Standardised measurement of self-awareness deficits in FTD and AD
Craig Williamson, Oscar Alcantar, Johannes Rothlind, et al.

J Neurol Neurosurg Psychiatry 2010 81: 140-145 originally published online
February 9, 2009
doi: 10.1136/jnnp.2008.166041

Updated information and services can be found at:
http://jnnp.bmj.com/content/81/2/140.full.html

These include:

Data Supplement
"Lay Summary"
http://jnnp.bmj.com/content/suppl/2010/01/28/jnnp.2008.166041.DC1.html

References
This article cites 38 articles, 15 of which can be accessed free at:
http://jnnp.bmj.com/content/81/2/140.full.html#ref-list-1

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://journals.bmj.com/cgi/ep
Standardised measurement of self-awareness deficits in FTD and AD

Craig Williamson, 2 Oscar Alcantar, 1,2 Johannes Rothlind, 3 Deborah Cahn-Weiner, 1 Bruce L Miller, 1,2 Howard J Rosen 1,2

ABSTRACT
Background Diminished ability to perceive one’s own impairments, whether cognitive or social, is common in dementia, in particular frontotemporal dementia (FTD), where ‘lack of insight’ is listed as a core diagnostic feature. Yet, there is no currently accepted method for measuring insight in dementia. The most commonly used methods, which involve comparing patients’ opinions of their level of impairment with the opinions of care givers or close family members, are subjective and require the participation of a knowledgeable informant. Here, the authors introduce a new method that allows objective quantification of an individual’s awareness of their cognitive abilities and relies upon objective bedside testing.

Methods The authors administered several tests of everyday, real-world functions to patients with FTD (n=10) and Alzheimer’s disease (AD, n=10) and to control subjects (n=10). Prior to the tasks, participants were asked to predict their performance using a percentile-based rating system. They were also asked to estimate their performance after task completion. Differences between their self-rated and actual performances were calculated.

Results Whereas the control group showed very little discrepancy between pretest predictions, post-task estimates and actual performance (mean difference of 3.9 percentile points for prediction/3.0 percentile points for post-task estimate), both patient groups overpredicted and overestimated their performance, with a significantly greater discrepancy for FTD (49.0/54.3 percentile points) than AD (27.2/28.3 percentile points).

Discussion Failures of insight and self-awareness of cognitive dysfunction can be objectively measured in dementia without the assistance of an informant, which will facilitate further study of this key component of higher cognitive functioning.

INTRODUCTION
Awareness of our own capabilities is a critical component of normal cognition that gives us the ability to recognise our limits and choose our activities accordingly. This function is commonly compromised in neurological diseases, including dementia, where impaired awareness of one’s deficits makes an important contribution to functional decline. 1 In AD, lack of awareness of one’s difficulties is common and increases in severity as the disease advances. 2-4 In frontotemporal dementia (FTD), impaired self-awareness occurs early in the illness and was included as one of five core diagnostic features in the Neary criteria. 5 Impaired awareness of deficits in dementia has been shown to influence patients’ willingness to seek and comply with treatment and to increase care giver burden. 6,7

While impaired self-awareness is common with many neurological diseases, a thorough understanding of this phenomenon has been hampered by the absence of objective and reliable methods for measuring insight. A variety of methods have been used, 1 but the most common has been to administer questionnaires to patients asking them about their current abilities and to compare their answers with those of knowledgeable informants (family members, hospital staff, etc). 10-19 This method is less than ideal for several reasons. Most importantly, it relies on the opinion of an informant, whose assessment can be influenced by factors including how well they know the individual, how distressing they find the behaviours and their own cognitive abilities. 20

A less commonly used approach has been to ask patients to rate their abilities and to compare these ratings with their performance on neuropsychological tests. 21-24 This approach has taken two general forms. One method is to obtain patient ratings on their general abilities in several domains (memory, language, etc) and compare the ratings with scores on representative tasks chosen. 21-23

Another technique has been to describe a task and have patients predict their performance prior to administration, or rate their performance on a task they have just completed. 24 As they have been implemented, these approaches remain problematic. When patients’ ratings of their abilities are compared with test scores, it is not always clear how questionnaire responses should be scaled to correspond with the scores. 20 Conversely, if patients are asked to predict or estimate actual performance on individual tests, they may have difficulty rating themselves because the test has little relationship with everyday tasks. For instance, tasks such as verbal fluency tap into frontal lobe, language and generative abilities, but it may be very difficult for patients to predict, a priori, how well they would do on such a task. In contrast, the use of tasks with better ecological validity may allow people to predict their performance by extrapolating from their daily experiences.

