

# Qualitative neuropsychological performance characteristics in frontotemporal dementia and Alzheimer's disease

J C Thompson, C L Stopford, J S Snowden, D Neary

*J Neurol Neurosurg Psychiatry* 2005;76:920–927. doi: 10.1136/jnnp.2003.033779

See end of article for authors' affiliations

Correspondence to:  
Jennifer Thompson,  
Cerebral Function Unit,  
Greater Manchester  
Neuroscience Centre,  
Hope Hospital, Salford M6  
8HD, UK; jennifer.  
thompson@manchester.  
ac.uk

Received 3 December 2003  
In revised form  
19 October 2004  
Accepted 20 October 2004

**Background:** Frontotemporal dementia (FTD) and Alzheimer's disease are clinically distinct disorders, yet neuropsychological studies have had variable success in distinguishing them. A possible reason is that studies typically rely on overall accuracy scores, which may obscure differences in reasons for failure.

**Objectives:** To explore the hypothesis that analysis of qualitative performance characteristics and error types, in addition to overall numerical scores, would enhance the neuropsychological distinction between FTD and Alzheimer's disease.

**Methods:** 38 patients with FTD and 73 with Alzheimer's disease underwent assessment of language, visuospatial abilities, memory, and executive function, using a neuropsychological screening instrument and standard neuropsychological tests. In each of these cognitive domains, performance characteristics and error types were documented, in addition to numerical scores on tests.

**Results:** Whereas comparison of neuropsychological test scores revealed some group differences, these did not occur consistently across tests within cognitive domains. However, analysis of performance characteristics and error types revealed qualitative differences between the two groups. In particular, FTD patients displayed features associated with frontal lobe dysfunction, such as concrete thought, perseveration, confabulation, and poor organisation, which disrupted performance across the range of neuropsychological tests.

**Conclusions:** Numerical scores on neuropsychological tests alone are of limited value in differentiating FTD and Alzheimer's disease, but performance characteristics and error types enhance the distinction between the two disorders. FTD is associated with a profound behavioural syndrome that affects performance on cognitive assessment, obscuring group differences. Qualitative information should be included in neuropsychological research and clinical assessments.

Frontotemporal dementia (FTD) is the most common clinical syndrome associated with frontotemporal lobar degeneration (FTLD). It is associated with a non-Alzheimer pathology that primarily affects the frontal and temporal cortex. Alterations in personality and behaviour are the most salient clinical features of FTD and are pivotal to the diagnostic criteria.<sup>1 2</sup> This contrasts sharply with Alzheimer's disease, where the diagnosis is principally dependent on the presence of cognitive symptoms.<sup>3</sup>

Despite the striking behavioural abnormalities, cognitive impairment is also a central feature of FTD. Patients show "frontal" executive impairment and clinical criteria suggest that this is in the absence of severe amnesia, aphasia, or perceptuospatial disorder.<sup>1 2 4</sup> This pattern differs from the cognitive changes observed in Alzheimer's disease—namely, deficits in episodic memory, language and perceptuospatial function.<sup>5</sup> These cognitive differences are consistent with functional imaging studies that have shown anterior hemisphere abnormalities in FTD and posterior hemisphere abnormalities in Alzheimer's disease.<sup>6 7</sup>

Given the contrasting clinical profiles and patterns of CBF abnormality, one might expect direct neuropsychological comparison of FTD and Alzheimer's disease to yield a double dissociation, with FTD patients performing poorly on tests of executive function, but showing relatively preserved performance on tests of memory and perceptuospatial abilities, and Alzheimer patients showing the reverse pattern of findings. It is surprising, then, that comparative neuropsychological studies have produced inconsistent and sometimes counter-intuitive findings. Some investigators have reported impaired performance on tests of perceptuospatial function in Alzheimer's disease relative to FTD,<sup>8–12</sup> whereas others have

reported no difference.<sup>13–17</sup> Similarly, some studies have shown worse performance on tests of memory in Alzheimer's disease than in FTD,<sup>10 14–16 18</sup> whereas others have found commensurate performance in both groups.<sup>8 9 11–13</sup> Greater impairment of executive function in FTD has been found by some investigators,<sup>10 11 14–16</sup> but a few studies have shown no difference between FTD and Alzheimer's disease on these tasks.<sup>12 13 17 19</sup>

Differences in study design are likely to account for some of these conflicting results. However, the lack of clear differentiation between these two clinically distinct disorders may reflect the fundamental nature of neuropsychological assessment. Regardless of how well standardised and validated a neuropsychological instrument may be, there is no one to one correspondence between a specific test and a neuropsychological function. Tasks may certainly be sensitive to one aspect of cognition, but are never *specific* to that cognitive domain. For example, counting arrays of dots is a simple test of spatial function, included in a battery of tests designed to tap perceptual and spatial function.<sup>20</sup> However, an inattentive, impulsive patient might also fail to count dots accurately for reasons of inattentiveness and lack of concern for accuracy. Thus a spatial task may be failed for executive reasons. Similarly, a putative test of executive function that makes spatial demands may be failed for spatial rather than executive reasons. It is possible, then, that FTD and

**Abbreviations:** FTLD, frontotemporal lobar degeneration; FTD, frontotemporal dementia; MMSE, mini mental state examination; SPECT, single photon emission tomography; VOSP, visual object and space perception

Alzheimer patients may achieve similar scores on tasks for different underlying reasons.

Neuropsychological test scores are essentially summary statements about a fairly narrowly defined behavioural response.<sup>21</sup> Despite being ostensibly objective, test scores rarely reveal information about *how* a particular task was carried out, or the types of error committed. Such information could shed light on the underlying reason for impairment on a given test. Consideration of qualitative aspects of test performance may be especially pertinent in FTD, given the behavioural abnormalities that characterise the condition, which include disinhibition, mental rigidity and inflexibility, concrete thought, distractibility, impersistence, economy of mental effort, and perseverative behaviour.<sup>22</sup> It would be reasonable to suppose that such behavioural abnormalities would affect performance on tests across a variety of cognitive domains.

