Journal of the International Neuropsychological Society (1999), **5**, 676–684.

Well-organized conceptual domains in Alzheimer's disease

BETH A. OBER AND GREGORY K. SHENAUT
Human Development, University of California, Davis and Department of Veterans Affairs,
Northern California Health Care System, Martinez

(Received December 14, 1998; Revised April 2, 1999; Accepted June 1, 1999)

Abstract

We used a novel apparatus called the *flags board* to elicit similarity judgments from 32 Alzheimer's disease (AD) patients and 32 elderly normal (EN) controls for two 12-member conceptual domains, ANIMALS and (musical) INSTRUMENTS. Based on Pathfinder and multidimensional scaling (MDS) analyses, performance by AD patients was nearly identical to that of EN controls for ANIMALS. Performance differed for INSTRUMENTS, but the AD group's Pathfinder network was found to agree with the intuitions of a panel of 18 raters as well as the EN group's. MDS analysis showed no deficit on abstract dimensions for the AD group, for either domain. The results are discussed in the context of degradation *versus* preservation of semantic memory in AD. (*JINS*, 1999, *5*, 676–684.)

Keywords: Alzheimer's disease, Conceptual domains, Semantic memory, Semantic priming

INTRODUCTION

Abnormal performance by Alzheimer's disease (AD) patients on certain tasks thought to measure intactness of semantic memory arises only when cognitive resources impaired by the disease are critically involved. When different, less cognitively demanding, procedures are used, performance deficits may be reduced or disappear.

For example, for a period of time there was a perplexing lack of consistency in the AD literature regarding semantic priming. Some researchers reported normal priming (e.g., Nebes et al., 1984; Ober et al., 1991), while others found an abnormally increased priming effect (hyperpriming; e.g., Chertkow et al., 1989; Martin, 1992). Various explanations were put forward to account for each of these findings, but none of them could explain both findings. In a meta-analysis of 21 semantic priming experiments, we (Ober & Shenaut, 1995) discovered that when experimental conditions such as pairwise priming, long stimulus-onset asynchrony (SOA), and high relatedness proportion encouraged the use of strategies and controlled processing, AD participants consistently showed hyperpriming. However, when experimental conditions such as continuous priming, short SOA, and low relatedness proportion discouraged such controlled processing, AD participants' performance did not differ from that of elderly normal (EN) controls. This cognitive-resource explanation was bolstered by a later finding, in a dual-SOA pairwise lexical decision priming procedure, that a single group of AD patients could be made to exhibit both normal priming under relatively automatic conditions and hyperpriming under relatively controlled conditions (Shenaut & Ober, 1996).

In another example, poor performance on picture naming tasks is often taken to support a fundamental deficit in semantic memory in AD, rather than as a sign of word-retrieval difficulties. However, we recently have completed a study involving a multiple-choice version of the Boston Naming Test with 45 AD participants in which we found almost perfect accuracy (Ober & Shenaut, 1998). Other researchers have shown greatly improved performance with name recognition *versus* name recall versions of confrontation naming tests (e.g., LaBarge et al., 1992). Again, this suggests that AD patients may retain basic knowledge, but are prevented from accessing it under certain test conditions.

Chan and her coworkers (Chan et al., 1993b, 1995; the same AD data is included in these two papers) have presented findings that purport to demonstrate structural abnormalities in the semantic networks of AD patients. They used a triadic comparison task to elicit similarity judgments concerning 12 four-footed animals from AD participants, EN controls, and several other patient groups. Based primarily upon differences in multidimensional scaling (MDS)

Reprint requests to: Beth A. Ober, Department of Human and Community Development, University of California, One Shields Avenue, Davis, CA 95616. E-mail: baober@ucdavis.edu

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

dimensionality, they concluded that AD participants' semantic networks differ in systematic ways from the others: In particular, that they were less abstract and more concrete (i.e., based relatively more on perceptual attributes). The data is not unambiguous concerning AD patients' semantic networks; Bonilla and Johnson (1995) found, in a card-arrangement task, that the spatial network of a group of mild AD patients did not differ from that of a group of EN controls.

Barsalou (1993) has proposed that concept formation is a dynamic, flexible process that uses selective attention to produce context-specific, somewhat idiosyncratic categories and semantic networks in working memory. Unlike theories proposing that concepts and semantic networks are constants of long-term memory, under Barsalou's formulation, *ad-hoc* semantic networks are generated using perception-based knowledge but also depending on the details of the current context, the participants' goals, and the participants' prior experiences. Barsalou has found support for this, for example, in that participants' descriptions and definitions of concepts sometimes differ between participants to a surprising degree, and that even the same participants' descriptions or definitions can vary from session to session.

That selective attention, goals, and other cognitive processes are presumably involved in the generation of these temporary semantic networks suggests that groups of individuals such as AD patients, who have deficits in attentional processes or working memory, may have some difficulty in creating such conceptual structures in spite of a relatively preserved underlying knowledge base. This would present a picture of semantic–conceptual normality in tasks that do not require generation of these dynamic conceptual structures (relatively automatic and implicit tasks) but semantic–conceptual abnormality when higher-level tasks, which require these dynamic structures, are used.