In the present study, we attempted to develop an objective and reliable method for evaluating cognitive self-awareness, suitable for administration to patients with dementia, which would not depend on an informant. We chose several tasks from the Neuropsychological Assessment Battery...
(NAB\cite{25}), which uses age, sex and education normed tasks to assess domains such as memory and visuospatial function in ways that are similar to people's everyday experience. In what is probably the most novel aspect of the task, instead of asking patients to estimate their actual score, we asked them to rate themselves on a percentile scale, allowing a direct comparison of predicted/estimated with actual performance using the same scale. In administering this task to FTD and Alzheimer's disease (AD) patients, and controls, we reasoned that discrepancies should be larger for patients than controls and largest for FTD because the clinical criteria for FTD stress the importance of poor insight.

**METHODS**

**Subjects**

We recruited 10 patients with probable AD and 10 patients with the behavioural variant of FTD from the Memory and Aging Center at the University of California, San Francisco. All patients underwent a standardised assessment that included a history and physical exam, care giver interview and neuropsychological testing using a standard protocol\cite{26}. Patients were diagnosed using published criteria.\cite{27} Ten normal controls served as a comparison group. Controls primarily consisted of volunteers from the community who had a normal neurological exam, clinical dementia rating scale (CDR) equal to zero, mini-mental status examination (MMSE) greater than or equal to 28 and normal performance of neuropsychological testing, or were members of the University of California, San Francisco Memory and Aging Center staff (n=3). The study was approved by the local IRB.

**Experimental procedure**

After informed consent, subjects were administered tasks of everyday function from the NAB\cite{25}. The NAB includes five modules that assess memory, language and calculations, judgement, spatial ability and attentional/executive functions. Each module contains a task in which the subject is asked to solve a real-world problem. Specifically, the memory module tests the ability to recall medication instructions and a name and address; the language and calculations module requires one to interpret an electric bill, write a check and balance a chequebook; the spatial module requires subjects to navigate from place to place using a map; the attention/executive module contains a task in which the subject is asked to solve a judgement task requiring subjects to explain the reasoning behind certain types of behaviour (eg, 'why is it important to brush your teeth?').

At the beginning of the session, subjects were informed that they would perform a series of tasks that would be described to them, and that they would be asked to predict their performance relative to a hypothetical sample of people their age, sex and education, based on the description of the task. They were shown a picture of a bell curve with corresponding percentile rankings at the bottom of the page (figure 1). They were reminded that on a typical task, the majority of healthy age-matched peers would perform at the 50th percentile, with smaller numbers performing above or below average (corresponding locations were pointed to by the experimenter). They were told that, after hearing about each task, they would have to predict how they would do by pointing to where they would be on the bell curve picture.

The session then commenced. Prior to each task, the test was described to the subject in concrete terms that were very similar to the instructions dictated in the NAB manual. The task was explicitly described with specific references to the type of stimuli that would be presented (eg, memorising a name and address), with examples as appropriate. The bell curve picture was then produced, and the subject was asked to predict how they would perform in terms of a percentile rank. Subjects were instructed to rate themselves based on how they thought they performed on similar tasks in daily life. Then, the task instructions were reread as outlined in the NAB procedure manual, and the task administered. Once the task was completed, the bell curve picture was produced again, and the subject was asked to estimate how they had just performed, again in terms of a percentile rank, compared with healthy age-matched peers.

