; $K = 57$), and one for studies that also assess semantic fluency ($r = 0.79$, $K = 27$). Each fluency mean effect was also re-calculated for these comparisons. It should be noted that for all mean effects, *Q* was significant (all $P < 0.05$).

The results of inferential statistics are presented in Table 4. For each comparison the correlation between the two measures of interest in terms of the PVAF by group

Table 3
Performance on phonemic fluency, semantic fluency, VIQ, TMT A BNT and ILM for DAT patients versus healthy controls

	M	*K*	*N*[a]	S.E.	95% CI of mean		Z^*	PVAF	Q^*	σ_θ^2	S.D.	95% CI of mean effects		
					Lower	Upper						Lower	Upper	
Studies that include phonemic fluency														
Semantic fluency	0.73	70	2674	0.017	0.69	0.76	42.7	52.7	630.6	0.016	0.128	0.47	0.98	0.57[b] (*K* = 70)
VIQ	0.63	20	481	0.033	0.56	0.69	19.2	39.3	67.4	0.014	0.117	0.40	0.86	0.63[b] (*K* = 20)
TMT A	0.49	39	1374	0.036	0.42	0.56	13.8	24.4	280.4	0.040	0.200	0.10	0.89	0.53[b] (*K* = 39)
BNT	0.60	81	2782	0.017	0.56	0.63	34.9	35.8	394.0	0.016	0.128	0.35	0.85	0.55[b] (*K* = 81)
ILM	0.75	57	2171	0.015	0.72	0.78	51.1	56.2	374.1	0.009	0.096	0.56	0.94	0.52[b] (*K* = 57)
Studies that include semantic fluency														
Phonemic fluency	0.57	70	2674	0.024	0.52	0.62	23.8	32.6	600.3	0.033	0.180	0.22	0.92	0.73[c] (*K* = 70)
VIQ	0.62	11	228	0.045	0.53	0.71	13.7	38.9	24.1	0.012	0.109	0.41	0.84	0.76[c] (*K* = 11)
TMT A	0.52	15	761	0.045	0.43	0.61	11.6	27.1	83.3	0.022	0.150	0.23	0.81	0.69[c] (*K* = 15)
BNT	0.61	63	4008	0.015	0.58	0.64	40.2	36.7	227.1	0.008	0.091	0.43	0.78	0.73[c] (*K* = 63)
ILM	0.79	27	1150	0.014	0.76	0.82	55.4	62.1	69.9	0.003	0.054	0.68	0.89	0.71[c] (*K* = 27)

Note: For conducting inferential statistics, the mean effects for phonemic and semantic fluency were recalculated for each comparison of interest. For example, 11 studies included both semantic fluency and VIQ. In addition to calculating the mean effect for VIQ from these 11 studies ($r = 0.62$), the mean effect for semantic fluency was also recalculated based only on these 11 studies (i.e. $r = 0.76$). Thus, in each comparison exactly the same participants have been tested upon each of the measures of interest, effectively 'controlling' for any substantive differences between studies, such as level of dementia severity.

[a] *N* refers to the patient group.

[b] Phonemic fluency (*M*).

[c] Semantic fluency (*M*).

* All values of *Q* and *Z* are significant (i.e. $P < 0.05$).

membership, as well as the mean difference between the two measures in this respect (ΔPVAF) is also presented. Semantic fluency was significantly more impaired than phonemic fluency ($r = 0.73$ versus 0.57, respectively, $P < 0.001$), with the presence of DAT accounting for 20.04% more variance upon this measure. Whilst the deficit for phonemic fluency did not differ from the deficits for VIQ or psychomotor speed ($P = 0.802$ and 0.140, respectively), relative to both of these measures semantic fluency was significantly more impaired ($P = 0.021$ and 0.001, respectively).

In terms of the PVAF by group membership, VIQ was moderately correlated with phonemic fluency but not at all correlated with semantic fluency ($r = 0.37$ and -0.02, respectively); however, the correlation between the TMT A and both measures of fluency was large ($r = 0.67$ and 0.60, respectively). Thus, whilst the magnitude of the deficit on both fluency measures is related to DAT patients' level of psychomotor speed, the level of impairment upon phonemic but not semantic fluency is also related to level of VIQ.