While triadic comparison, used by Chan et al. (1993a), is a well-known task for the elicitation of similarity judgments, the fact that card sorting, used by Bonilla and Johnson (1995), produced a more normal result for AD participants suggests that task characteristics may have an effect on the characteristics of the semantic network created in order to perform the task. Gammack (1990) found such a result with five different tasks used to elicit semantic networks from an expert participant.

In what follows, we will explore the issue of task effects on empirically derived semantic networks in AD a little further. We will describe a new task and the results of using it to elicit similarity information from AD and EN participants for stimuli from two different semantic domains.

METHODS

Research Participants

The 32 AD participants were referred from the U.C. Davis Alzheimer's Disease Clinical Center (Sacramento and Berkeley sites). All of these individuals had undergone thorough evaluations by a neurologist, neuropsychologist, and nurse practitioner. All met the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria for probable AD (McKhann et al., 1984). This AD sample was by and large a mildly-to-moderately impaired sample, with 9 out of the 32 AD participants scoring less than 20 on the Mini-Mental State Exam (MMSE; Folstein et al., 1975). The 32 EN participants were recruited from the community (Sacramento, Davis, and Bay Area) and from among spouses of AD participants. The EN participants met all of the same exclusionary criteria as the AD participants (no history of heavy alcohol consumption, vascular disease, diabetes, etc.). The demographic information is presented in Table 1. Participant groups did not differ significantly in years of age, education, or in the number of errors on the American version of the Nelson Adult Reading Test (AMNART; Grober & Sliwinski, 1991), but did differ significantly on the MMSE [$F(1,62) = 81.124$, $p < .001$]. (It should be noted that a number of studies have reported normal or close-to-normal (AM)NART performance in AD, especially in the early stages of the disease; for a brief review see O'Carrol, 1992.)

Apparatus and Procedure

We developed a new testing apparatus for this experiment. It consisted of a 46-cm square of plastic, with a rectangular arrangement of 12 holes drilled at one end, and a square 10×10 arrangement of holes drilled in the center of the remaining area. There were coordinate labels in the margin of the 10×10 square, digits for one dimension, letters for the other. Each stimulus was typed on two adhesive labels stuck together back-to-back on one end of a small plastic dowel that fit into the holes on the board. Because these resembled small flags, we dubbed the apparatus the *flags board*.

Table 1. Summary of participant demographics

Group	Male	Female	Age (years) *M*	Age (years) (*SD*)	Education (years) *M*	Education (years) (*SD*)	MMSE *M*	MMSE (*SD*)	AMNART *M*	AMNART (*SD*)
AD	18	14	76.44	(7.17)	14.59	(3.30)	22.47	(4.27)	32.50	(7.94)
EN	9	23	73.94	(3.55)	14.56	(1.61)	29.44	(0.98)	34.75	(5.75)

Note. MMSE score is number correct out of 30; AMNART score is number of correctly pronounced words out of 45.

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

For each conceptual domain, we placed the 12 flags containing the stimuli, in alphabetical order, into the 12 "starting" holes on one end of the flags board (the end farthest away from the participant), and directed the participant to place each flag, one at a time, into the 10×10 grid. The instructions to participants were as follows:

This task involves grouping items according to how similar they are. You will be arranging these animals (musical instruments) on this board in a way that reflects their similarity to each other. That is, if you believe that items are similar, put them close together. If you believe that items are not similar, put them further apart. Please use the entire board.

After having the task described to them, participants were given a list of all of the stimuli for the given domain, typed in large lettering on a single sheet, in the same order that they were placed onto the flags board at the beginning of the session; participants were asked to read through the entire list before beginning the flag placement procedure. Participants were free to relocate flags if necessary; the goal was to have the best arrangement possible in terms of overall similarity among the 12 flags. No time limits were imposed; the approximate times to completion, for each domain, were 2.5 to 5 min for AD patients and 1.5 to 3 min for EN controls.

In the flags board procedure, after the participants were satisfied with their flag placement, the coordinates were transcribed by the experimenter, and the Euclidean distances between pairs of flags were used as raw proximity data in subsequent analyses. After each participant was finished placing the flags into the flags board, the experimenter asked, "Can you tell me how you organized the animals (or musical instruments) on the board?" The participant's response to this question was recorded by the experimenter; there was no time limit for conveyance of this strategy information. For each of the strategy protocols, we tabulated all of the organizational principles named by the participant, and whether the principles identified general strategies used to organize the whole domain ("shape," "appearance"), or specific subsets of the domain ["wild," "string(ed instrument)"].

Stimuli

We used word stimuli in each of two conceptual domains: ANIMALS (*bear*, *cat*, *cow*, *dog*, *elephant*, *giraffe*, *horse*, *lion*, *rabbit*, *sheep*, *tiger*, and *zebra*), and (musical) INSTRUMENTS (*bassoon*, *cello*, *clarinet*, *drum*, *flute*, *harp*, *oboe*, *piano*, *trombone*, *trumpet*, *tuba*, and *violin*). Our ANIMALS were identical to those used by Chan et al. (1993a) in an MDS assessment of verbal fluency data obtained from AD patients for the category "animals"; these stimuli may be symmetrically classified into subgroups using the features domestic–wild, carnivore–herbivore, and small–large. (It should be noted that in Chan et al., 1993b, 1995, *pig* replaced *sheep*, which somewhat reduces the symmetry of the attributes of the domain.) The musical INSTRUMENTS are less easily classified, although there are features such as family (string, brass, woodwind, percussion), musical range (high, low), relative physical location in the symphony orchestra seating chart, and relative familiarity (e.g., piano or trumpet *vs.* oboe or bassoon), that might be used for classification. All participants performed the task for both ANIMALS and INSTRUMENTS; the order was counterbalanced across participants.