Our aim in this study was to examine both quantitative and qualitative aspects of neuropsychological test performance in a group of clinically diagnosed FTD and Alzheimer patients. We predicted that a greater memory and perceptuospatial disorder in Alzheimer's disease and a greater executive impairment in FTD would lead to some performance differences between the two groups on tests tapping these domains. However, we expected that such performance differences would not be consistent or clear cut, but that analysis of error types and performance characteristics would enhance the differentiation between FTD and Alzheimer's disease.

## METHODS

### Patients

The study group comprised 111 patients with a clinical diagnosis of Alzheimer's disease ( $n = 73$ ) or frontotemporal dementia ( $n = 38$ ). The Alzheimer's disease group contained a significantly greater proportion of female patients than the FTD group, which was more evenly matched in terms of sex (FTD, 16 female; 22 male; Alzheimer's disease, 45 female; 28 male ( $p = 0.047$ )). Duration of symptoms did not differ significantly between the groups (mean (SD): FTD, 3.2 (2.2) years; Alzheimer's disease, 2.6 (1.6) years). The mini-mental state examination (MMSE) score was significantly higher in the FTD group (FTD, 22 (6); Alzheimer's disease, 18 (5) ( $p = 0.002$ )).

All patients underwent detailed history taking and neurological examination as part of their diagnostic work up. Patients were interviewed with their principal carer by a consultant neurologist in order to determine the evolution of their condition, with particular reference to symptoms arising from problems in language, visuospatial function, memory, and behaviour. The neurological examination was directed at eliciting signs of visuospatial disorientation, speech impairment, alterations in power, tone, and coordination, and the presence or absence of primitive reflexes. Diagnoses were made on the basis of historical information, clinical interview, and neurological examination, and were supported by findings on structural (magnetic resonance) and functional imaging (single photon emission computed tomography (SPECT)). All patients met currently accepted diagnostic criteria for FTD or Alzheimer's disease.<sup>1-3</sup> The cognitive and behavioural features are those that can be elicited from clinical interview.

### Inclusion criteria

The study group comprised consecutive referrals to a regional diagnostic dementia clinic presenting between February 1999 and February 2003, in whom a full set of neuropsychological data was available. Only patients in whom a diagnosis of FTD or Alzheimer's disease was evident on the basis of clinical history and supported by neuroimaging were included in the

study, to avoid the inherent circularity of defining patients on the basis of neuropsychology. The FTD group contained only individuals with the behavioural syndrome of frontotemporal dementia, and not semantic dementia or progressive non-fluent aphasia. Individuals with a mixed clinical syndrome were not included in the study.

### Evaluation

Data collection was based on the patients' performance on neuropsychological tests administered as the time of their initial assessment. Each patient underwent evaluation using a locally constructed neuropsychological screening instrument,<sup>23,24</sup> tapping language, perception, spatial function, praxis, memory, and executive abilities. Examiners record both accuracy scores and qualitative aspects of performance, such as error types and behavioural characteristics of the patient. Neuropsychological profiles derived from the instrument are strongly related to the patterns of abnormality on SPECT<sup>25</sup> and, when combined with historical and neurological information, have yielded a 97% accuracy rate of clinical diagnosis in over 200 pathologically confirmed cases of dementia in this centre. In addition to the neuropsychological screening instrument, standardised neuropsychological tests of language and perceptuospatial function were also included in the initial assessment. The neuropsychological tests included in the present study were selected to represent language, perceptuospatial function, memory, and executive function. For each of these domains of cognitive function, performance characteristics and error types were recorded, in addition to numerical accuracy scores.

### Neuropsychological assessment

As noted in the introduction, it is intrinsic to the nature of neuropsychological tests that they tap multiple cognitive functions. For example, presentation of line drawings can be used to assess visual perception, semantic access/knowledge, and naming. Similarly, presentation of photographs of celebrities can be used to assess visual perception, semantic access/knowledge, naming, and memory. For convenience, tests have been grouped under language, perceptuospatial function, memory, and executive function.

### Language

#### Accuracy scores

- *Graded naming test.* Total naming score on the graded naming test (/30).
- *Boston naming test.* Number of items named on a shortened version of this test (/10).
- *Naming famous faces.* Ability to name famous faces taken from the entertainment media and political sphere (/9).
- *Repetition.* Ability to repeat monosyllabic and polysyllabic word strings and short phrases (/6).
- *Sentence comprehension.* Ability to interpret sentences correctly, based on individual word semantics, syntactic structure, and relational terms (/11).
- *Metaphor interpretation.* Ability to abstract meanings of common metaphors and proverbs (/4).

#### Performance characteristics and error types

- *Graded naming test error types.* Errors were categorised as follows: superordinate semantic category (for example, "animal" for kangaroo); coordinate semantic category ("koala" for kangaroo); phonological ("cuttlshock" for shuttlecock); functional description ("to open a bottle" for corkscrew); omission ("don't know"/no response); perceptual error ("glasses" for handcuffs); perseverative response (repetition of an earlier response).

- *Repetition error types.* The following were recorded: omissions of words; word order errors; phonological errors; concrete responses (for example, an instruction to repeat “an uncommon animal” elicits the response “giraffe”).
- *Sentence comprehension.* The presence or absence of concrete responses was recorded. (for example, the question “would you use a saw to cut wood?” elicits the response “I’ve got a gas fire so I don’t need to cut wood.”)
- *Metaphor interpretation.* The presence or absence of concrete explanations was recorded (for example, an interpretation of “every cloud has a silver lining” as “something to do with the weather”).