That the semantic fluency deficit predominantly reflects a problem with semantic memory storage and not difficulties with effortful retrieval is also suggested by the fact that the deficit for the BNT is significantly in excess of the phonemic fluency deficit ($P = 0.018$). Moreover, whilst in terms of the PVAF by group membership the BNT is strongly correlated with both phonemic and semantic fluency, the correlation is substantially larger for the latter ($r = 0.50$ and 0.68, respectively). This suggests that, whilst performance on both phonemic and semantic fluency is dependent upon the integrity of semantic memory, the relative contribution of this aspect of cognition to performance on the latter is substantially greater. However, since the semantic fluency deficit is also significantly larger than that for the BNT ($P < 0.001$), this suggests that semantic memory deficits are particularly

Table 4
Inferential statistics comparing phonemic and semantic fluency deficits with one another and with deficits on other cognitive measures

Comparison		Correlation	ΔPVAF	*t*	d.f.	*P*
Phonemic fluency	Semantic fluency	0.62	20.04	11.04	69	<0.001
Phonemic fluency	VIQ	0.37	0.03	0.25	19	0.802
Semantic fluency	VIQ	−0.02	18.96	2.73	10	0.021
Phonemic fluency	TMT A	0.67	4.08	1.51	38	0.140
Semantic fluency	TMT A	0.60	20.06	4.36	14	0.001
Phonemic fluency	BNT	0.50	5.47	2.43	80	0.018
Semantic fluency	BNT	0.68	17.18	9.68	62	<0.001
Phonemic fluency	ILM	0.55	28.77	11.82	56	<0.001
Semantic fluency	ILM	0.56	11.51	3.16	26	0.004

pronounced when there are additional demands upon effortful retrieval. Both phonemic and semantic fluency were significantly less impaired than ILM ($P < 0.001$ and $P = 0.004$, respectively).

3.4. Phonemic and semantic fluency deficits as a function of dementia progression

The ΔPVAF by group membership upon semantic versus phonemic fluency was not significantly related to mean scores on the MMSE ($r = 0.16$, $K = 45$, $P = 0.285$), suggesting that there is a quantitative but not a qualitative difference with regard to the relative prominence of retrieval slowing and storage degradation as the disorder progresses.

3.5. Assessing the possibility of publication bias

A number of validity threats have been identified that may lead to imprecise conclusions in both non-quantitative and meta-analytic reviews. Particularly problematic is 'the file drawer problem' which refers to the fact that significant results are more likely to be published than non-significant results. To assess whether this bias posed a threat to the results of the present study, funnel plot diagrams were constructed; for none of the variables was there evidence of this bias operating.

4. Discussion

4.1. Storage degradation versus retrieval slowing

For both phonemic and semantic fluency the presence of DAT was associated with mean deficits that were significantly different from zero ($P < 0.001$), and large in practical importance according to Cohen's (1977) criteria. However, for studies that assessed both measures, the semantic fluency deficit was significantly larger. As discussed earlier, it is thought that, relative to phonemic fluency, semantic fluency imposes greater demands upon the integrity of the semantic network, but that the two measures are equivalent in sensitivity to executive control processes such as effortful retrieval (Henry & Crawford, in press). The significantly larger deficit for semantic relative to phonemic fluency found in the present study therefore suggests that the semantic memory deficit associated with DAT reflects a degradation or disorganisation of the semantic store, and not retrieval slowing (see also; Hodges et al., 1992).

It is however, important to recognise that in interpreting data of this sort there is the possibility of identity fallacies. Thus, whilst it is suggested that a similar pattern in phonemic and semantic fluency effect sizes for patients with DAT and patients with focal temporal lobe damage provides evidence that DAT patients are like temporal patients and therefore have greater semantic than executive dysfunction, other interpretations of these data are possible. It may, for instance, be that the retrieval of semantic items in semantic fluency tasks instead depends on retrieval mechanisms that are different to those required in phonemic retrieval (e.g. if depth of encoding differs, retrieval mechanisms may differ as well), and thus that patients with DAT are more impaired on semantic than phonemic fluency for reasons that are unrelated to a degradation of the semantic store. However, the present authors favour this particular interpretation because there is a great deal of other evidence that is consistent with it.