Network Analysis Methods

Spatial models of similarity represent objects as points in multidimensional space, with distances between them representing dissimilarity. The standard technique for converting similarity judgments on a set of objects to locations in semantic space is multidimensional scaling (MDS). Spatial models cannot easily be made to represent such relations because points in multidimensional space cannot be arranged in violation of fundamental geometric axioms. These models lost favor in the late 1970's when it was pointed out (Tversky, 1977) that common similarity relations frequently violate such axioms of metricity as the triangle inequality (the triangle inequality states that each side of a triangle is shorter than the sum of the other sides).

Another way to express the triangle inequality is that a straight line is the shortest path between two points, therefore, no indirect path through a third point can be shorter. An example of a violation of the triangle inequality which was often cited during the Cold War concerned similarity judgments on the countries Cuba, Jamaica, and Russia. Cuba and Jamaica are very similar geographically, both being island nations; Cuba and Russia were very similar politically, both being Communist countries; Jamaica and Russia were very dissimilar. Based on participants' ratings of the similarity of each pair of items, the path from Russia to Jamaica *via* Cuba is shorter than the direct path, which violates the Euclidean triangle inequality—it cannot be represented in two dimensions.

In response to this criticism, several other techniques have been developed for representing similarity relationships. For example, the *Pathfinder* method (Dearholt & Schvaneveldt, 1990) is a direct response to the problem of triangle inequalities: it produces network representations (called PFNETs, for *Pathfinder networks*) of domains by deleting all direct links for which there is a shorter indirect path (i.e., all violations of the triangle inequality are eliminated). There are two parameters that control the operation of the Pathfinder algorithm. The r parameter is the Minkowski exponent, and allows the distance computation to be nonlinear; the q parameter sets the maximum number of links in indirect paths examined for violations of the triangle inequality. For psychological and sociological data, the optimal parameters are $r = \infty$ and $q = N - 1$ (Durso & Coggins, 1990). With these settings, only ordinal assumptions are made regarding distances, and triangle inequalities of any path length are found and eliminated. This produces the sparsest PFNETs, which have been found to correspond maximally to intuitions about similarity relations.

RESULTS

Pathfinder Networks

The initial analysis of the data from the flags board consisted of creating PFNETs (with $r = \infty$ and $q = N - 1$) for the average proximity data for each domain and each group. These networks can be found in Figures 1 and 2. As can be seen, the ANIMALS domain networks for the two groups are nearly identical: The only difference is that *zebra* is connected to *lion* for the EN group, and to *tiger* for the AD group. However, the two groups' networks for INSTRUMENTS differed substantially: AD participants connected *violin* rather than *cello* to *harp*, *flute* rather than *clarinet* to *oboe*, *trombone* rather than *trumpet* to *clarinet*, and *bassoon* rather than *drum* and *trombone* to *tuba*.

A common and very useful statistic available for comparing two PFNETs is the *PFC* (*Pathfinder closeness*) statistic, which is the number of connections in both PFNETs divided by the number of connections in either PFNET (*PFC* is also known as the intersection:union ratio). We determined the probability of two random PFNETs with 12 nodes (when each link is saved *vs.* discarded, by chance, from each of the two networks, with the restriction that 11 links remain in each network) having an equal or higher *PFC* value by using a Monte Carlo simulation with over 16 million pairs of PFNETs. For the ANIMALS domain, the two PFNETs are significantly more similar than chance ($PFC = 0.8333$, $p < .001$), but for the INSTRUMENTS domain, the similarity is not significantly different from chance ($PFC = 0.3750$, $p = .268$).

However, in examining the differences between the EN and AD networks for INSTRUMENTS, we were struck by the fact that the connections that differed between them did not seem to indicate abnormality in one and normality in the other. To examine this intuition, we produced five spatially different layouts of the graph in Figure 2, with no labels indicating which links were AD, EN, or both, keeping the same connection between pairs of nodes. We then asked 18 students and staff members from the authors' academic department to indicate their agreement with each of the 16 connections. Each connection in a graph was rated on a 3-point scale, where 1 = *I would definitely connect this pair*, 2 = *I might connect this pair*, and 3 = *I would never connect this pair*. The average rating was 1.58 ($SD = .38$) for links occurring in both AD and EN PFNETs, 1.86 ($SD = .49$) for AD-only links, and 1.79 ($SD = .40$) for EN-only links. There was no significant difference between the extent to which raters agreed with a link as a function of whether it was found in AD participants' or EN participants' PFNETs [$t(17) = 1.00$, $p > .5$], and raters agreed significantly more with links in both PFNETs than with links in only one [$t(17) = 3.79$, $p < .01$].

