Chapter 16

The Dutch Adult Reading Test (DART): A measure of (premorbid) intelligence?

M. de Lange, R. W. H. M. Ponsé, N. Razendaal, P. J. Huus, and J. Jolles

ABSTRACT

Reading tests have been proposed as a simple measure for obtaining an estimation of premorbid IQ, especially for use in the clinical diagnosis of brain injury and dementia in elderly patients. The Dutch version of the National Adult Reading Test was investigated in the MAAS-A panel with 359 subjects aged 25–80 years. Intra-rater and inter-rater reliability were studied. The results show that the Dutch Adult Reading Test (DART) has good reliability but systematically underestimates IQ compared to that measured with a standard intelligence test. This underestimation was particularly prominent when the DART raw scores were low. The data show that caution is needed when the DART is used for clinical purposes.

INTRODUCTION

A problem in clinical practice is that it is difficult to obtain estimates of premorbid IQ in patients with dementia. This is of importance in clinical neuropsychology because a difference between estimated premorbid intelligence and measured intelligence is one of the criteria for the diagnosis of dementia according to DSM-IV (American Psychiatric Association, 1994). Besides, norms for (neuro)psychological tests are often related to intelligence scores.

Geber & Stiwinski (1997) describe several methods for estimating premorbid IQ such as the use of demographic data (education and occupation), as well as methods that focus on overlearned skills, such as reading performance. These skills are highly correlated with intelligence and are relatively robust in the absence of focal brain damage. Nelson & O'Connell (1978) developed a test to measure premorbid IQ based on reading skills, the National Adult Reading Test (NART). The NART is a word-reading test; it requires the reading of 50 irregular words which
cannot be correctly pronounced by applying the usual rules that map spelling into sound. The NART gives a good indication of the vocabulary of a patient and is for this reason a good estimator of the verbal premorbid IQ.

The Dutch Adult Reading Test (DART; Schmand, Lindeboom and Harshamp, 1992) is the Dutch version of the NART. The main findings of the NART were replicated with the DART (Schmand, Bakker, Saan & Louman, 1991). A high correlation (r=.85) with verbal intelligence was found in healthy controls (n=22) and the insensitivity of the DART to cerebral deterioration was proven in a group of brain damaged and demented patients (n=53). However, when the DART was used in a neuropsychological assessment battery in standard clinical practice in the Maastricht Memory Clinic (MMC, see Chapter 17), it was found that the IQ obtained with the Groninger Intelligence Test (Luteijn & van der Ploeg, 1982) was much higher than the IQ obtained with the DART. This made the validity of the DART questionable. Therefore, the DART was investigated in MAAS, in order to see whether it is a good predictor of (premorbid) IQ in a healthy population. This study was completed with the DART data from a group of 110 patients of the MMC. Three persons independently scored the DART in order to obtain information on inter-rater reliability.

METHOD

Subjects

In total 469 subjects participated in the first panel study A1 of MAAS. DART scores were obtained from 359 subjects. The mean age in this group was 52.4 years (range 24-85 years); men made up 54% of the group. All subjects were screened for several characteristics before they were invited for neuropsychological assessment (see Section 3.4 for MAAS exclusion criteria). The patient population consisted of 110 patients who attended the MMC for neuropsychiatric and neuropsychological assessment. The average age in this group was 53.7 years (range 26-83 years); men made up 63% of the group.

Test administration

All subjects completed the DART and GIT (see Section 4.3). The audiotaped responses of the subjects on the DART were scored independently afterwards by three neuropsychologists (MdL, RP, PH). For the scoring, the pronunciation rules as proposed by Schmand et al. (1992) were used. These rules were adapted for words that are very sensitive to the Limburg dialect. These words were: circulaire, accessoire, funiculaire, turquoise, and clairvoyance.
RESULTS

Inter- and intra-rater reliability

Inter-rater reliability, expressed as Kappa, was high (> .75). In Table 16.1, Kappa scores of clusters of five words for each group of subjects are shown. The intra-rater reliability was also very satisfactory. The Pearson correlation coefficient between two ratings of the DART of 32 subjects from the MAAS–A₁ panel ranged from .98 to .99 for the three raters.

Comparison of GIT-IQ and DART-IQ

The mean DART-IQ of the patient group and the MAAS group was 99 and 97, respectively. These DART-IQ scores were below the IQ scores obtained by the GIT. The patient group had a GIT-IQ of 110 whereas the GIT-IQ of the MAAS group was 113. Further analysis showed that this underestimation was greater for low IQ scores of the DART. GIT-IQ scores in the higher range were better predicted by the DART IQ score than lower GIT-IQ scores were. In the healthy population (MAAS group), there was a steady decline in the index score (DART raw score divided by GIT-IQ) with lower DART scores (see Figure 16.1). The underestimation of GIT-IQ was about 30% for DART raw scores above 80 and rose to 70% for DART raw scores below 51.

<table>
<thead>
<tr>
<th>Word no.</th>
<th>DART</th>
<th>Patients MMC</th>
<th>Subjects MAAS–A₁</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5</td>
<td>.61</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td>5-10</td>
<td>.83</td>
<td>.75</td>
<td></td>
</tr>
<tr>
<td>11-15</td>
<td>.84</td>
<td>.78</td>
<td></td>
</tr>
<tr>
<td>16-20</td>
<td>.63</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>21-25</td>
<td>.76</td>
<td>.74</td>
<td></td>
</tr>
<tr>
<td>26-30</td>
<td>.79</td>
<td>.80</td>
<td></td>
</tr>
<tr>
<td>31-35</td>
<td>.72</td>
<td>.72</td>
<td></td>
</tr>
<tr>
<td>36-40</td>
<td>.86</td>
<td>.82</td>
<td></td>
</tr>
<tr>
<td>41-45</td>
<td>.71</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>46-50</td>
<td>.79</td>
<td>.84</td>
<td></td>
</tr>
</tbody>
</table>

DISCUSSION

The results of this study show that the DART as proposed by Schmand et al. (1991) is not reliable as a measure for premorbid IQ. Although the IQ obtained for this group should be about the same level as the
Fig. 16.1. Index score DART–IQ divided by GIT–IQ (+SD). At the horizontal axis 10 classes of DART IQ are shown: approximately equal numbers of subjects in each IQ class according to the cumulative frequency distribution. Class 1 = (DART IQ) 65–81; 2 = 82–87; 3 = 88–91; 4 = 92–95; 5 = 96–99; 6 = 100–102; 7 = 103–105; 8 = 106–107; 9 = 108–

Premorbid IQ assessed by the DART; this study shows that the DART-IQ systematically underestimated IQ when compared with the actual IQ score in a healthy control group. Even more striking is the fact that the underestimation increased when the IQ score decreased. Thus the usefulness of the DART as a measure of premorbid IQ is open to discussion. These results suggest that the DART should be evaluated extensively with regard to its reliability and validity as well as its usefulness for measuring premorbid IQ in a demented patient population.

REFERENCES