Nebes (1989), for instance, points out that semantic tasks that impose similar demands upon the retrieval of semantic information are not always comparably impaired. Moreover, multidimensional scaling has revealed DAT to be associated with a distortion in semantic space (Chan et al., 1993), indicating that there is a fundamental difference in the organisation of DAT patients' semantic representations relative to healthy elderly. Inter-item consistency across different modes and methods of access has also been reported with respect to the errors made on different tests of semantic knowledge, suggesting that DAT is associated with a loss of knowledge for specific semantic constructs (Chertkow & Bub, 1990; Hodges et al., 1992). A retrieval deficit is unlikely to account for this pattern of performance, as different measures are presumed to invoke different methods of retrieval.

It is of note that in the literature relating to the nature of the semantic memory deficit in DAT, particular importance has been assigned to studies involving semantic priming. This is the only paradigm used with DAT patients that allows the assessment of semantic memory implicitly, minimising the influence of non-semantic cognitive processes, such as attention (Giffard, Desgranges, Kerrouche, Piolino, & Eustache, 2003). Equivalent priming for patients with DAT and healthy controls has been reported (Nebes et al., 1984), and it has been suggested that this provides evidence that the semantic network in DAT is intact. However, not all studies have found evidence of equivalent priming; others have instead reported hypopriming (i.e. reduced priming; Ober & Shenaut, 1988), or, counter-intuitively, hyperpriming (i.e. increased priming; Nebes, Brady, & Huff, 1989). Conflicting explanations have been proposed to explain this hyperpriming effect. Chertkow et al., (1989), for instance, found that hyperpriming was particularly noticeable for those items that explicit memory tasks had revealed to be degraded, but equivalent priming effects for patients and controls were found for items that were not degraded. It was suggested that hyperpriming therefore reflects a breakdown in the integrity of semantic memory. However, Nebes et al. (1989) argue that hyperpriming in DAT is consistent with the semantic store being intact, and suggested that it reflects an artefact of general cognitive slowing.

In an attempt to better understand the hyperpriming phenomenon, Giffard et al. (2002) conducted a longitudinal study of patients with DAT in which, in addition to explicit tests of semantic memory, a lexical decision task was administered to assess implicit semantic memory. In the lexical decision task related pairs of words had either coordinate (e.g. cat-dog) or attribute (cat-paw) relationships. It was found that only for the coordinate condition was hyperpriming observed, before reducing. For the attribute condition, equivalent priming was at first observed, gradually decreasing as time went on (hypopriming). When compared with the results on the explicit tests of semantic memory, it was found that the degree of semantic priming was dependent upon the level of semantic memory degradation, with hyperpriming reflecting the start of this process at the attribute level of semantic knowledge. Giffard et al. (2002) suggest that when the attributes of concepts start being lost, the ability to distinguish between similar concepts such as cat and dog is impaired, and patients with DAT therefore treat the semantic priming as repetition priming. Thus, intact semantic priming effects do not provide evidence of a preserved semantic store; instead, the profile of semantic priming effects is not static, but evolves across the course of the disease, and reflects the level of semantic memory deterioration. Thus, in addition to the results of the present meta-analysis, a great deal of other evidence has accumulated that suggests that DAT is associated with a disorganisation or degradation of semantic representational knowledge.