The closeness of two averaged, group-wise PFNETs doesn't take into account the range of variation within the groups. One standard way to address this issue is to use a common baseline against which to compare participants from both groups. In order to derive a baseline, we averaged the proximity values for the EN participants in each domain, and compared the EN group-wise PFNET to each individual participant's PFNET (with $r = \infty$ and $q = 5$; PFNETs produced with $q = 5$ are nearly as sparse as those with $q =$

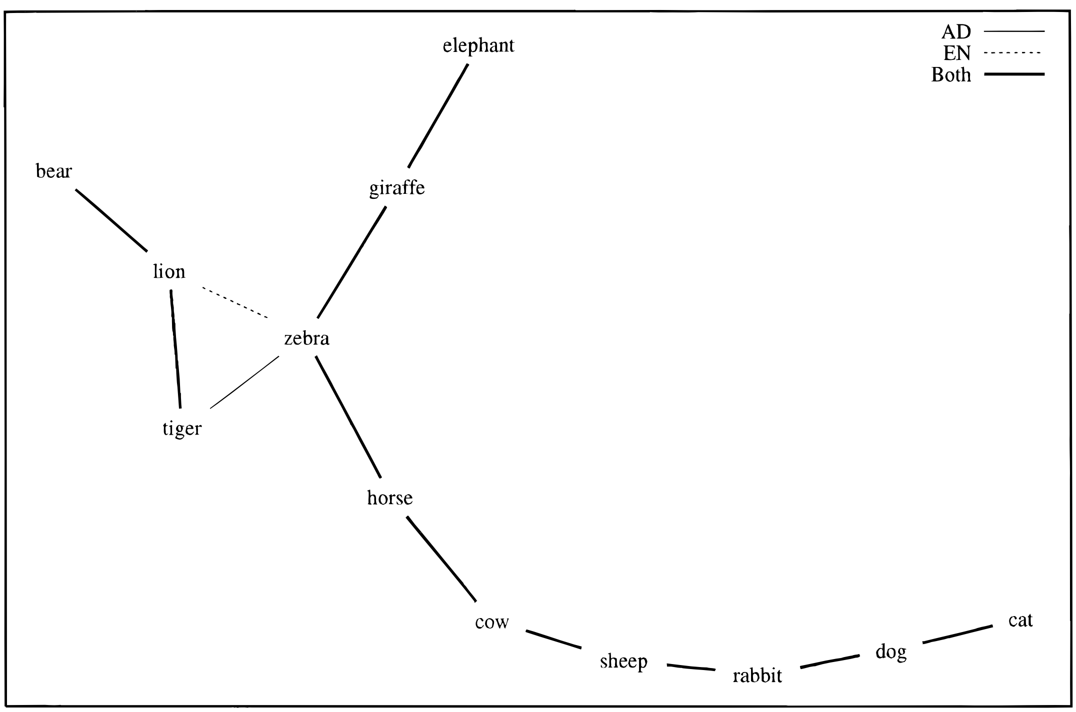

Fig. 1. Composite ANIMALS PFNET, with $r = \infty$ and $q = N - 1$.

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

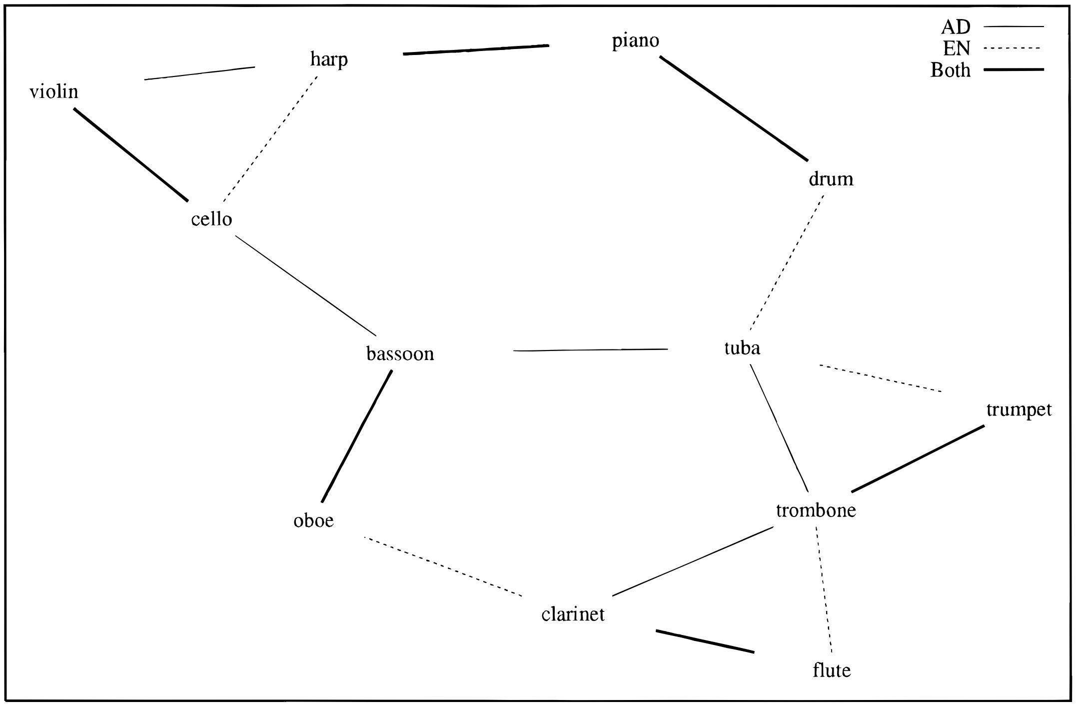

Fig. 2. Composite INSTRUMENTS PFNET, with r = ∞ and $q = N - 1$.

$N - 1$, and take seconds rather than an hour to compute). The results, which are summarized in the upper half of Table 2, were analyzed using a 2 (group) × 2 (domain) ANOVA. We found significant effects of group (less similarity to the EN baseline for the AD compared to EN participants) [$F(1,62) = 6.969$, $p < .05$], and domain [less similarity to the EN baseline for INSTRUMENTS compared to ANIMALS; $F(1,62) = 9.902$, $p < .01$], as well as a significant Domain × Group interaction [a bigger difference between participant groups in similarity to the EN baseline for INSTRUMENTS compared to ANIMALS; $F(1,62) = 8.087$, $p < .01$].

Table 2. PFC analysis of PFNETs ($r = \infty$, $q = 5$) for two domains

	PFC	*SD*	Corr. w/MMSE
Individual PFNETs compared to composite EN baseline PFNET			
AD			
ANIMALS	.30	.10	.10
INSTRUMENTS	.19**	.10	.31
EN			
ANIMALS	.30	.09	N.A.
INSTRUMENTS	.29	.14	N.A.
Individual PFNETs compared to individual EN PFNETs			
AD			
ANIMALS	.20	.04	.17
INSTRUMENTS	.16***	.04	.34*
EN			
ANIMALS	.20	.04	N.A.
INSTRUMENTS	.20	.04	N.A.

Note. PFC = Pathfinder closeness. PFNET = Pathfinder network. The asterisks in the PFC column indicate significant differences (** for $p < .01$, *** for $p < .001$) between the AD PFC and EN PFC values. The asterisk (* for $p < .05$) in the Corr. w/MMSE column indicates a significant correlation of the AD PFC with MMSE score. $N = 32$ for both AD and EN groups. N.A. = not applicable.

Since we are specifically interested in a model wherein individual participants' semantic networks may differ substantially and perhaps qualitatively from one another, it is somewhat questionable to depend on a single, averaged set of proximity values in the baseline. Therefore, we did a second ANOVA, this time based on a comparison of every participant's PFNET ($r = \infty$, $q = 5$) with every individual EN participant's PFNET. The resulting average PFCs are presented in the lower half of Table 2. We again found significant effects of group, $F(1,62) = 10.204$, $p < .01$], domain [$F(1,62) = 5.793$, $p < .05$], and Group × Domain [$F(1,62) = 8.362$, $p < .01$]. The direction of the differences was the same as in the previous ANOVA.

In order to determine whether the individual AD PFNETs were differentially similar to those of the EN participants as a function of dementia severity, we performed correlations between the *PFC* statistics and the MMSE scores of the AD participants. These data are summarized in the far-right column of Table 2. Only correlations for INSTRUMENTS were significant or approached significance; both were positive, which is consistent with the idea that less demented AD participants tend to be more similar to EN participants.

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

MDS Analyses

In spite of the controversial nature of the implied spatial model of semantic memory, MDS methods have recently been applied to similarity data obtained from AD patients by Chan and colleagues (e.g., Chan et al., 1993a, 1995). Therefore, to facilitate comparison, we submitted our participants' data to two-dimensional MDS analysis using the SINDSCAL program (also used by Chan and colleagues). For ANIMALS, the two groups produced very similar spatial maps. The correlation between the two groups' coordinates was .99 for the primary dimension ($p < .001$) and .90 for the secondary dimension ($p < .001$). By inspection, the primary dimension corresponded well to domestic-wild and the secondary dimension corresponded to herbivore-carnivore. For INSTRUMENTS, the spatial maps for the two groups were much less similar. The correlation between the primary dimensions was .54 and between the secondary dimensions was .51 (both ps $< .10$, i.e., approaching significance); however, there was a highly significant correlation between the AD group's primary dimension and the EN group's secondary dimension ($r = .88$, $p < .001$), indicating a dimensional swap between the two groups. The dimension in common between them corresponded well to a wind–nonwind classification. The AD group's secondary dimension could be seen as a band–nonband classification (*drum*, *tuba*, *trumpet*, *trombone*, *flute*, *clarinet vs. harp*, *piano*, *violin*, *cello*, *bassoon*, *oboe*). The EN group's primary dimension defied classification—it was essentially the same as the AD group's band–nonband classification, except that *oboe* and *bassoon* moved from the nonband side into the band side. In summary, the overall picture for the MDS analysis paralleled the Pathfinder analysis: The ANIMALS networks were virtually identical for the AD and EN groups, whereas the INSTRUMENTS networks, although overlapping in part, differed considerably between groups.

Strategy Protocols

As described earlier, after the completion of the flags task with a given domain, the participants were asked to describe how they went about organizing the flags on the flags board. Table 3 summarizes what the participants said regarding the classification schemes by which they organized their flag placement.

There are two points worth noting about the classification data (the implications of these points will be addressed in the Discussion section). First, although the AD participants supplied fewer classification names than did the EN participants (see next paragraph for statistical analyses), the top four classifications and their rank order were the same for the two participant groups for ANIMALS, and three out of the top four EN classifications for INSTRUMENTS were also in the top four of the AD participants. (The exception to this was the classification "reed," which is a much lower frequency word than others in the top four of the EN group.) Second, abstract classifications (i.e., those that cannot be determined exclusively by visual–perceptual features of the named object) were as likely to make it to the top four or five classifications for AD as for EN participants (e.g., wild, domestic, pets, and farm for ANIMALS; wind, blow, and percussion for INSTRUMENTS).