To ensure that subjects were using the percentile rankings in a way that was generally consistent with their true opinions about their performance, we preceded each percentile-based prediction with a question asking them to tell us how they thought they would do on the task by saying average, below average or above average. These estimates were converted into numbers (1—below average, 2—average, 3—above average).

**Analysis**

Raw scores were converted into percentiles using the procedures outlined in the NAB manual. Pretest and post-test performance estimates were subtracted from actual percentile scores. This calculation yields a negative discrepancy for participants who overestimate their performance and a positive discrepancy for those who underestimate. Our primary objective was to compare average discrepancies across domains, rather than attempt to discern whether there were any differences in self-awareness between domains. Pretest and post-test discrepancies for each task were averaged within subjects, resulting in a mean discrepancy score for each participant. These discrepancy scores were compared across diagnostic groups using one-way analysis of variance (ANOVA). Tukey post-hoc tests were used to identify significant differences between individual groups.

Theoretically, if all subjects rated themselves at the same level, self-awareness score would be totally determined by performance, and if self-appraisal were conservative, with patients rating themselves to be at a level achieved by the average person, those with the poorest performance would have the worst self-awareness discrepancy score. Indeed, the two patient groups would be expected to show a greater error in self-appraisal just because they are actually impaired. To investigate whether our hypothesised group differences were due to abnormal distortions in self-appraisal rather than group differences in performance, an analysis of covariance was performed with actual performance as a covariate.

To assess whether group differences in distortions of self-awareness were comparable across tasks, the discrepancy scores for each of the five performance domains were also examined and compared across groups.

Lastly, to gauge the validity of self-appraisal distortions in dementia as indicators of reduced insight, we examined the correlation between the NAB self-appraisal discrepancy and discrepancy between patient and care giver (or other knowledgeable informant) ratings of patient's problems in everyday functioning. At the end of the experiment, we asked patients a series of nine questions about the changes in level of everyday difficulty in the areas of memory, judgement, financial management, attention and navigation (eg, 'Compared with 5–10 years ago, are you having any problems with your memory—for example remembering recent events or conversations?'). They responded on a four-point scale (0—no problems, 1—mild problems,
RESULTS

Subject characteristics

The FTD, AD and control groups did not significantly differ with respect to age, sex or years of education (table 1). AD subjects performed most poorly on the MMSE, while FTD subjects’ performance was between AD and controls.

Neuropsychological assessment battery prediction and performance

AD patients made the lowest predictions of their performance, while FTD subjects predicted that their performance would be slightly higher than controls (figure 2). The same pattern held for post-test self-assessment. Differences in pretest prediction were not statistically significant across groups, while post-test performance predictions approached statistical significance (p=0.057). The increased difference between groups in post-test estimation was due to the fact that AD patients downgraded their estimates slightly after the task, whereas controls and FTD patients did not.

Both AD and FTD subjects performed significantly worse than normal controls on all NAB modules, with the exception of navigation (table 2). On average, FTD subjects scored lower than AD subjects on all modules except memory, but there were no statistically significant differences between patient groups in any of the domains or in the average across the NAB tasks.

Discrepancy scores

The average pretest and post-test discrepancies were significantly greater in FTD than in AD and controls (figures 3, 4, left side). The mean pretest discrepancies were −49.0 (±25.5) in FTD, −27.2 (±18.1) in AD and 3.9 (±16.5) in controls. Post-test discrepancies were −54.3 (±17.9) in FTD, −28.3 (±15.5) in AD, and −3.0 (±15.5) in controls. Differences between groups for pre- and post-test discrepancies were significant for FTD versus controls, AD versus controls and FTD vs AD. Scatter plots of individual discrepancy scores within each group (figures 3, 4, right side) revealed no major outliers.

ANCOVA with post-test discrepancy as the dependent variable, diagnosis as the independent variable and mean NAB percentile score as a covariate was significant ($R^2=0.758$) with significant effects of diagnosis (p=0.004) and NAB score (p=0.002). Post-hoc Bonferroni comparisons revealed that
discrepancies were significantly greater in FTD compared with AD (p=0.006) and controls (p=0.022), but the difference between AD and controls was no longer significant. This result, combined with the finding that overall NAB performance was not significantly different across patient groups, suggests that self-appraisal was less determined by actual performance in FTD than it was in AD.