## Perceptuospatial

### Accuracy scores

- *Line drawing identification.* Ability to provide appropriate identifying information for line drawings of objects taken from the Boston naming test (/10).
- *Celebrity identification.* Ability to provide appropriate identifying information for famous faces taken from the entertainment media and political sphere (/9).
- *Visual object and spatial perception battery.* Total scores taken from the following subtests: screening; incomplete letters; silhouette perception; object decision; dot counting; position discrimination; number location; cube analysis.
- *Visual construction.* Ability to construct accurate copies of four line drawings: interlocking pentagons taken from the MMSE; a flower; a simple abstract design, and a simplified version of the Rey–Osterreth figure (/4).

### Performance characteristics and error types

Qualitative assessment of perceptuospatial ability was limited by lack of consistency in reporting error types in some tests. Reliable data could be obtained from the visual construction test.

- *Visual construction error types.* The presence or absence of the following was recorded: spatial errors (that is, drawing individual elements of a figure in isolation or in incorrect spatial alignment), perseverative strokes (repeated redrawing of part of the figure in the same location; omission of detail; organisational/planning error (for example, individual elements drawn before overall outline leading to poor “fit” or leaving insufficient space on the page to complete design; overelaboration of detail (such as addition of extra details not present in the original). Attribution of error type was based on evaluation of drawings in addition to written report of the original neuropsychological assessor.

## Memory

### Accuracy scores

- *Immediate memory.* Forward and reverse digit span.
- *Story recall.* Number of elements recalled from a short fable-like story. Free and cued recall was assessed immediately after presentation and again following a 30 minute delay. The proportion of information retained following a delay was calculated by dividing delayed recall by immediate recall.
- *Celebrity recall.* Photographs of celebrities were presented as part of the language/perception assessment. Free recall of these celebrities was assessed 30 minutes after presentation, by means of naming or provision of identifying information (/9).

### Performance characteristics and error types

- *Story recall error types.* The presence or absence of the following was recorded: intrusions (responses irrelevant to the task relating to an earlier test); confabulation (introduction of fictional elaborative elements in recall); misconstruction (misordering or misconstruing elements resulting in a confused, disordered account).

## Executive functions

### Accuracy scores

- *Category fluency.* Number of animal names produced in one minute.
- *Letter fluency.* Number of words beginning with “F” produced in one minute. Proper nouns and derivations of the same word stem (for example, fill, filling, filler) were not admissible.
- *Weigl’s test.* Grouping a set of blocks according to colour, shape, and motif. Three points were awarded for each correct uncued sort, two points if the patient could complete the sort following a partial grouping by the examiner, and one point if the patient succeeded in identifying the sorting dimension following a complete sort by the examiner (/9).

### Performance characteristics and error scores

- *Verbal fluency error types.* The number of repetitions and rule violations was recorded (for example, production of a proper noun in the letter fluency task).
- *Weigl’s test errors.* The presence or absence of: perseveration of a particular sorting category; dismantling of the blocks when partially/completely sorted by the examiner and regrouping according to a previously relevant dimension, or idiosyncratic sorting rules.

## Behavioural observations

Examiners’ reports comprise a written description of patients’ behaviour during the assessment, in which it is noted whether their behaviour was socially appropriate, whether they cooperated with the assessment, and whether they engaged in tests with appropriate effort and concern for accuracy and maintained attention. Conversely, it is also noted if patients are easily distractible, are disinhibited, or show economy of effort—that is, if they give rapid, cursory responses to questions or abandon tests readily. For the purposes of this study, the presence or absence of social appropriateness, disinhibited behaviour, and economy of effort was recorded.

## Procedure

Two independent raters, JT and CS, documented numerical accuracy scores, performance characteristics, and error types based on raw test data and documentation in clinical reports.

## Statistical analysis

Data analyses were carried out using SPSS (version 10) and StatsDirect (version 2.3.7). Logistic regression analyses with clinical diagnosis as the dichotomous outcome variable (FTD = 1/Alzheimer’s disease = 0; classification cut off point = 0.5) was used to calculate the degree to which individual items yielded an increase or a decrease in the odds of the diagnosis being FTD as opposed to Alzheimer’s disease. Increases or decreases in odds were considered statistically significant at  $p < 0.05$ . Continuity corrections were applied to the data where items were only present in one or other group. Items that yielded a significant increase or decrease in odds were entered into further logistic regression analyses to determine the predictive value of quantitative and qualitative

data. The frequency distributions of the two groups of patients according to the examiners' behavioural observations were compared using  $\chi^2$  tests.

## RESULTS

### Neuropsychological assessment

#### Language

Total scores on the graded naming test were of no value in predicting diagnosis (table 1). However, error responses were of significant value: phonological errors were predictive of Alzheimer's disease, whereas perseverative responses increased the probability of a FTD diagnosis.

Naming of either line drawings from the Boston naming test or celebrities did not differentiate the groups. Overall accuracy on a test of repetition was commensurate, but the presence of concrete errors significantly increased the likelihood of an FTD diagnosis, while phonological errors increased the likelihood of an Alzheimer diagnosis. FTD patients' performance was poorer on sentence comprehension and metaphor interpretation. The presence of concrete errors on these tasks was a significant predictor of FTD.

#### Perceptuospatial

The ability to identify line drawings or celebrities did not differentiate the groups (table 2). On the visual object and space perception (VOSP) battery, Alzheimer patients showed poorer performance on some subtests (incomplete letters, position discrimination, cube analysis), but not others (silhouette identification, object decision, dot counting, number location). On a visual construction task, Alzheimer patients produced fewer correct drawings. There were also differences in the type of errors committed. A large proportion of Alzheimer patients made spatial errors, whereas none of the FTD patients did. Thus spatial errors were highly predictive of Alzheimer's disease. By contrast, organisational errors, perseverations, or

overelaborated copies were significant predictors of FTD. Figure 1 shows examples of drawings completed by the two patient groups. The copies carried out by an FTD patient showed poor organisation but preserved spatial configuration. By contrast, the copies carried out by an Alzheimer patient showed poor spatial configuration, with individual elements drawn in isolation.