Table 4 provides the mean numbers of classification schemes given for each domain by each participant group, and the percent of participants who reported using a gen-

Table 3. Classification schemes in two domains by AD and EN participants

Animals	
Both (EN/AD): _wild_ 20/14, _domestic_ 13/10, _pets_ 12/8, _farm_ 12/7, _cat family_ 6/3, _herbivore_ 5/2, _jungle_ 3/4, _african_ 4/2, _provide products or services_ 4/2, _wild but nondangerous_ 3/3, _carnivore_ 4/1, _zoo_ 2/3, _docile_ 2/2, _horse family_ 3/1, _size_* 1/3, _wild but dangerous_ 2/2, _circus_ 1/1, _give milk_ 1/1, _hooves_ 1/1, _tamest_ 1/1	AD Only: _how animals get along with each other_* 2, _biters_ 1, _color_* 1, _domesticity_* 1, _fur_ 1, _indoor pet_ 1, _long neck_ 1, _outdoor pet_ 1, _see regularly_ 1, _shape_* 1, _special_ 1, _typical_ 1
	EN Only: _bovine_ 2, _rodent_ 2, _woods or forest_ 2, _biological families_* 1, _claws_ 1, _native_ 1, _nonnative_ 1, _ruminant_ 1, _ungulate_ 1, _wild but trainable_ 1
Musical instruments	
Both (EN/AD): _string_ 24/19, _wind_ 14/9, _blow_ 10/10, _percussion_ 7/6, _reed_ 8/1, _noisy or loud_ 4/4, _strings played with bow_ 7/1, _strike or beat_ 3/3, _brass_ 4/1, _horns_ 3/2, _woodwind_ 4/1, _keys_ 2/2, _strings played with fingers_ 3/1	AD Only: _appearance_* 1, _easy to handle_ 1, _how you play them_* 1, _poke_ 1, _range of notes_* 1, _similarity_* 1, _strum_ 1, _take lessons for_* 1, _tones_* 1
	EN Only: _low tone_ 2, _soprano tone_ 2

Note. Numbers after classifications reflect the number of participants who provided it. For classifications given by both AD and EN, the first number is the EN count, the second is the AD count. Classifications marked with "*" are general principles; others identify a specific subset of items.

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

Table 4. Summary data on strategies described by participants after completing flags placement

	Total named strategies				General principles	
	Animals		Musical instruments		Animals	Musical Instruments
Group	*M*	(*SD*)	*M*	(*SD*)	%	%
AD	2.88	(1.45)	2.22**	(1.41)	28	19
EN	3.53	(1.34)	3.19	(1.20)	6	0

Note. Total named strategies included both named groups (e.g., pets, farm animals) and general principles (domain-wide classification schemes). Percentages are of participants who provided a general principle.
** $p < .01$ for the AD − EN difference in total named strategies.

eral principle (a strategy used to organize the domain as a whole). Note that most participants stated at most one general principle; 2 AD patients stated two of them for ANIMALS. A 2 (group) × 2 (domain) ANOVA on the number of strategies named did indeed show a significant decrease for AD compared to EN participants [$F(1,62) = 8.94$, $p < .01$] and a significant decrease for INSTRUMENTS compared to ANIMALS [$F(1,62) = 6.11$, $p < .05$] with a nonsignificant Group × Domain interaction ($F < 1$).

For both groups and domains, classifications that delineated a subset of items were specified more often than a general organizational strategy. However, the AD participants were much more likely to give general principles than were the EN participants, for both semantic domains; in fact, for INSTRUMENTS, no EN participant gave a general strategy.

We correlated the number of classifications reported by each participant with mental status (as measured by the MMSE) of the AD participants. The correlation for the INSTRUMENTS domain approached significance ($r = .29$, $p < .093$); whereas the correlation for the ANIMALS domain was not significant.

In addition, we correlated the number of classifications given with Pathfinder PFC indices. As described earlier, the PFC indices were derived in each of two ways: Individual PFNETs were compared to those of EN individuals (individual PFC), and individual PFNETs were compared to the PFNET from an overall, EN baseline (baseline PFC). For the AD patients, the number of classifications given was positively correlated with the PFCs: For ANIMALS, $r = .33$, $p = .056$, for individual PFC, and $r = .39$, $p < .05$, for baseline PFC; for INSTRUMENTS, $r = .32$, $p < .05$, for individual PFC, with a nonsignificant correlation for baseline PFC. For EN participants, the number of classifications was positively correlated with individual PFCs for ANIMALS ($r = .31$, $p = .08$) and with baseline PFCs for INSTRUMENTS ($r = .33$, $p = .058$).

DISCUSSION

Probably the clearest result here is that change of context, and in particular, change of task, can produce a profound effect on the connections within the semantic network produced. For the same ANIMALS stimuli (or, the same except for one item, as explained earlier) Chan and colleagues (e.g., Chan et al., 1993a, 1995) found AD networks to be quite different from those of EN controls, while we found their networks to be virtually identical. Furthermore, there is strong evidence that for the same task, AD participants may differ relatively more or less from controls, as a function of semantic domain: For ANIMALS, there was almost no group difference, but for INSTRUMENTS, group differences were apparent. However, based on the intuitions of 18 raters, the connections that occurred only in the AD PFNET do not seem any less correct than those that occurred only in the EN PFNET, for INSTRUMENTS. An MDS analysis did not reveal a deficiency in dealing with abstract dimensions of classification: AD participants' MDS results appeared to reflect the use of four abstract dimensions (tame–wild, herbivore–carnivore, wind–nonwind, and band–nonband), possibly one more than the EN group.

An analysis of strategy protocols showed that many more different strategies were identified by both AD and EN participants for ANIMALS than for INSTRUMENTS—about 50% more. In addition, AD participants reported relatively more general principles that had been used to organize entire domains than EN participants, although both groups reported far more specific classification schemes (identifying a subgroup of items) than general strategies. Furthermore, it is striking how many different strategies were used by both groups in each domain: There was only one classifier in each domain that was used by more than half the participants in either group.

These results are incomprehensible if semantic networks are constant, unvarying mental structures; for example, how can two such different PFNETs as those of our AD and EN participants for INSTRUMENTS both be equally correct? How is it possible (in the case of ANIMALS) for AD participants to produce PFNETs differing considerably from those of EN participants with triadic comparison (as in Chan et al., 1995) yet to produce virtually identical PFNETs with the flags board task? Why are there so many different classification strategies reported by participants? Clearly, more is at work here than simple degradation or absolute preservation of semantic knowledge in AD.

The model suggested by Barsalou can be used to account for the pattern of results we found. The Barsalou model is based on perceptual frames, which are structured representations of perceptual and linguistic experiences. These frames are activated, either by task demands or by controlled searches of memory, and the information in them which is relevant to the task is used to construct a semantic network in working memory. This network, which can expand, contract, or be altered in arbitrary ways as processing continues, is used as a relational database during processing, more or less in the same way that a permanent semantic network might be used in other models. In particular, such networks can be used to generate consistent, reasonable judgments regarding similarities among the items in a domain such as

https://doi.org/10.1017/S1355617799577102 Published online by Cambridge University Press

ANIMALS or INSTRUMENTS in the present experiment. This process may be seen as taking place in two stages. The first is an initial, relatively automatic stage, during which perceptual and linguistic frames are activated, and certain salient relations among them are organized into a simple, initial semantic network. The second stage takes place when the initial semantic network is not sufficient to perform the task: In this case, controlled processes are used to search perceptual and linguistic memory, and also to derive inferences from those items already present in the working-memory semantic network using deductive logic.

There are several points during the process of compiling and using working-memory semantic networks that might affect similarity judgments. First, it may be that there are relatively few perceptual or linguistic experiences that apply to the items in the domain: Perhaps they are unfamiliar, or very abstract. This would cause the initial semantic network to be sparse and poorly organized. Controlled processes might be able to expand the initial networks by finding more relevant perceptual and linguistic frames, or by using existing nodes to infer additional ones. Second, it may be that the items chosen for the domain do not lend themselves to the creation of a simple, consistent network, perhaps because they are not closely related to each other, or perhaps because certain of the relations are highly context-sensitive. In this case, it is unlikely that a complex initial network will be constructed. Induction may be used to create connections between seemingly unrelated nodes. Finally, experimental conditions (or a disease) may limit the attentional resources available to participants for the purpose of searching for or deriving additional information, or for retaining a network of a given complexity in working memory. This would tend to force participants to use a less complete and less consistent semantic network.

The first finding to account for is the high degree of similarity between the AD and EN groups for our ANIMALS domain. Since the items in the domain are familiar, with widely shared perceptual knowledge, and since the items were chosen to fit together well on several obvious attribute classes, the initial semantic networks generated automatically by participants in both groups were probably quite adequate for the task, and little additional controlled processing would have been required in order to elaborate them. Therefore, the result we observed, of very similar networks for the two groups, would be expected.

The second finding to account for is that the INSTRUMENTS domain produced different networks for the two groups. Note that the items are less familiar in this domain—in fact, several participants in both groups needed a quick refresher on the definition of certain INSTRUMENTS (e.g., oboe, bassoon, cello). Therefore, less perceptual–linguistic information was presumably available to participants initially, producing relatively impoverished initial networks. That this occurred is also suggested by the fact that only about two-thirds as many classification strategies were named by participants in both groups for INSTRUMENTS as for ANIMALS. However, the EN group would have been able to use subsequent controlled processes to augment the initial networks somewhat, resulting in a more consistent, more complete database for the task; therefore, their networks tended to differ from those of the AD participants.

The third finding to be accounted for is the difference between the outcome reported in Chan et al. (1993b, 1995) and that of the present experiment. Chan and colleagues used a set of stimuli that was very close to our ANIMALS, but with a Triadic Comparison task. While we think that their set was somewhat less well organized than ours (the substitution of *pig* for *sheep* reduced the domain's symmetry), it is also true that in Triadic Comparison, but not in Flags, participants must depend more upon retaining a temporary semantic network in working memory (in Flags, the names of all the items are always before the participant). Both of these differences probably underlie the different outcomes.

In summary, the present results with the flags board, taken together with those reported in Chan et al. (1993b, 1995) do not lead us to a picture of impoverished semantic knowledge on the part of AD individuals. Rather, it indicates that the construct of a permanent semantic network may not be adequate to explain all of the variation encountered, but instead, a model such as that suggested by Barsalou (1993), involving working-memory implementation of dynamic, task- and individual-specific semantic networks from perceptual and linguistic frames, might be more explanatory. We have suggested that some aspects of the initial construction of these working networks are relatively automatic, and therefore much less affected by AD, but that there are relatively effortful, controlled subsequent processes that are strongly affected by the disease. This brings results reported for semantic memory organization in AD into alignment with those reported for primed lexical decision: When experimental conditions downplay controlled processing, AD participants do not differ from controls; when effortful processing is required, there are large differences (Ober & Shenaut, 1995; Shenaut & Ober, 1996).

The present results have identified task and domain as being critical elements affecting performance by AD participants in network elicitation experiments, but only two domains and a single task have been directly compared. By utilizing several different tasks and several different domains, it should be possible to get a much better understanding of the range of variation in semantic organization. The present experiment is a first step in that direction; the authors are presently carrying out a much larger study involving eight conceptual domains and five tasks (one of which is the flags board used in the present study). This larger study should provide a more complete picture of how the similarity of semantic memory organization in AD and EN participants varies under different circumstances.

ACKNOWLEDGMENTS

This study was supported by grants to B.A. Ober from the National Institute on Aging (R29-AG10848) and the Department of

Veterans Affairs (Medical Research Merit Review Grant). Recruitment of the Alzheimer participants was made possible by the University of California, Davis, Alzheimer's Disease Center, which is supported by the National Institute on Aging (grants AG10129 and AG12435) and the California Department of Health Services Alzheimer's Disease Program (contracts 98-14970 and 98-14971). The authors thank Esther Lara, Jean Coleman, and Carole Bibeau for assisting with participant recruitment, and Jennifer Brustrum, Kimberly Carter, and Victoria Lorrekovich for assisting with participant testing.

REFERENCES

Barsalou, L.W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A.F. Collins, S.E. Gathercole, M.A. Conway, & P.E. Morris (Eds.), *Theories of memory* (pp. 29–102). Hove, U.K.: Lawrence Erlbaum.

Bonilla, J.L. & Johnson, M.K. (1995). Semantic space in Alzheimer's disease patients. *Neuropsychology*, *9*, 345–353.

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

Chan, A.S., Butters, N., Salmon, D.P., Johnson, S.A., Paulsen, J.S., & Swenson, M.R. (1995). Comparison of the semantic networks in patients with dementia and amnesia. *Neuropsychology*, *9*, 177–186.

Chan, A.S., Butters, N., Salmon, D.P., & McGuire, K.A. (1993b). Dimensionality and clustering in the semantic network of patients with Alzheimer's disease. *Psychology and Aging*, *8*, 411–419.

Chertkow, H., Bub, D., & Seidenberg, M. (1989). Priming and semantic memory loss in Alzheimer's disease. *Brain and Language*, *36*, 420–446.

Dearholt, D.W. & Schvaneveldt, R.W. (1990). Properties of Pathfinder Networks. In R.W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 1–30). Norwood, NJ: Ablex Publishing Corporation.

Durso, F.T. & Coggins, K.A. (1990). Graphs in the social and psychological sciences: Empirical contributions of Pathfinder. In R.W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 31–52). Norwood, NJ: Ablex Publishing Corporation.

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

Gammack, J.G. (1990). Expert conceptual structure: The stability of Pathfinder representations. In R.W. Schvaneveldt (Ed.), *Pathfinder associative networks: Studies in knowledge organization* (pp. 1–30). Norwood, NJ: Ablex Publishing Corporation.

Grober, E. & Sliwinski, M. (1991). Development and evaluation of a model for estimating premorbid verbal intelligence in the elderly. *Journal of Clinical and Experimental Neuropsychology*, *13*, 933–949.

LaBarge, E., Balota, D.A., Storandt, M., & Smith, D.S. (1992). An analysis of confrontation naming errors in senile dementia of the Alzheimer type. *Neuropsychology*, *6*, 77–95.

Martin, A. (1992). Degraded knowledge representations in patients with Alzheimer's disease: Implications for models of semantic and repetition priming. In L.R. Squire & N. Butters (Eds.), *Neuropsychology of memory* (2nd ed., pp. 220–232). New York: Guilford Press.

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*, 939–944.

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. & Shenaut, G.K. (1995). Semantic priming in Alzheimer's disease: Meta-analysis and theoretical evaluation. In P. Allen & T.R. Bashore (Eds.), *Age differences in word and language processing* (pp. 247–271). Amsterdam: Elsevier Science B.V.

Ober, B.A. & Shenaut, G.K. (1998, April). *Alzheimer's disease individuals show normal performance on multiple-choice recognition versions of the Boston Naming Test*. Paper presented at the Cognitive Aging Conference, Atlanta, Georgia.

Ober, B.A., Shenaut, G.K., Jagust, W.J., & Stillman, R.C. (1991). Automatic semantic priming with varying types of category relationships in Alzheimer's disease and normal aging. *Psychology and Aging*, *6*, 647–660.

O'Carrol, R.E. (1992). Predicting premorbid intellectual ability in dementia. *Clinical Neuropsychologist*, *6*, 113–115.

Shenaut, G.K. & Ober, B.A. (1996). Methodological control of semantic priming in Alzheimer's disease. *Psychology and Aging*, *11*, 443–448.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327–352.