The greater discrepancy in FTD than in AD was a general finding across tasks, and there did not appear to be any major differences across domains (table 3).

**Relationship between percentile-based ratings and other estimates**

In order to ensure that patients’ percentile rankings were truly reflective of how they felt about their performance, we examined the correlation between these percentile rankings and the verbal ratings (average, below average, above average) given just prior to each percentile estimate. The percentile rankings were strongly and significantly correlated with these verbal ratings (r=0.914 for pretest predictions, r=0.885 for post-test estimates), indicating that the percentile scores were being used in a manner consistent with the way patients saw their own performance.

We also examined correlations of the discrepancies between the estimated and actual percentile rankings, and the discrepancy between patient and informant estimation of patients’ daily functioning. In the 13 patients with self and informant ratings of daily function, these were significantly correlated with percentile-based discrepancies (r=0.645 for pretest predictions, r=0.556 for post-test estimations), indicating that those patients who were the worst at estimating their performance on cognitive testing also showed the least awareness of their difficulties in daily life.

**DISCUSSION**

These data indicate that the method described here, which uses a percentile-based ranking system and objective neuropsychological tasks tapping everyday function, is a feasible approach for measuring self-awareness in dementia and yields patterns of self-awareness across diagnostic groups consistent with established clinical characterisations. Specifically, both FTD and AD patients showed impaired self-awareness of cognitive deficits relative to controls. Moreover, self-awareness measured using this percentile-based approach is significantly correlated with verbal estimates of one’s abilities, and with awareness of one’s impairments in daily functioning. While self-awareness deficits were associated with a lower performance in AD, the impaired performance of the patients with FTD accounted less for their abnormally inflated self appraisal. Even after adjusting for actual performance, FTD patients were significantly less aware of their deficits than AD patients. Comparison across domains suggested greater distortions in self-appraisal in FTD patients on a task on which they performed slightly better than the AD patients (Memory), and also on tasks on which they performed less well. FTD patients displayed more severe distortions of self-appraisal than AD regardless of question format and evaluation context (pre- vs post-testing). AD patients appeared somewhat more sensitive to their performance than FTD, as indicated by the fact that they were the only group to downgrade their estimate after performing the tasks (although not to a large degree).

These findings would be expected, as many prior studies have found reduced self-awareness in AD, and FTD is characterised by early and prominent loss of insight. Our results are generally consistent with previous results but differ in some ways. Several studies have examined awareness of deficits in both FTD and AD patients by comparing patient and informant assessments. Eslinger et al examined several domains of cognition and behaviour, and found that, compared with normal controls, social-behavioural-type FTD patients were significantly impaired in 10 of the 17 domains assessed, while AD patients were significantly impaired in only two of 17. Rankin et al used personality measures to demonstrate that FTD patients had more impaired insight into their personality changes than AD patients. Salmon et al showed a similar trend that was not statistically significant. Banks and Weintraub recently demonstrated impaired insight into behaviour using the Frontal Behavioural Inventory (FBI) in both AD and FTD compared with controls, but showed no differences across the these patient groups, but their result may have differed from others’ because they compared the informants’ FBI scores with the patients’ overall assessment of behaviour, rather than comparing responses to the same questions. Thus, overall, studies using patient vs informant discrepancy approaches have supported the clinical characterisation of FTD as having more impaired self-awareness than AD.