#### Memory

Alzheimer patients were impaired in story recall (both immediate and delayed) and celebrity recall relative to FTD patients (table 3). Of note was that there was a marked difference in the amount of information retained over a delay in story recall. In the Alzheimer group, patients recalled a median of 0% of the information produced on immediate recall, whereas FTD patients produced 100% of their original recall. There were also qualitative differences between the groups: misconstrued or confabulatory accounts significantly increased the odds that the patient had FTD, whereas intrusions of unrelated information increased the odds that the patient had Alzheimer's disease.

#### Executive functions

On a letter fluency task, FTD patients produced fewer correct responses and made more rule violations, although this was not statistically significant (table 4). Alzheimer patients produced more repetition errors on a category fluency test. Overall scores on the Weigl's test did not differentiate the groups. However, when individuals dismantled blocks partially sorted by the examiner, and regrouped them perseveratively this increased the likelihood of an FTD diagnosis.

#### Effect of dementia severity

In order to ensure that the observed pattern of findings did not reflect the inclusion of more severely affected individuals,

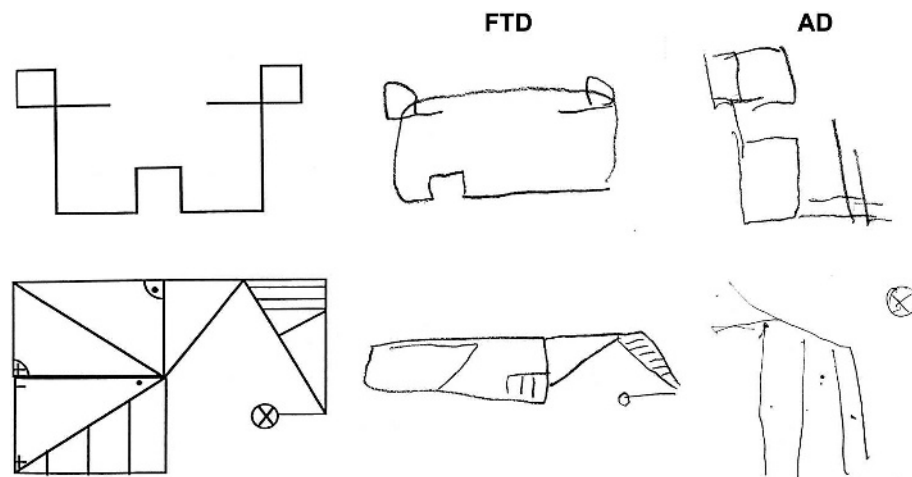
**Table 1** Language

Language	Median score		Range		p Value
	FTD	AD	FTD	AD	
<b>No of correct answers</b>					
Graded naming test (/30)	15	14	0 to 25	0 to 27	0.805
Naming line drawings (BNT) (/10)	10	10	1 to 10	4 to 10	0.819
Naming famous faces (/9)	6	4	0 to 9	1 to 9	0.100
Repetition (/6)	5	4	0 to 6	0 to 6	0.256
Sentence comprehension (/11)	<b>9</b>	<b>10</b>	<b>0 to 11</b>	<b>5 to 11</b>	<b>0.001*</b>
Metaphor comprehension (/4)	<b>1</b>	<b>3</b>	<b>0 to 4</b>	<b>0 to 4</b>	<b>0.000*</b>
<b>Error types</b>					
<i>Graded naming test errors</i>					
Superordinate category	1	1	0 to 7	0 to 15	0.736
Coordinate category	3	2	0 to 11	0 to 7	0.080
Phonological	<b>0</b>	<b>0</b>	<b>0 to 3</b>	<b>0 to 6</b>	<b>0.019†</b>
Functional descriptions	4	5	0 to 12	0 to 17	0.569
Omissions	3	3	0 to 28	0 to 18	0.285
Perceptual	1	1	0 to 4	0 to 15	0.052
Perseverative	<b>0</b>	<b>0</b>	<b>0 to 18</b>	<b>0</b>	<b>0.006*</b>
<b>Errors</b>					
<i>Repetition errors</i>					
<b>Per cent of subjects</b>					
Omission errors in repetition	31%	43%			0.273
Word order errors in repetition	11%	12%			0.893
Phonological errors in repetition	<b>9%</b>	<b>43%</b>			<b>0.001†</b>
Concrete errors in repetition	<b>17%</b>	<b>4%</b>			<b>0.034*</b>
<i>Comprehension errors</i>					
Concrete sentence interpretation	<b>11%</b>	<b>0</b>			<b>0.000*</b>
Concrete metaphor interpretation	<b>56%</b>	<b>24%</b>			<b>0.001*</b>

Bold type indicates that the item significantly changed the odds of one or other diagnosis: \*FTD performance worse than AD; †AD performance worse than FTD.

Analyses were based on a minimum of 73 Alzheimer patients and 33 FTD patients.

AD, Alzheimer's disease; BNT, Boston naming test; FTD, frontotemporal dementia.



**Figure 1** Examples of visual construction in frontotemporal dementia (FTD) and Alzheimer's disease (AD).

the analyses were repeated excluding those with MMSE scores of 15 and under, leaving 48 Alzheimer and 29 FTD patients. Although the pattern of results was broadly similar, a few differences emerged. The number of line drawings from the Boston naming test now differentiated the groups ( $p = 0.045$ ), with Alzheimer patients naming fewer than FTD patients. Overall score on the Weigl's test also differentiated the groups ( $p = 0.033$ ), with FTD patients achieving lower scores than Alzheimer patients.

The majority of tests and qualitative features remained statistically significant; however, the following no longer differentiated the groups: number of phonological errors on the graded naming test, and the presence of perseverative strokes and poor organisation on visual construction.

