Differential impairment of semantic and episodic memory in Alzheimer’s and Huntington’s diseases: a controlled prospective study

John R Hodges, David P Salmon, Nelson Butters

Abstract
A controlled prospective study compared the performance of 14 patients with dementia of Alzheimer type (DAT) and 14 patients with Huntington’s Disease (HD), who were matched for overall level of dementia, on a battery of semantic and episodic memory tests. The DAT patients were significantly more impaired on measures of delayed verbal and figural episodic memory, and in addition showed a more rapid rate of decline on tests which depend upon the integrity of semantic knowledge (naming, number information, similarities and category fluency). In contrast, the HD patients were significantly worse, and showed a more rapid decline on the letter fluency test, a task especially sensitive to deficiencies in retrieval. The HD patients were also more impaired than DAT patients on a vocabulary test and on copying geometric figures. The observed double dissociations offer compelling evidence that aetio logically distinct forms of dementia result in different patterns of cognitive impairment.

Whether aetio logically distinct forms of dementia result in qualitatively different patterns of cognitive impairment is a question of major theoretical and practical importance. Albert et al and Cumming and Benson have proposed that dementia associated with primarily subcortical neuropathological changes (for example, progressive supranuclear palsy, Parkinson’s Disease, Huntington’s Disease) is characterised by a profile of intellectual impairment distinct from that seen in dementias which have their initial neuropathology in association cortices (for example, Alzheimer’s Disease). Some studies, utilising standardised clinical batteries with patients in the advanced stages of their illnesses failed to uncover these proposed cognitive differences. However, the more recent application of theoretically based tests, borrowed from experimental and cognitive psychology, have reported important qualitative distinctions among some types of dementia, especially between dementia of Alzheimer type (DAT) and Huntington’s Disease (HD).

The latter findings suggest that DAT involves a severe impairment in episodic memory (that is, memory for temporally specific events or episodes). This deficit has been attributed to defective encoding, an increased sensitivity to proactive interference (that is, the effects of previously learnt material) and a rapid rate of forgetting. Furthermore, DAT patients’ numerous deficiencies in semantic memory, which include impaired naming and difficulty in generating specific exemplars (for example, dog, cat, lion) of abstract categories (for example, animals) on verbal fluency tasks have been associated with a disruption of the hierarchical structure of semantic knowledge. In contrast to patients with DAT, the poor performance of patients with HD on episodic and semantic memory tasks has been associated with a general retrieval deficit. That is, although HD patients appear able to store and retain new information, they exhibit difficulty in initiating systematic retrieval strategies on fluency and recall tasks, whilst the structure of their semantic knowledge per se appears to remain relatively intact.

Despite this growing evidence for a cortical-subcortical distinction, there have been a number of criticisms of the dissociation. Whitehouse stated that “the supporting evidence for the concept is weak, adequate systematic studies comparing different dementias have not been performed” and that “parallel comparison of several types of dementia, using neuropsychological instruments sensitive to a wide range of cognitive abilities” are required. After an extensive review of the literature, similar conclusions were reached by Brown and Marsden. They also addressed the importance of equating groups for their overall levels of dementia and suggested that long-term prospective studies of carefully matched cases are required to validate the cortical-subcortical distinction.

Although many of the most recent comparisons of DAT and HD have matched groups for overall level of dementia, none have followed their patients longitudinally. To assess the issue of progression, our study compared the performances of DAT and HD patients, matched for overall level of dementia, on a battery of episodic and semantic memory tests on two occasions separated by at least 12 months. It was anticipated that DAT and HD patients would show their most rapid decline over the 12 month period on those tests most sensitive to the processes underlying their respective memory disorders. For instance, DAT patients should demonstrate their most rapid decline on tests assessing rate of forgetting and the integrity of semantic knowledge, whereas HD patients should be
most affected over time on tests dependent heavily upon retrieval mechanisms. Tests of naming, verbal abstraction, vocabulary and number information were employed as indices of the status of semantic knowledge, whereas the learning and retention of a word list and figural stimuli served as measures of episodic memory. The subjects' capacity to retrieve information from semantic memory was evaluated with letter and category fluency tests.

Methods

Subjects

Three groups consisting of a total of 42 subjects participated in the study: 14 HD patients (seven males, seven females), 14 patients with DAT (three males, 11 females) and 14 neurologically intact normal control subjects (seven males, seven females). Written informed consent was obtained from all subjects or the caregivers, where appropriate. All subjects were evaluated on entry to the study (Year 1) and 12 months later (Year 2).