4.2. Differential verbal fluency deficits in DAT

Of the two types of fluency, only semantic fluency was associated with a deficit significantly in excess of the deficits for VIQ and the TMT A, indicating that DAT is not characterised by a differential executive deficit, at least as indexed by phonemic fluency, but is characterised by a differential deficit in semantic memory storage. Indeed, a non-speeded, relatively non-effortful measure of semantic memory (the BNT), was associated with a significantly larger deficit than phonemic fluency. Failure to find a differential deficit on phonemic fluency for patients with DAT contrasts with the results for patients with focal cortical lesions, for whom deficits on phonemic fluency have been found to substantially exceed deficits on VIQ and the TMT A (Henry & Crawford, in press). It is of interest that the results across the present and Henry and Crawford's (in press) focal lesion meta-analytic study are entirely consistent with Miller's (1984) study, in which it was found that whilst both focal frontals and patients with DAT are significantly impaired on phonemic fluency, when predicted fluency scores are derived from measures of verbal intelligence there is a substantial discrepancy between predicted and actual scores for the frontal groups, but not for those with DAT. Thus, frontal patients exhibit a differential deficit that cannot be attributed to decrements in verbal intelligence, but rather, reflects a specific problem with executive processing. In Miller's (1984) study, as in the present meta-analysis, there is no evidence of this type of selective deficit in DAT; instead, performance levels on tests of phonemic fluency are entirely consistent with patients' overall level of cognitive functioning.

4.3. Additive deficits in semantic memory and retrieval slowing

Relative to semantic fluency, performance on the BNT was significantly less impaired. There is a great deal of evidence that confrontation naming is very sensitive to the integrity of the semantic system, but the present results indicate that it is far less sensitive to the presence of DAT than semantic fluency. The literature relating to semantic memory performance in DAT has indicated that there is great variability in the prominence of deficits on different measures of this construct, and it has been suggested that a distinction between implicit and explicit tasks, or the level of automaticity versus controlled strategic processing may be critical factors in explaining this variability (for a review see; Nebes, 1989). Since the BNT and semantic memory are both explicit tasks of memory, the former distinction does not apply, and we would suggest that the critical factor distinguishing these two tasks is the degree to which they impose demands upon effortful retrieval.

A prominent view in the cognitive ageing literature is that the magnitude of deficits on tests of memory may be at least partially determined by the degree to which a measure invokes self-initiated processing, with smaller deficits to be expected on measures in which the task and stimulus materials themselves initiate and guide appropriate processing (Craik, 1986). This theoretical model has been successfully applied to resolve discrepancies in the prospective memory literature (Henry, Macleod, Phillips, & Crawford, in press) and it has been suggested that it may also explain discrepancies in the semantic memory literature relating to DAT (Nebes, 1989). Huff et al. (1986) for instance, argues that object-naming tests such as the BNT are less impaired than measures of semantic fluency in DAT, because the former provides more support to the lexical search process through providing information from the visual stimulus to guide the search. Thus, since semantic fluency but not confrontation naming imposes substantial demands upon effortful self-initiated retrieval in addition to semantic memory, it may be that deficits in semantic memory and effortful retrieval are additive. It should be noted that this present rationale does not require that the BNT imposes no demands upon effortful self-initiated retrieval processes (especially since performing this task may certainly *seem* effortful to patients with DAT), simply that, relative to tests of semantic fluency, these particular demands are reduced.

It is of interest that it has been found that the level of impairment associated with DAT differs for tasks that require the same semantic information, but vary in terms of their relative dependence upon effortful retrieval processes. Nebes and Brady (1988) for instance, found that although patients with DAT were capable of accurately deciding whether a particular attribute was related to a particular concept, they were significantly impaired when required to generate attributes of a concept. Such results have traditionally been interpreted as evidence that the semantic memory deficit in

DAT reflects retrieval slowing, and not a degradation of the semantic store. We would suggest that the present results are consistent with the possibility that whilst degradation of the semantic store underlies the semantic memory deficit in DAT (as indicated by significantly larger deficits upon tests of semantic fluency and the BNT relative to tests of phonemic fluency), if a task additionally imposes substantial demands upon effortful retrieval these deficits may be additive.

4.4. Sources of heterogeneity in DAT

The difference in the PVAF by semantic and phonemic fluency was not related to mean scores on the MMSE ($r = 0.16$, $P = 0.29$). This suggests that deficits in semantic memory storage exceed executive dysfunction to a comparable degree at every stage of the disorder, and that the relative contributions of executive versus semantic memory storage impairment can be regarded as broadly equivalent at each stage of the disease. Thus, although it has been suggested that executive deficits are increasingly prominent later in the course of the illness (Duke & Kaszniak, 2000) and that differences in dementia severity may account for discrepancies between studies with regard to the relative prominence of deficits upon tests of phonemic and semantic fluency (Butters et al., 1987; Margolin et al., 1990), the present results suggest that this is unlikely to be the case.