#### Predictive value of quantitative and qualitative data

A logistic regression analysis was conducted to determine the statistical value of numerical scores and error types in predicting diagnosis. Items that significantly increased or

decreased the odds of FTD as opposed to Alzheimer's disease were included in this analysis. Items that *only* occurred in one or other of the patient groups were excluded from the analysis to avoid statistical overfitting (in visual construction, spatial errors occurred only in the Alzheimer's disease group and overelaboration occurred only in the FTD group; in repetition, concrete errors occurred only in the FTD group). Items from the VOSP were excluded from this analysis as, because of time constraints at the time of test administration, data were not available for a relatively large number of patients. Although immediate and delayed free and cued recall of a story all reached significance, only immediate and delayed free recall were retained in order to reduce the number of variables in the analysis.

A logistic regression analysis with diagnosis as the dichotomous outcome variable (FTD = 1/Alzheimer's disease = 0), and a classification cut off point of 0.5 was carried out in two stages, the first involving quantitative test scores only and the second with the inclusion of qualitative

**Table 2** Perceptuospatial function

Perceptuospatial performance	Median score		Range		p Value
	FTD	AD	FTD	AD	
<b>No of correct answers</b>					
Line drawing identification (/10)	10	10	5 to 10	9 to 10	0.075
Celebrity identification (/9)	9	9	1 to 9	2 to 9	0.242
Shape detection screening test (/20)	<b>20</b>	<b>19</b>	<b>13 to 20</b>	<b>0 to 20</b>	0.180
Incomplete letters (/20)	<b>19</b>	<b>15</b>	<b>13 to 20</b>	<b>0 to 20</b>	<b>0.003†</b>
Silhouettes (/30)	16	14	5 to 23	2 to 26	0.884
Object decision (/20)	16	15	8 to 20	7 to 20	0.178
Dot counting (/10)	10	9	1 to 10	0 to 10	0.310
Position discrimination (/20)	<b>20</b>	<b>18</b>	<b>11 to 20</b>	<b>0 to 20</b>	<b>0.033†</b>
Number location (/10)	8	7	0 to 10	0 to 10	0.372
Cube analysis (/10)	<b>9</b>	<b>6</b>	<b>1 to 10</b>	<b>0 to 10</b>	<b>0.021†</b>
Visual construction (/4)	<b>2</b>	<b>1</b>	<b>0 to 4</b>	<b>0 to 4</b>	<b>0.002†</b>
<b>Error types</b>	<b>Per cent of subjects</b>				
Spatial errors in drawing	<b>0</b>	<b>66%</b>			<b>0.000†</b>
Perseverative strokes in drawing	<b>19%</b>	<b>4%</b>			<b>0.019*</b>
Omission of detail in drawing	76%	77%			0.904
Poor organisation of drawing	<b>51%</b>	<b>23%</b>			<b>0.008*</b>
Overelaboration in drawing	11%	0			0.099

Bold type indicates that the item significantly changed the odds of one or other diagnosis: \*FTD performance worse than AD; †AD worse than FTD.

Analyses were based on a minimum of 73 Alzheimer patients and 33 FTD patients, except for the VOSP subtests, for which data were absent for larger numbers of patients owing to time constraints at the time of test administration (screening test: FTD 20, AD 63; incomplete letters: FTD 23; AD 68; silhouettes: FTD 20, AD 69; object decision: FTD 19, AD 60; dot counting: FTD 26, AD 72; position discrimination: FTD 14, AD 26; number location: FTD 18; AD 48; cube analysis: FTD 20, AD 48).

AD, Alzheimer's disease; FTD, frontotemporal dementia.

**Table 3** Memory

Memory	Median score		Range		p Value
	FTD	AD	FTD	AD	
<b>No of correct answers</b>					
Digit span forwards	6	6	0 to 6	2 to 9	0.954
Digit span backwards	3	3	0 to 10	0 to 7	0.601
Story: free recall	<b>5</b>	<b>2</b>	<b>0 to 10</b>	<b>0 to 10</b>	<b>0.007†</b>
Story: cued recall	<b>7</b>	<b>3</b>	<b>0 to 10</b>	<b>0 to 10</b>	<b>0.001†</b>
Story: delayed free recall					<b>0.000†</b>
Story: delayed cued recall					<b>0.000†</b>
Story: % retained with delay (free)	<b>100</b>	<b>0</b>	<b>0 to 175%‡</b>	<b>0 to 400%‡</b>	<b>0.000†</b>
Story: % retained with delay (cued)	<b>97</b>	<b>50</b>	<b>0 to 160%‡</b>	<b>0 to 400%‡</b>	<b>0.001†</b>
Celebrity recall (/9)	<b>4</b>	<b>1</b>	<b>0 to 8</b>	<b>0 to 9</b>	<b>0.000†</b>
<b>Error types</b>					
Per cent of subjects					
Intrusions in story recall	<b>6</b>	<b>31</b>			<b>0.012†</b>
Confabulations in story recall	<b>24</b>	<b>8</b>			0.038*
Misconstruction in story recall	<b>50</b>	<b>9</b>			0.000*

Bold type indicates that the item significantly changed the odds of one or other diagnosis: \*FTD performance worse than AD; †AD performance worse than FTD.

Analyses were based on a minimum of 73 Alzheimer's disease patients and 33 FTD patients.

‡Delayed recall is expressed as a percentage of immediate recall; figures are greater than 100% where the amount of information produced on delayed recall exceeds that produced on immediate recall.

AD, Alzheimer's disease; FTD, frontotemporal dementia.

characteristics. The following quantitative variables were entered into the first stage of analysis: sentence repetition, metaphor interpretation, visual construction, recall of celebrities, and free recall of a story (immediate and delayed). The resulting logistic regression model correctly classified 71% of FTD patients and 93% of Alzheimer patients, yielding a sensitivity of 0.71 and a specificity of 0.96.