The HD patients in the study had been previously diagnosed by a senior neurologist on the basis of a positive family history of the disease, the presence of involuntary choreiform movements and the presence of dementia. Their functional capacity was assessed with Shoulson and Fahn's scale, which rates functional disability from one (minimal) to five (total). Two of the HD patients were rated at Stage 1, six at Stage 2, five at Stage 3 and one at Stage 4.

The diagnosis of probable DAT was made by two senior staff neurologists according to the criteria developed by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), which consist of inclusion and exclusion criteria. All patients presented with progressive cognitive impairment affecting predominantly memory and had a history of between one and four years. All achieved a score of four or less on the Hachinski scale, thus reducing the possibility of multi-infarct dementia.

Normal control subjects were either spouses of patients or volunteers obtained through newspaper advertisements. Subjects with a history of alcoholism, drug abuse, learning disability, serious neurological or psychiatric illness were excluded.

To allow for the level of overall dementia, each HD patient was individually matched with a DAT patient on the basis of the initial (Year 1) scores on the Mini-Mental State Examination (MMSE). The DAT patients were chosen from a data base of approximately 80 patients undergoing prospective evaluation at the University of San Diego Alzheimer's Research Centre (ADRC). As shown in table 1, the mean MMSE scores of these groups were identical at 24.4 (ranges: DAT = 19 to 28; HD = 20 to 29). The Dementia Rating Scale (DRS) scores of the HD group (mean 110-3, range 88 to 124) was lower than that of the DAT group (mean 114.9 range 102 to 134), but this difference was not significant. The HD patients were, as expected, significantly (p < 0.05) younger than the DAT patients, but there was no significant difference between the educational levels of the two patient groups.

The normal controls and DAT patients were also individually matched for age and educational level. The mean age of the normal controls was virtually identical to that of the DAT patients (table 1). The mean score of the normal controls on the MMSE was 28.2 (range 25-30) and on DRS was 138.3 (range 133-144).

Table 1  Mean (SD) age, education, MMSE and DRS scores for the DAT and HD patients and for normal control subjects

<table>
<thead>
<tr>
<th></th>
<th>DAT cases (n = 14)</th>
<th>HD cases (n = 14)</th>
<th>Normal controls (n = 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>73.6 (6.6)</td>
<td>52.6 (12.5)††</td>
<td>73.1 (6.0)</td>
</tr>
<tr>
<td>Education (years)</td>
<td>12.9 (2.7)</td>
<td>14.1 (2.6)</td>
<td>13.3 (1.4)</td>
</tr>
<tr>
<td>MMSE score</td>
<td>24.4 (2.6)*</td>
<td>24.4 (2.7)*</td>
<td>28.2 (3.2)</td>
</tr>
<tr>
<td>DRS score</td>
<td>114.9 (10.4)*</td>
<td>110.3 (13.0)*</td>
<td>138.3 (3.6)</td>
</tr>
</tbody>
</table>

* = significant difference (p < 0.05) between normal controls and patient group.
† = significant difference (p < 0.05) between DAT and HD patient groups.

NEUROPSYCHOLOGICAL TESTS

Semantic memory/language tests

NUMBER INFORMATION TESTS. This test of numerical semantic memory consists of 24 general knowledge questions, each of which has a number for the answer. For example, "How many minutes are there in an hour?" and "At what temperature does water freeze?"

BOSTON NAMING TEST (BNT). Only half of the full 60 outline drawings comprising the BNT were presented at each test session. On the first occasion (Year 1), the even numbered items from the test were administered. Subsequently, at Year 2, the odd numbered items were administered. If the subjects were unable to name an object, a predetermined stimulus cue was given, and if still unable to name the item, a phonemic cue was provided. Two scores were obtained: the total number of items named spontaneously and the total number named after stimulus cueing. There was no evidence of bias in the selection of items for the shortened version of the test since the normal control subjects obtained virtually identical mean scores on the two versions.

VOCABULARY ('WAIS-R'). This test requires subjects to define 35 words, the frequency of which ranges from very common to moderately uncommon. No time constraint was applied. The subject's responses (that is, definitions) were scored in accordance with the criteria described by Wechsler.

SIMILARITIES TEST ('WAIS-R'). This test assesses both semantic knowledge and abstraction, and requires subjects to state the way in which 14 pairs of stimulus items are the same (for example, apple-orange, poem-statue). Subjects' responses were scored in accordance with the criteria described by Wechsler.