Efforts to measure insight without clinician or informant bias have yielded less consistent findings. Eslinger et al used this approach, asking patients to assess their performance on standard neuropsychological tasks such as verbal fluency (eg, asking

**Table 2** Mean performance on individual Neuropsychological Assessment Battery subtests

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Alzheimer’s disease</th>
<th>Frontotemporal dementia</th>
<th>Normal controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>8.3 ± 13.4</td>
<td>14.5 ± 25.9</td>
<td>54.6 ± 13.6</td>
</tr>
<tr>
<td>Bill paying</td>
<td>20.4 ± 28.3</td>
<td>7.9 ± 17.8</td>
<td>62.5 ± 7.4</td>
</tr>
<tr>
<td>Judgement</td>
<td>56.4 ± 35.0</td>
<td>28.5 ± 33.2</td>
<td>89.1 ± 24.3</td>
</tr>
<tr>
<td>Navigation</td>
<td>35.3 ± 31.4</td>
<td>25.6 ± 26.1</td>
<td>59.1 ± 36.4†</td>
</tr>
<tr>
<td>Driving</td>
<td>13.9 ± 26.6</td>
<td>12.8 ± 24.7</td>
<td>71.6 ± 27.6</td>
</tr>
</tbody>
</table>

Percentile = SD. *p<0.001; †p=0.068 by ANOVA.

Figure 3 Left-hand side: mean and standard error of the discrepancies between pretest performance prediction and actual Neuropsychological Assessment Battery (NAB) performance on a percentile scale. Right-hand side: scatter plots of these discrepancies in each group.
patients to predict how they will do based on a scale going from ‘no words’ to ‘a lot of words’). They found poor correlations between pretest ratings and performance even in controls, consistent with the idea that it is difficult for people to make predictions about tests that have relatively little resemblance to tasks performed in daily life. However, post-test ratings were significantly correlated with performance in controls and AD, while FTD patients’ post-test ratings were correlated with performance on two of three tasks. O’Keefe et al. failed to find any significant difference between FTD and AD on most assessments of insight, although patients with FTD showed less awareness of their errors on a Stroop task. While indicating poor monitoring of performance, this result does not speak directly to their level of overall self-awareness. Lastly, Banks and Weintraub examined self-awareness for memory and naming in controls, FTD and AD by asking patients to answer questions such as ‘how good is your memory?’ on an analogue scale ranging from ‘no ability’ through ‘average’ to ‘perfect.’ They found both FTD and AD to be impaired at self-awareness for memory, with no difference between FTD and AD. Thus, studies attempting to measure insight without the aid of an informant have not consistently revealed differences in insight between FTD and AD, which is at odds with the clinical characterisation of FTD.

We attribute the fact that these prior studies were only marginally able to demonstrate more impaired insight in FTD than AD to methodological factors that we tried to address with the current study. First, our use of tasks simulating real-world situations allowed patients to use their experience in everyday function as a guide to rating their performance. The fact that controls were highly accurate at predicting their performance supports the validity of this approach. Also, the use of a percentile-based scoring system may have advantages over other approaches and may increase measurement accuracy by forcing the examiner and study participant to use a common metric.

A potential concern for our measure is that it reflects overall estimation ability more than awareness of personal performance deficits per se. We were reassured by the fact that percentile-based discrepancy scores were highly correlated with non-numerical estimates of task performance, and also correlated with informant-based estimates of patients’ awareness of their overall functioning in daily life. Although we did not include a control task, Banks and Weintraub found that dementia patients were not impaired in their ability to estimate their weight and eyesight, suggesting that impaired self-awareness in dementia is due to a specific deficit in estimating cognitive impairments. Also, previous studies have highlighted the loss of awareness of social functions in FTD. This was not an issue addressed in this study, because there are few established tests of social functioning that are normed in older individuals. Once developed, such tasks should be adaptable for assessment of self-awareness in social domains in a manner similar to that used here.