A composite score for this regression algorithm was created and entered into the next stage of the logistic regression analysis along with the following qualitative variables: phonological errors on the graded naming test; errors of phonology and concreteness in sentence repetition; concrete errors in metaphor interpretation; errors of perseveration and organisation in copying drawings; intrusions, confabulations and misconstructions in story recall; and dismantling on the Weigl's test. The purpose of entering a composite score of the "quantitative" logistic regression equation instead of the individual items was to minimise the number of variables in the analysis. The resulting logistic regression model correctly classified 96% of FTD patients and 93% of Alzheimer patients, yielding a sensitivity of 0.91 and a specificity of 0.96.

### Behavioural observations

Ninety one per cent of Alzheimer patients were noted to behave appropriately during the assessment—that is, they were cooperative and engaged in tests with normal concern for accuracy—compared with only 30% of FTD patients ( $p < 0.0005$ ); 56% of FTD patients showed disinhibited behaviour compared with just 4% of Alzheimer patients ( $p < 0.0005$ ); and 35% of FTD patients were recorded as showing economy of effort when carrying out tasks compared with 3% of Alzheimer patients ( $p < 0.0005$ ).

### DISCUSSION

We predicted that neuropsychological test scores alone would only be partially successful in distinguishing FTD and Alzheimer's disease because test scores mask differences in underlying reason for test failure. In contrast, we predicted that analysis of qualitative aspects of performance would enhance the differentiation between the two groups and highlight clinically distinct patterns of performance.

Some overall test scores, as predicted, did differentiate the groups: FTD patients performed worse on a letter fluency test

**Table 4** Executive function

Executive function	Median score		Range		p Value
	FTD	AD	FTD	AD	
<b>No of correct answers</b>					
Category fluency	10	9	0 to 26	2 to 21	0.762
Letter fluency	<b>4</b>	<b>7</b>	<b>0 to 18</b>	<b>1 to 29</b>	<b>0.016*</b>
Weigl's test /9	4	6	0 to 9	0 to 9	0.157
<b>Error types</b>					
Letter fluency: repetitions	0	0	0 to 3	0 to 4	0.451
Letter fluency: rule violations	0	0	0 to 3	0 to 3	0.096
Category fluency: repetitions	<b>0</b>	<b>0</b>	<b>0 to 4</b>	<b>0 to 9</b>	<b>0.036†</b>
Category fluency: rule violations	0	0	0 to 1	0 to 4	0.785
Per cent of subjects					
Perseveration on Weigl's test	44	29			0.117
Dismantling on Weigl's test	<b>33</b>	<b>15</b>			<b>0.031*</b>

Bold type indicates that the item significantly changed the odds of one or other diagnosis. \*FTD performance worse than AD; †AD performance worse than FTD.

Analyses were based on a minimum of 73 Alzheimer's disease patients and 33 FTD patients.

AD, Alzheimer's disease; FTD, frontotemporal dementia.

and on sentence and metaphor comprehension than Alzheimer patients. By contrast, Alzheimer patients performed worse on drawing tests, on some subtests of a perceptuospatial battery, and on tests of memory. These findings are entirely in keeping with clinical observations that FTD predominantly affects frontal executive skills, whereas in Alzheimer's disease there is prominent involvement of memory and perceptuospatial abilities. They are in keeping too with neuroimaging findings, which indicate that FTD is a disorder of the anterior hemispheres, whereas Alzheimer's disease involves the medial temporal lobes and posterior hemispheres.<sup>6,7</sup> Nevertheless, as expected, despite the evident clinical differences between the two groups, performance on executive and perceptuospatial tasks did not consistently differentiate the groups. Category fluency and performance on the Weigl's test did not differentiate the groups (although the latter did reach significance when patients with MMSE scores of 15 or under were excluded from the analysis). Similarly, comparable performance was elicited on some of the perceptual and spatial tests, including the simple dot counting task. Comparable performance occurred despite the fact that the Alzheimer patients showed no behavioural features to suggest frontal lobe disorder, and the FTD patients had no history of difficulty negotiating or becoming lost in the environment to suggest spatial disorder.

Our analysis of performance characteristics and error types showed that even when test scores did not differentiate the two groups, there are qualitative differences in performance, supporting the notion that failure arises for different underlying reasons. The groups were not differentiated by their total scores on a standard naming test, yet certain error types were strongly predictive of one or other diagnosis: perseverative errors increased the probability of FTD, whereas phonological errors were predictive of Alzheimer's disease. Overall repetition scores did not differentiate the two groups but error types did: the presence of concrete responses increased the odds of a FTD diagnosis, and phonological errors were predictive of an Alzheimer's disease diagnosis. Overall scores on the Weigl's test did not differentiate the groups, yet the FTD patients were more likely to show "dismantling" behaviour. That is, they failed to identify the sorting rule when blocks were grouped by the examiner, but dismantled the blocks, rearranging them according to a previous criterion or their own idiosyncratic sorting method. Such behaviour indicates an inability to shift mental set. In category fluency, repetition of early items—suggesting an inability to remember items previously produced—was a significant predictor of Alzheimer's disease.

A logistic regression analysis indicated that both numerical scores and error types were of value in predicting the diagnosis of individual cases. When test scores only were entered into the analysis, the resultant logistic regression model correctly classified 87% of patients. However, the accuracy of this classification was not equal between the two groups. Test scores correctly classified 93% of Alzheimer patients, but just 71% of FTD patients. When error types were added to the analysis, the classification accuracy remained unchanged for Alzheimer's disease, but increased to 96% for FTD. This suggests that although test scores are of considerable statistical value in predicting diagnosis, the inclusion of error types in addition to test scores greatly enhances the differentiation of the two conditions.