VERBAL FLUENCY TESTS. A letter fluency task and a category fluency task were administered to each subject. For the letter fluency test, the subjects were asked three letters "F", "A", and...
“S” sequentially and asked to produce “as many different words as you can think of” that begin with the given letter. They were instructed that proper nouns and the repetition of words with a different suffix (for example, find, finding, finder) were not acceptable. On the category fluency task, the subjects were asked to produce as many exemplars as possible from the animal category, followed by the vegetables and fruit categories. For each of the three letters and three categories, the subjects were allowed 60 seconds to generate words orally.

**Episodic Memory Tests**

**Buschke-Fuld Selective Reminding Test (BSRT).** Subjects were read a list of 10 unrelated words at a rate of one word every two seconds and then immediately asked to recall the entire list. On the second trial the subjects were read only those words they failed to recall on the first trial. At the end of the second trial they were again asked to recall the entire list. This procedure of only presenting those words subjects failed to recall on the preceding trial was followed for six trials. Scores were thereby obtained for the total number of words correctly recalled over the six trials, the total number of words recalled from short-term storage (that it, those items recalled that were on the immediately preceding list), and the total number of words recalled from long-term storage (that is, those items correctly recalled that were not on the immediately preceding list).

**Russell’s Adaptation of the Visual Reproduction Test.** This modification of the Visual Reproduction Test (VRT) from the Wechsler Memory Scale assesses memory for geometric shapes. On each of three trials, the subjects were shown for 10 seconds a card with an outline drawing of a complex geometric figure. Immediately following the presentation, the subjects were asked to reproduce the figures from memory on a blank piece of paper. On the first two trials, each presentation card had one figure; the presentation card for the third trial had two geometric patterns. After a delay of 30 minutes, filled with unrelated testing, the subjects were asked without previous warning to reproduce the three figures. Finally, subjects were instructed to simply copy the stimulus figures to assess any visuospatial dysfunction that might have contaminated visual memory performance. Using a modification of the scoring criteria developed by Wechsler, scores for immediate recall, delayed recall and copying were obtained on the basis of the number of accurately drawn components from the original figures.

**Statistical Methods**

To compare the initial (Year 1) results in the two patient groups and the normal controls one-way analyses of variance (ANOVAs) with post-hoc two tailed t tests were used. To compare the Year 1 and Year 2 results in DAT and HD cases, two-way repeated ANOVAs with post-hoc two tailed t tests were employed.

**Results**

Table 2 presents the mean test scores of the three groups (HD, DAT, normal controls) on entry to the study (Year 1). One-way ANOVAs revealed highly significant group effects for all tests (p’s < 0.01 to < 0.001, see table 2 for F values). Post-hoc comparisons showed that both patient groups were significantly impaired (p < 0.01) relative to controls on all tests, with the exception of the copying procedure of the VRT which was performed normally by the DAT patients and recall from short-term memory on the BSRT which was normal in the HD patients. Comparisons of the two patient groups showed that the HD patients’ performances were significantly worse than those of the DAT patients on letter fluency and category fluency (p’s < 0.05), Vocabulary and the copying part of the VRT (p’s < 0.01). The performances of the patients with DAT on the delayed recall of the VRT (p < 0.01) and recall from long-term memory on the BSRT (p < 0.05) were significantly worse than those of the HD patients.

Since there were no significant changes in the normal controls’ mean scores obtained on
Years 1 and 2, longitudinal analyses were limited to the two patient groups. A comparison of the DAT and HD patients' Year 1 and Year 2 data are presented in figs 1 to 4.

On the following measures of semantic memory the DAT patients declined more rapidly than the HD patients: BNT, Number Information Test and Similarities Test (fig 1). A two (groups) × two (years) ANOVA of the total BNT scores yielded a significant interaction [F (1,24) = 5.27, p < 0.05], but no significant main effects for groups or years. Post-hoc t tests revealed a significant difference between the DAT group's Year 1 and Year 2 scores [t (13) = 2.36, p < 0.05] on the BNT, but no difference between the HD group's two test scores on the naming test. Likewise, analysis of the scores on the Number Information Test revealed a significant group × year interaction [F (1,24) = 5.35, p < 0.05], but in this case there was also a significant year effect [F (1,24) = 9.74, p < 0.01]. Post-hoc t tests showed a significant difference between the DAT group's Year 1 and Year 2 scores [t (13) = 3.30, p < 0.01], but no significant change in the HD group's performance over this period. For the Similarities test there was also a significant group × year interaction effect [F (1,22) = 4.70, p < 0.05], but no significant main effects for groups or years. Post hoc t tests indicated that the difference between the DAT group's Year 1 and Year 2 scores was significant [t (13) = 2.56, p < 0.05], but again the change in the HD patients' performance did not reach statistical significance.

For Vocabulary, a different pattern of results emerged. A two × two ANOVA yielded a significant group [F (1,20) = 4.98, p < 0.05] and year effect [F (1, 20) = 18.77, p < 0.001], with no significant interaction between these two main factors. Thus as seen in fig 1, the HD patients had lower vocabulary scores than the DAT patients, and there was no evidence of a differential rate of decline between the groups.

As shown in fig 2, the HD patients also obtained inferior scores on both measures of verbal fluency. A 2 × 2 ANOVA of the letter fluency (FAS) scores revealed a significant main effect of groups [F (1,24) = 9.05, p < 0.01] and for years [F (1,24) = 7.99, p < 0.01], but the interaction between these two factors was not significant. A 2 × 2 ANOVA of the category fluency scores (animals, fruit, vegetables) revealed a significant year effect [F (1,24) = 11.56, p < 0.01] and the group effect approached significance [F (1,24) = 3.61, p = 0.07]. The interaction between group and year factors was not significant. Planned analyses of change over time on the two fluency measures suggested a pattern of double dissociation. The HD group's performance on the letter fluency test declined significantly between Years 1 and 2 [t (13) = 2.77, p < 0.05] but they exhibited no significant change in category fluency over this time period [t (13) = 1.37, p > 0.05]. For the DAT group the opposite pattern was found; their performance on the category fluency test declined significantly between Year 1 and 2 [t (13) = 2.68, p < 0.05] but they demonstrated a lack of significant change in letter fluency [t (13) = 0.98, p > 0.05].

In contrast to the verbal fluency tests, the DAT patients were significantly impaired relative to the HD patients on measures of delayed verbal and figural memory (figs 3 and 4). A two × two ANOVA of the total words recalled over the six trials of the BSRT (fig 3) revealed a significant main effect of years [F (1,18) = 5.10, p < 0.05], but no significant effects for groups or the interaction between groups and years. However, individual comparisons of the items correctly retrieved from short- and long-term memory did show significant group differences. DAT patients recalled more items from short-term memory [F (1,18) = 5.35, p < 0.05] than did the HD patients, whereas HD patients retrieved more items from long-term memory than did the DAT patients [F (1,18) = 6.35, p < 0.05]. There was also a significant year effect [F (1,18) = 7.08, p < 0.05] for long-term memory. The failure to find a significant interaction between groups and years on long-term memory may be due to floor effects associated with the DAT patients' severely impaired performance on Year 1.

A two × two ANOVA of the immediate recall scores from the VRT (fig 4) failed to show significant group, year, and interaction...
on a letter fluency task often used to assess individual's retrieval capacities. Since the HD patients' performances on a number of tests sensitive to disruption of semantic memory (BNT, Number Information, Similarities and category fluency) remained stable over time, it would appear that the actual structure of their semantic knowledge remained relatively intact despite their progressive retrieval deficit.

Although the longitudinal differences between the two patient groups' performance on the letter and category fluency tests were small, and barely reached significance, they do deserve comment since they may provide important cues about the differential effects of the disease processes on these two tasks. Category fluency requires the systematic retrieval of hierarchically organised information from semantic memory, whereas letter fluency can be successfully performed by relying solely upon phonemic or lexical cues to guide the retrieval process. This distinction in the two tests dependence on the structure of semantic knowledge may explain why DAT patients show significant decline on category, but not letter, fluency over a one-year period. Since the structure of semantic knowledge is severely affected in DAT, the patients' rapid decline on category fluency may be indicative of a progressive deterioration in this domain. Similarly, the HD patients' very impaired performance on both fluency tasks may reflect the importance of retrieval processes for all such word generation tests. If retrieval per se is impaired, it should matter very little whether one uses phonemic, lexical or semantic cues to guide the search processes.