However, the mean effects for all of the cognitive measures assessed in the present study were associated with significant heterogeneity. Since sampling error, the most serious source of artefactual variance, had been removed, substantive differences between studies remain. Some of this variance will almost certainly be attributable to differences in dementia severity between patients. Although not significantly related to the *relative* magnitude of deficits on the two measures, increased dementia severity was significantly related to effect size magnitude for each of the two measures individually ($r = -0.55$ and -0.45 for phonemic and semantic fluency, respectively). However, it is important to stress that DAT is an extremely heterogeneous disorder, and it is therefore likely that other factors in addition to dementia severity will also moderate effect size magnitude upon tests of fluency as well as other cognitive measures across individual studies.

Thus, whilst patients with DAT are generally more impaired on semantic relative to phonemic fluency, for certain sub-groups, the reverse, or a pattern of comparable impairment may emerge, and as noted previously, this has been reported in some studies (Nebes et al., 1984; Ober et al., 1986; Suhr & Jones, 1998). Indeed, whilst temporo-parietal abnormalities are typically the earliest, most prominent pathological features of DAT (Braak & Braak, 1996; Gomez-Isla et al., 1996), an executive variant has been identified (see e.g.; Back-Madruga et al., 2002) that is characterised by concomitant frontal disturbances. Moreover, it has been suggested that DAT patients who present with a rapid rate of progression may be particularly impaired on tests designed to capture executive dysfunction (Coen et al., 1996; Mann, Mohr, Gearing, & Chase, 1992). Thus, the present results do not rule out the possibility that distinct sub-types exist that differ with respect to their level of frontal and executive impairment. However, much of this variance will be bundled up within, rather than between studies, and thus the specific influence of each cannot be explored in the present study (for example, there were insufficient studies that were restricted only to a rapid versus a slow rate of progression, etc.). The heterogeneity statistic Q quantifies the degree of heterogeneity between studies but cannot address the degree of heterogeneity within each of the studies contributing to a mean. However, it is recommended that if future primary research breaks down their samples more fully, meta-analysis should be conducted to address which variables moderate performance on tests of verbal fluency.

Finally, it could be that floor effects in the more severely demented patients have obscured differences in performance on the two tasks, and thereby attenuated the discrepancy between phonemic and semantic fluency. Without the individual data for each case in each study, it is impossible to rule out definitively whether or not patients scoring zero have contributed to each mean, although it is to be hoped that the authors of the primary studies would have excluded such cases. However, evidence that floor effects will not have markedly influenced the present results comes from examining scores on the most commonly employed phonemic and semantic fluency tasks. For the former (the FAS) the grand raw score mean for the DAT samples was 14.07 (S.D. = 7.63) and in only four studies did the mean fall below 10.00 (the lowest mean was 7.50). For semantic fluency, a single trial of animal fluency was the most commonly administered variant; the grand raw score mean for the DAT samples was 12.56 (S.D. = 5.91) and in only four studies did the mean fall below 5.0. Thus, there are no indications that floor effects are a serious feature of these data. Furthermore, at the level of meta-analysis, and as noted above, the *difference* in the relative magnitude of the phonemic versus semantic fluency deficit did not correlate with dementia severity.

4.5. Summary and conclusions

DAT patients were significantly more impaired on both semantic fluency and the BNT than on phonemic fluency, and for semantic but not phonemic fluency the deficit qualified as a differential deficit relative to verbal intelligence and psychomotor speed. This suggests that the deficit in semantic memory reflects a degradation in the integrity of the semantic store, and not retrieval slowing. However, since the deficit for the BNT was significantly smaller than the deficit for semantic fluency, semantic memory impairment appears to be especially pronounced when there are additionally substantial demands on effortful retrieval. Moreover, episodic memory appears to be most disrupted by the disorder, as the deficit for a measure of this construct was significantly in excess of the deficits for both phonemic and semantic fluency. Finally, although the effect sizes for both phonemic and semantic fluency were significantly positively related to dementia severity, the difference in terms of the PVAF by these measures was not, suggesting that the relative prominence of retrieval slowing and semantic store degradation is comparable at each stage of DAT.