Across the range of tasks, features that are typically associated with frontal lobe failure were predictive of FTD: concreteness of thought, perseveration, confabulation and misconstruction, and poor organisation. Concrete responses occurred across various tasks, including interpretation of metaphor and proverb, a task often used to measure the capacity for abstract thought. Concrete responses also occurred

on tasks that ostensibly may not be thought to require abstract thought: interpretation of complex sentences and a simple task of repetition. These findings indicate that concrete and over-literal responses are a salient characteristic of FTD. Perseveration on naming and drawing tests significantly increased the odds of an FTD diagnosis. On the Weigl's test of abstraction and set shifting, the "dismantling" form of perseveration previously described increased the odds of an FTD diagnosis. Confabulations and misconstructions were features of FTD patients' memory performance. Of note was also that the amount of information retained over a delay varied considerably between the two groups: FTD patients retained a median of 100% of their immediate recall compared to 0% in the Alzheimer group. These qualitative features strongly suggest a "frontal" form of memory loss in FTD, contrasting with an amnesia secondary to medial temporal lobe pathology in Alzheimer's disease. In FTD, drawing errors most commonly arose because of organisational failure. Frankly spatial errors were notably absent in FTD.

The relatively high rate of phonological errors observed in the Alzheimer group deserves comment, as it is a generally held view that phonological aspects of language are well preserved in Alzheimer's disease. Phonological errors on the graded naming test significantly increased the probability of an Alzheimer's disease diagnosis. However, when analysis was restricted to those patients scoring 15 or more on the MMSE this did not reach significance, suggesting that this finding reflected the inclusion of individuals with more severe cognitive impairment. Nevertheless, on a test of word and sentence repetition, 43% of Alzheimer patients made phonological errors, compared with just 9% of FTD patients. This feature remained highly predictive of Alzheimer's disease when analysis was restricted to those scoring over 15 on the MMSE. Thus repetition appears to be a particularly sensitive task in eliciting phonological impairment. Our findings are consistent with other published reports. Alzheimer patients have been reported to make no more phonological errors than healthy controls in spontaneous speech.<sup>26</sup> By contrast, on repetition tasks Alzheimer patients are more prone to committing phonological errors,<sup>27</sup> the rate of which appears to be related to the phonological and syntactic complexity of the material used.<sup>28,29</sup>

Our data enable us to resolve discrepancies in published reports. Previous neuropsychological studies have not always shown significant differences between FTD and Alzheimer's disease. Furthermore, there has been inconsistency regarding the neuropsychological measures reported to differentiate the two conditions. We have shown that neuropsychological test scores can actually obscure differences between patient groups because they mask the underlying reasons for test failure. In FTD there is profound behavioural disturbance, characterised by impulsive and disinhibited behaviour, perseveration, impersistence, inattention, and economy of effort. These behaviours affect test performance in a way that is not apparent, and is actually disguised when numerical scores alone are considered.

FTD is characterised behaviourally by a "frontal lobe" syndrome. By contrast, patients with Alzheimer's disease are usually socially appropriate and do not generally display behaviours suggestive of frontal lobe dysfunction. In the present study, the performance of Alzheimer patients was not coloured by perseveration, concrete thought, confabulation, or economy of effort, and their behaviour was generally appropriate during testing. Nevertheless, there were few quantitative differences between FTD and Alzheimer patients on executive tasks. The finding that Alzheimer patients perform similarly to FTD patients in terms of their overall scores, despite lacking a "frontal lobe" syndrome, suggests that they are failing these tasks for non-frontal reasons. Executive tasks are complex and make demands on multiple

cognitive abilities including language, working memory, and often spatial abilities. Weigl's test requires the ability to perceive an array of blocks and evaluate any differences between them. To carry out this task efficiently requires intact visual perception and spatial abilities, functions that are compromised in Alzheimer's disease. Similarly, the impairment of short term memory characteristic of Alzheimer's disease may result in a patient sorting blocks according to a previous sorting rule because of failure to keep track, and not because of a "frontal" form of executive dysfunction.

Some methodological issues merit discussion. The patients involved in the study were consecutive referrals to a specialist dementia clinic, providing a representative sample of each patient group, covering a range of illness severities. Consequently, patients were not matched for age, sex, or severity of illness. It has been customary when conducting research with individuals who have degenerative neurological disorders to match patient groups for dementia severity according to their scores on the MMSE. However, we did not believe this was appropriate because this bedside test is heavily weighted towards memory impairment, with 16 of the 30 points allocated to memory items. Given the prominent memory disorder associated with Alzheimer's disease, matching across the two patient groups for MMSE scores in an attempt to yield groups of equivalent overall severity could not be meaningfully achieved. Indeed, predictably, scores on the MMSE were significantly lower in Alzheimer's disease than in FTD. Moreover, a notional "severity of dementia" is most unlikely to account for our findings, as the FTD group was more impaired on some measures, whereas the Alzheimer group was more impaired on others. The duration of symptoms was broadly equivalent in each group, so in this regard the groups were well matched.

Our study strongly indicates that reliance on numerical scores alone does not reliably differentiate FTD and Alzheimer's disease. FTD is associated with a profound behavioural disorder that affects performance on all cognitive tests, obscuring qualitative group differences. When performance characteristics and error types are properly taken into account, the accuracy of differentiation is enhanced. Knowledge of the performance characteristics and error types that characterise FTD ought to be of significant value to clinicians and may assist in the differential diagnosis of this condition and in the evaluation of individual patients. The study attests to the value of obtaining qualitative information in general neuropsychological research. Test scores must be viewed and interpreted in the context of both cognitive and behavioural performance to avoid misconceptions that can arise from a purely quantitative assessment. The latter generates "scores", which can serve to obscure rather than highlight significant qualitative differences between patients with different forms of dementia. These characteristic distinctions are clearly manifest in the clinical history of cognitive and behavioural symptoms and the findings of the neurological examination, corresponding to the two distinct clinical neuropsychological profiles of FTD and Alzheimer's disease.

## ACKNOWLEDGEMENTS

We thank Andy Vail, lecturer in the Biostatistics Group, University of Manchester, for statistical advice.