The vocabulary test yielded somewhat unexpected results. Although performance on this test also relies upon semantic knowledge the HD patients were significantly more impaired than the DAT patients. This result cannot be explained on the basis of different premorbid educational levels. Several factors may have contributed; since scoring criteria for the test systematically reward more complete and elaborate responses, general psychomotor slowing, retardation in the activation of associative semantic networks in HD and dysarthria may all play some role in limiting these patients' responses. This test probably places greater demands on initiation and retrieval mechanisms than the other measures of semantic memory employed in the study. It is highly unlikely that a motoric articulatory defect accounts entirely for the impairment in vocabulary since the HD patients' degree of dysarthria was minimal or mild in all cases, the vocabulary test was performed without time constraint, their performance on other language based tests (BNT and Number Information) remained stable over the course of the study and there was no deterioration in their performance on the copy part of the VRT which clearly reflects motor function.

As predicted, the DAT patients' performance on tests of verbal and figural episodic memory remained inferior to that of the DAT patients' over the time course of the study. The
results for the BSRT indicated that the DAT patients, in comparison to the HD patients, had proportionately greater impairment of long-term memory than of short-term memory. This difference is consistent with the fate of encoding, storage and retrieval processes in the two diseases. Since HD involves the relative preservation of encoding and storage and BSRT, in combination with severe retrieval deficits, access to information in both short- and long-term memory are negatively influenced. On the other hand, since DAT is characterised by severe deficiencies in encoding and storage, very little new information is successfully transferred to long-term memory. This difficulty in establishing long-term stores results in the DAT patients’ short-term recall being relatively superior to their long-term storage. For the VRT, the significant difference between the DAT and HD groups on the delayed, but not on immediate, recall of the figural stimuli confirms other reports of rapid rates of forgetting in DAT patients. The failure to demonstrate a significantly different rate of decline between the patient groups on the BSRT and VRT was likely due to the floor effects imposed by the DAT patients’ extremely impaired performance at Year 1.

Despite the DAT patients’ significantly worse recall of the geometric figures on the VRT, their performance on the copy condition was significantly better than that of the HD patients. Although the HD patients’ movement disorder clearly contributed to their copying impairment, these patients’ well-documented visuo-perceptual and visuo-spatial deficits may also have played a significant role in their poor drawings.

The present data have special relevance to the methodological criticisms recently levelled against neuropsychological studies in which groups of patients with aetiologicaly distinct dementing diseases are compared. As noted by Brown and Marsden, it is essential that all subjects fulfil strict diagnostic criteria and that groups be matched for their overall level of dementia. Furthermore, they stressed that although cross-sectional studies can offer a degree of support for the proposed qualitative distinctions between types of dementia, longitudinal studies are required to confirm the validity of such findings. This study addressed these issues. The patients met rigorous diagnostic criteria, and were carefully matched on an individual matched-pair basis for their overall level of dementia. The longitudinal analyses showed that differences in test results between the groups either remained consistent or became accentuated with time. Since DAT and HD characteristically afflict patients of very different ages it is not possible to perform cross-sectional or longitudinal studies on age-matched patients with the same overall level of dementia. However, the difference in age between the groups actually accentuates the cognitive distinction and thereby strengthens the overall conclusions. On the basis of age, it would be expected that the DAT groups’ performance on virtually all tests would be inferior to that of the HD group, whereas the reverse was found on a number of tests. Furthermore, the more rapid decline in the DAT patients’ performance on measures of semantic memory cannot be explained on the basis of ageing since the age-matched normal controls showed no decline on any of the tests over the course of the study.

Our results, therefore, highlight the limitations of brief global cognitive assessment scales in patients with early dementia. Although many of the DAT patients scored above the normally accepted cut-off (24) for dementia on the MMSE when more sensitive tests of episodic and semantic memory were applied all subjects showed significant impairment.

In conclusion, the observed double dissociation between semantic and episodic memory processes in DAT and HD offer compelling evidence that these diseases do indeed produce qualitatively different patterns of cognitive impairment. The observed support to the storage-retrieval dichotomy proposed for cortical and subcortical dementing diseases. Semantic knowledge may depend primarily upon the integrity of cortical association areas, whereas the ability to initiate systematic retrieval may be compromised by basal ganglia dysfunction. Before the concept of cortical and subcortical dementia can be fully accepted, however, further detailed longitudinal comparisons are needed using patients with other subcortical diseases, such as progressive supranuclear palsy and Parkinson’s Disease.

This study was supported in part by funds from the VA Medical Research Service and by NIA grants AG-05131 and AG-08204, and was performed whilst one of the authors (JRH) was in receipt of a Medical Research Council Travelling Fellowship.

Differential impairment of semantic and episodic memory in Alzheimer’s and Huntington’s diseases