References

American Psychological Association (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC.

Amieva, H., Phillips, L. H., Della Sala, S., & Henry, J. D. (in press). Inhibitory functioning in Alzheimer's Disease. *Brain*.

Back-Madruga, C., Boone, K. B., Briere, J., Cummings, J., McPherson, S., & Fairbanks, L. et al., (2002). Functional ability in executive variant Alzheimer's disease and typical Alzheimer's disease. *Clinical Neuropsychologist*, *16*, 331–340.

Balota, D. A., & Ferraro, F. R. (1996). Lexical, sublexical, and implicit memory processes in healthy young and healthy older adults and in individuals with dementia of the Alzheimer type. *Neuropsychology*, *10*, 82–95.

Bayles, K. A., Tomoeda, C. K., Kaszniak, A. W., & Trosset, M. W. (1991). Alzheimer's disease's effects on semantic memory: loss of structure or impaired processing? *Journal of Cognitive Neuroscience*, *3*, 166–182.

Braak, H., & Braak, E. (1996). Evolution of neuropathology of Alzheimer's disease. *Acta Neurologica Scandinavica*, *165*(Supplement), 3–12.

Butters, N., Granholm, E., Salmon, D. P., Grant, I., & Wolfe, J. (1987). Episodic and semantic memory: a comparison of amnesic and demented patients. *Journal of Clinical and Experimental Neuropsychology*, *9*, 479–497.

Chan, A. S., Butters, N., Paulsen, J. S., Salmon, D. P., Swenson, M. R., & Maloney, L. T. (1993). An assessment of the semantic network in patients with Alzheimer's disease. *Journal of Cognitive Neuroscience*, *5*, 254–261.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of the Alzheimer's type: what do various measures measure? *Brain*, *113*, 397–417.

Coen, R. F., Maguire, C., Swanwick, G. R., Kirby, M., Burke, T., & Lawlor, B. A. et al., (1996). Letter and category fluency in Alzheimer's disease: A prognostic indicator of progression? *Dementia*, *7*, 246–250.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Revised ed.). New York: Academic Press.

Craik, F. I. M. (1986). A functional account of age differences in memory. In F. Klix & H. Hagendorf (Eds.), *Human memory and cognitive capabilities*: *Mechanisms and performances* (pp. 409–422). Amsterdam, The Netherlands: Elsevier-North-Holland.

Crawford, J. R., & Henry, J. D. (in press). Assessment of executive deficits. In P.W. Halligan & N. Wade (Eds.), *The effectiveness of rehabilitation for cognitive deficits*. London: Oxford University Press.

Duke, L. M., & Kaszniak, A. W. (2000). Executive control functions in degenerative dementias: A comparative review. *Neuropsychology Review*, *10*, 75–99.

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

Giffard, B., Desgranges, B., Kerrouche, N., Piolino, P., & Eustache, F. (2003). The hyperpriming phenomenon in normal aging: A consequence of cognitive slowing? *Neuropsychology*, *17*, 594–601.

Giffard, B., Desgranges, B., Nore-Mary, F., Lalevee, C., Beaunieux, H., & de la Sayette, V. et al., (2002). The dynamic time course of semantic memory impairment in Alzheimer's disease: Clues from hyperpriming and hypopriming effects. *Brain*, *125*, 2044–2057.

Gomez-Isla, T., Price, J. L., McKeel, D. W., Morris, J. C., Growdon, J. H., & Hyman, B. T. (1996). Profound loss of layer II entorhinal cortex neurons in very early Alzheimer's disease. *Journal of Neuroscience*, *16*, 4491–4500.

Grande, L., McGlinchey-Berroth, R., Milberg, W. P., & D'Esposito, M. (1996). Facilitation of unattended semantic information in Alzheimer's disease: Evidence from a selective attention task. *Neuropsychology*, *10*, 475–484.

Grober, E., Buschke, H., Kawas, C., & Fuld, P. (1985). Impaired ranking of semantic attributes in dementia. *Brain and Language*, *26*, 276–286.

Henry, J. D., & Crawford, J. R. (in press). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*.

Henry, J. D., Macleod, M. S., Phillips, L.H., & Crawford, J. R. (in press). A meta-analytic review of prospective memory and aging. *Psychology and Aging*.

Hodges, J. R., Salmon, D. P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology Neurosurgery and Psychiatry*, *53*, 1089–1095.

Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease—Failure of access or degraded knowledge. *Neuropsychologia*, *30*, 301–314.

Huff, F. J., Corkin, S., & Growdon, J. H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and Language*, *28*, 235–249.

Kaplan, E. F., Goodglass, H., & Weintraub, S., 1983. *The Boston naming test*. Philadelphia: Lea & Febiger.

Mann, U. M., Mohr, E., Gearing, M., & Chase, T. N. (1992). Heterogeneity in Alzheimer's disease: Progression rate segregated by distinct neuropsychological and cerebral metabolic profiles. *Journal of Neurology, Neurosurgery and Psychiatry*, *55*, 956–959.

Margolin, D. I., Pate, D. S., Friedrich, F. J., & Elia, E. (1990). Dysnomia in dementia and in stroke patients: Different underlying cognitive deficits. *Journal of Clinical and Experimental Neuropsychology*, *12*(4), 597–612.

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. *Neurology*, *34*, 939–944.

Miller, E. (1984). Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *British Journal of Clinical Psychology*, *23*, 53–57.

Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49*, 1253–1258.

Monsch, A. U., Seifritz, E., Taylor, K. I., Ermini-Funfschilling, D., Stahelin, H. B., & Spiegel, R. (1997). Category fluency is also predominantly affected in Swiss Alzheimer's disease patients. *Acta Neurologica Scandinavica*, *95*, 81–84.

National Research Council (1992). *Combining information statistical issues and opportunities for research*. Washington DC: National Academy Press.

Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, *106*, 377–394.

Nebes, R. D. (1994). Contextual facilitation of lexical processing in Alzheimer's disease: intralexical priming or sentence-level priming? *Journal of Clinical and Experimental Neuropsychology*, *16*, 489–497.

Nebes, R. D., & Brady, C. B. (1988). Integrity of semantic fields in Alzheimer's disease. *Cortex*, *24*, 291–299.

Nebes, R. D., Brady, C. B., & Huff, F. J. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *11*, 219–230.

Nebes, R. D., Martin, D. C., & Horn, L. C. (1984). Sparing of semantic memory in Alzheimer's disease. *Journal of Abnormal Psychology*, *93*, 321–330.

Ober, B. A., Dronkers, N. F., Koss, E., Delis, D., & Friedland, R. F. (1986). Retrieval from semantic memory in Alzheimer type dementia. *Journal of Clinical and Experimental Neuropsychology*, *1*, 75–92.

Ober, B. A., & Shenaut, G. K. (1988). Lexical decision and priming in Alzheimer's disease. *Neuropsychologia*, *26*, 273–286.

Perry, R. J., Watson, P., & Hodges, J. R. (2000). The nature and staging of attention dysfunction in early (minimal and mild) Alzheimer's disease: Relationship to episodic and semantic memory impairment. *Neuropsychologia*, *38*, 252–271.

Reitan, W., 1990. *The Halstead-Reitan Neuropsychological Test Battery*: *Theory and clinical interpretation*. Neuropsychology Press.

Rohrer, D., Salmon, D. P., Wixted, J. T., & Paulsen, J. S. (1999). The disparate effects of Alzheimer's disease and Huntington's disease on semantic memory. *Neuropsychology*, *13*, 381–388.

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.

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Russell Sage Foundation.

Stanley, T. D. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, *15*, 131–150.

Stuss, D. T., & Benson, D. F. (1986). *The frontal lobes*. New York: Raven Press.

Suhr, J. A., & Jones, R. D. (1998). Letter and semantic fluency in Alzheimer's, Huntington's, and Parkinson's dementias. *Archives of Clinical Neuropsychology*, *13*, 447–454.

Wechsler, D. (1981). *WAIS-R manual*. New York: The Psychological Corporation.

Wechsler, D., & Stone, C. P. (1974). *Wechsler memory scale II manual*. New York: The Psychological Corporation.