## Authors' affiliations

J C Thompson, C L Stopford, J S Snowden, D Neary, Cerebral Function Unit, Greater Manchester Neuroscience Centre, Hope Hospital, Salford, UK

Competing interests: none declared

## REFERENCES

- 1 **The Lund and Manchester Groups.** Consensus Statement. Clinical and neuropathological criteria for fronto-temporal dementia. *J Neural Neurosurg Psychiatry* 1994;**4**:416-18.
- 2 **Neary D, Snowden JS, Gustafson L, et al.** Frontotemporal lobar degeneration. A consensus on clinical diagnostic criteria. *Neurology* 1998;**51**:1546-54.
- 3 **McKhann G, Drachman D, Folstein M, et al.** 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* 1984;**34**:939-44.
- 4 **Neary D, Snowden JS, Northern B, et al.** Dementia of frontal lobe type. *J Neural Neurosurg Psychiatry* 1988;**11**:592-8.
- 5 **Cummings JL, Benson DF.** *Dementia. A clinical approach.* Boston: Butterworth Heinemann, 1992.
- 6 **Talbot PR, Snowden JS, Lloyd JJ, et al.** The contribution of single photon emission tomography to the clinical differentiation of degenerative cortical brain disease. *J Neural* 1995;**242**:579-86.
- 7 **Talbot PR, Lloyd JJ, Snowden JS, et al.** A clinical role for <sup>99m</sup>Tc-HMPAO SPECT in the investigation of dementia? *J Neural Neurosurg Psychiatry* 1998;**64**:306-13.
- 8 **Efgrén C, Brun A, Gustafson L, et al.** Neuropsychological tests as discriminators between dementia of Alzheimer type and frontotemporal dementia. *Int J Geriatr Psychiatry* 1994;**9**:635-42.
- 9 **Mendez MF, Cherrier M, Perryman KM, et al.** Frontotemporal dementia versus Alzheimer's disease: differential cognitive features. *Neurology* 1996;**47**:1189-94.
- 10 **Hodges JR, Patterson K, Ward R, et al.** The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: a comparative neuropsychological study. *Neuropsychology* 1999;**13**:31-40.
- 11 **Razani J, Brauer-Boone K, Miller B, et al.** Neuropsychological performance of right and left frontotemporal dementia compared to Alzheimer's disease. *J Int Neuropsychol Soc* 2001;**7**:468-80.
- 12 **Rascovsky K, Salmon DP, Ho GJ, et al.** Cognitive profiles differ in autopsy confirmed frontotemporal dementia and Alzheimer's disease. *Neurology* 2002;**58**:1801-8.
- 13 **Pachana NA, Brauer-Boone K, Miller B, et al.** Comparison of neuropsychological functioning in Alzheimer's disease and frontotemporal dementia. *J Int Neuropsychol Soc* 1996;**2**:505-10.
- 14 **Lindau M, Almkvist O, Johansson SE, et al.** Cognitive and behavioural differentiation of frontal lobe degeneration of the non-Alzheimer type and Alzheimer's disease. *Dement Geriatr Cogn Disord* 1998;**9**:205-13.
- 15 **Perry RJ, Hodges JR.** Differentiating frontal and temporal variant frontotemporal dementia from Alzheimer's disease. *Neurology* 2000;**54**:2277-84.
- 16 **Mathuranath PS, Nestor PJ, Berrios GE, et al.** A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology* 2000;**55**:1613-20.
- 17 **Grossi D, Fragassi NA, Chiacchio L, et al.** Do visuospatial and constructional disturbances differentiate frontal variant of frontotemporal dementia and Alzheimer's disease? An experimental study of a clinical belief. *Int J Geriatr Psychiatry* 2002;**17**:641-8.
- 18 **Gregory CA, Orrell M, Sahakian B, et al.** Can frontotemporal dementia and Alzheimer's disease be differentiated using a brief battery of tests? *Int J Geriatr Psychiatry* 1997;**12**:375-83.
- 19 **Kumar A, Schapiro MB, Haxby JV, et al.** Cerebral metabolic and cognitive studies in dementia with frontal lobe behavioural features. *J Psychiatr Res* 1990;**24**:97-109.
- 20 **Warrington EK, James M.** *Visual object and space perception battery.* Bury St. Edmunds: Thames Valley Test Company, 1991.
- 21 **Lezak M.** *Neuropsychological assessment.* Oxford: Oxford University Press, 1995.
- 22 **Bathgate D, Snowden JS, Varma A, et al.** Behaviour in frontotemporal dementia, Alzheimer's disease and vascular dementia. *Acta Neurol Scand* 2001;**103**:367-78.
- 23 **Neary D, Snowden JS, Bowen DM, et al.** Neuropsychological syndromes in presenile dementia due to cerebral atrophy. *J Neural Neurosurg Psychiatry* 1986;**49**:1630-74.
- 24 **Snowden JS, Mann DM, Neary D.** Distinct neuropsychological characteristics in Creutzfeldt-Jakob disease. *J Neural Neurosurg Psychiatry* 2002;**73**:686-94.
- 25 **Neary D, Snowden JS, Shields RA, et al.** Single photon emission tomography using <sup>99m</sup>Tc-HM-PAO in the investigation of dementia. *J Neural Neurosurg Psychiatry* 1987;**50**:1101-9.
- 26 **Blanken G, Dittman J, Haas JC, et al.** Spontaneous speech in senile dementia and aphasia: implications for a neurolinguistic model of language production. *Cognition* 1987;**27**:247-74.
- 27 **Glosser G, Kohn, SE, Friedman RB, et al.** Repetition of single words and nonwords in Alzheimer's disease. *Cortex* 1997;**33**:653-66.
- 28 **Biassou N, Grossman M, Onishi K, et al.** Phonologic processing deficits in Alzheimer's disease. *Neurology* 1995;**45**:2165-9.
- 29 **Glosser G, Friedman RB, Kohn SE, et al.** Cognitive mechanisms for producing nonwords: evidence from Alzheimer's disease. *Brain Language* 1998;**63**:32-49.