It is also notable that our study found that dementia patients had deficits in predicting performance as well as in estimating recent performance—tasks that may, in part, depend on differing cognitive processes. Prediction likely requires the ability to form a mental representation of the upcoming task and may require episodic memory spanning weeks to months for recent examples of similar tasks. In contrast, post-task estimation likely takes advantage of cognitive monitoring abilities, probably dependent on the frontal lobes as well as relatively short-term memory lasting a few minutes. Given the differences between these tasks, the relative stability of patients’ self-appraisals are quite remarkable. The need for certain cognitive abilities for both tasks, including working memory (ie, for task instructions and the specific question recently asked), may be one explanation. Another may be that impaired cognitive monitoring abilities may affect post-task estimation in our sessions, and may also impede recognition of cognitive difficulties in daily life, so that task prediction in our sessions is influenced by cognitive monitoring failures occurring over the last few months. Also, it should not be assumed that deficits in self-awareness have the same cognitive basis in both AD and FTD. This is illustrated by the fact that our AD group showed a slight downgrading of performance estimation after task completion, suggesting that monitoring abilities are less impaired in AD than FTD. Future research will need to be directed at uncovering the cognitive underpinnings of self-awareness in dementia patients and should take into account potential differences in aetiology depending on diagnosis.

A more detailed study of self-awareness of cognitive and functional impairment in dementia is warranted for many reasons. First, absence of self-awareness of deficits likely contributes to functional impairment and care giver burden by impeding role adjustment and adherence to medically indicated.

Table 3 Mean pretest and post-test discrepancy scores on individual neuropsychological assessment battery subtests

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Alzheimer’s disease</th>
<th>Frontotemporal dementia</th>
<th>Normal controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory pretest</td>
<td>–37.1 ± 17.9</td>
<td>–42.1 ± 40.5</td>
<td>0.7 ± 24.8***</td>
</tr>
<tr>
<td>Memory post-test</td>
<td>–26.1 ± 17.3*</td>
<td>–54.3 ± 24.4</td>
<td>–16.7 ± 23.0**</td>
</tr>
<tr>
<td>Bill paying pretest</td>
<td>–42.3 ± 26.3</td>
<td>–50.8 ± 29.9</td>
<td>0.2 ± 32.6**</td>
</tr>
<tr>
<td>Bill paying post-test</td>
<td>–38.6 ± 18.0**</td>
<td>–70.7 ± 22.9</td>
<td>–11.7 ± 23.3**</td>
</tr>
<tr>
<td>Judgement pretest</td>
<td>–1.5 ± 34.0*</td>
<td>–40.7 ± 37.9</td>
<td>16.3 ± 22.2**</td>
</tr>
<tr>
<td>Judgement post-test</td>
<td>–10.3 ± 27.6*</td>
<td>–43.5 ± 34.4</td>
<td>18.8 ± 24.3***</td>
</tr>
<tr>
<td>Navigation pretest</td>
<td>–24.5 ± 33.6</td>
<td>–51.0 ± 30.0</td>
<td>–7.2 ± 34.1*</td>
</tr>
<tr>
<td>Navigation post-test</td>
<td>–24.4 ± 23.3</td>
<td>–48.7 ± 27.1</td>
<td>–8.8 ± 26.7*</td>
</tr>
<tr>
<td>Driving pretest</td>
<td>–30.7 ± 21.7</td>
<td>–55.6 ± 28.1</td>
<td>9.5 ± 26.1***</td>
</tr>
<tr>
<td>Driving post-test</td>
<td>–42.2 ± 15.8</td>
<td>–58.2 ± 27.7</td>
<td>3.6 ± 24.5***</td>
</tr>
</tbody>
</table>
treatment and lifestyle changes. Measurement of self-awareness may also allow stratification of patients with regard to injury risk and permit tailoring of behavioural therapies. Furthermore, accurate quantitative measurement is necessary to further study the cognitive and neural mechanisms mediating self-awareness. Though many studies have indicated an important role for the frontal lobes, particularly on the right, little is known about the specific frontal functions involved. The ability to accurately and objectively measure deficits in self-awareness of cognitive impairment in dementia will enhance the ability to study these issues in detail.

Funding This work was supported by the State of California DHS Alzheimer’s Disease Research Center of California (ARCC) grant 06-55318, NIH grants P1DA019724, P50-AG23501, grant no M01 RR00079 (UCSF General Clinical Research Center) and the Hillbom Network.

Competing interests None.

Ethics approval Ethics approval was provided by the University of California, San Francisco Committee on Human Research.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES