Subscale Validity of the Mattis Dementia Rating Scale


Department of Neurology and Alzheimer’s Disease Center, University of Alabama at Birmingham

The Mattis Dementia Rating Scale (DRS), a widely used dementia screening instrument, generates five subscale scores in the areas of Attention (ATT), Initiation-Perseveration (IP), Construction (CN), Conceptualization (CON), and Memory (MEM). The present study sought to determine the criterion validity of the DRS subscales in a sample of 50 patients with Alzheimer’s disease (25 with mild and 25 with moderate dementia). Subject performance on the five DRS subscales was correlated with performance on five well-validated neuropsychological criterion measures using Pearson r and stepwise regression. On a univariate level, each DRS subscale correlated most strongly with its assigned neuropsychological criterion measure. On a multivariate level, each DRS subscale emerged as the Step 1 predictor of its assigned criterion measure, with the exception of DRS CN, which was the Step 2 predictor. The results suggested that overall the DRS subscales are valid measures of their respective constructs and have value for both clinical and research purposes in mild and moderate dementia. © 1997 National Academy of Neuropsychology

Address correspondence to: Daniel C. Marson, Department of Neurology, JT1216, University of Alabama at Birmingham, Birmingham, AL 35294-0007.

This research was supported by an Alzheimer’s Disease Program Project grant (NIH, NIA 5 P01 AG06569-05), and an Alzheimer Disease Center Cores grant (NIH, NIA 1 P30 AG10163-1).
proven valuable for the screening and staging of dementia (Coblentz et al., 1973; Moss & Alberts, 1988; Salmon, Thal, Butters, & Heindel, 1990; Shay et al., 1991).

Whereas the DRS Total Score appears to be a valid and reliable measure of overall dementia level, no investigations have focused on the DRS subscales and their validity. Factor analytic studies have suggested that the DRS test items measure several distinct cognitive abilities (Colantonio, Becker, & Huff, 1993; Kessler, Roth, Kaplan, & Goode, 1992). However, no study to date has specifically examined whether the DRS subscales validly measure their respective constructs. Establishing the criterion validity of the DRS subscales would significantly enhance both the clinical and research value of the DRS.

In this study we sought to determine the criterion validity of the DRS subscales in a sample of patients diagnosed with Alzheimer’s disease (AD). Five well-validated neuropsychological measures representative of the cognitive domains of the five DRS subscales were selected as criterion measures. We then correlated DRS subscale scores with criterion measure scores to determine the criterion validity of the DRS subscales.

**METHODS**

**Subjects**

Subjects were 50 patients with possible or probable AD recruited from a Program Project in Alzheimer’s disease (NIA 5P01 AG06569-05) and from the UAB Alzheimer’s Disease Center Core (NIA 1P30 AG10163-1). Subjects ranged in age from 54 to 89 (M = 71.4, SD = 8.1) and in education from 9 to 16 years (M = 12.2, SD = 2.1). Patients were well characterized medically, neuropsychologically, and neuroradiologically. Diagnoses of possible or probable AD were made by consensus judgment of a neurologist, psychiatrist, and neuropsychologist, using NINCDS-ADRDA criteria (McKhann et al., 1984). The mean Mini Mental State Examination score (MMSE) (Folstein, Folstein, & McHugh, 1975) for the group was 18.4 (SD = 5.0). Using the Clinical Dementia Rating Scale (CDR) (Hughes, Berg, Danziger, Coben, & Martin, 1982), all subjects were classified as either mildly demented (CDR = 1.0) (n = 25) or moderately demented (CDR = 2.0) (n = 25). Informed consent was obtained from all subjects and their caregivers as part of this IRB approved research.

**Neuropsychologic Criterion Measures**

As part of a longitudinal study, subjects were annually administered the DRS and a neuropsychologic test battery designed for dementia assessment. For each DRS subscale, we selected from our battery a well-standardized neuropsychological test previously demonstrated to be a valid measure of the respective DRS construct. These criterion measures were:

1. **DRS ATT: WMS-R Attention-Concentration.** The Attention/Concentration Index of the Wechsler Memory Scale-Revised (WMS-R Attention) (Wechsler, 1987) was selected as the criterion measure for the DRS ATT subscale. WMS-R Attention is widely accepted as a valid measure of overall attentional capacity (Lezak, 1995). The index consists of tasks of digit span forward and backward, visual memory span forward and backward, and of timed recitation of overlearned material. The DRS ATT subscale consists of untimed tasks of digit span forward and backward, motor response to single commands, visual scanning, word list reading, and visual figure matching.

2. **DRS IP: Controlled Oral Word Association.** The Controlled Oral Word Association test (CFL) (Benton & Hamsher, 1989) was selected as the criterion measure for the DRS IP subscale. The CFL task, which involves time limited generation of words to three different phonemic letter cues, has been acknowledged to be a valid measure of verbal
fluency and initiation (Lezak, 1995). The DRS IP subscale consists of tasks of time
limited word generation to two different semantic categories (supermarket items;
pieces of examiner’s clothing), oral repetition of phonemically related syllables,
bilateral motor planning, and copying repetitive geometric patterns. CFL was chosen as
the criterion for DRS IP since a semantic fluency task like Animals was not available
in this particular dementia battery.

3. DRS CN: WAIS Block Design. The Block Design subtest of the Wechsler Adult
Intelligence Scale (WAIS) (Wechsler, 1955) was selected as the criterion measure for
the DRS CN subscale. WAIS Block Design is widely accepted as a valid measure of
visuospatial construction (Lezak, 1995). The task involves the manual manipulation of
colored blocks to replicate a series of pictured designs. The DRS CN subscale involves
the copying of five geometric figures of varying difficulty, and the writing of one’s
name. Although Block Design differs in significant respects from DRS CN, in partic-
ular in its manipulative use of three-dimensional stimuli in a timed format, it was the
most appropriate standardized measure available from the test battery for visuospatial
construction.

4. DRS CON: WAIS Similarities. The Similarities subtest of the WAIS (Wechsler, 1955)
was selected as the criterion measure for the DRS CON subscale. WAIS Similarities is
acknowledged to be a valid measure of abstract verbal concept formation and asso-
ciative thinking (Lezak, 1995). This task involves the verbal identification of similarities between pairs of verbally presented objects and concepts. The DRS CON subscale
consists of tasks of identification of similarities between pairs of objects, recognition of similarities between objects (multiple-choice format), primed inductive reasoning,
identification of nonmembers of semantic categories, identification of similarities and
differences among simple geometric figures, and generation of a simple sentence using
specified words.

5. DRS MEM: WMS-R Verbal Memory. The Verbal Memory Index of the WMS-R
(WMS-R Verbal Memory) (Wechsler, 1987) was selected as the criterion measure for
the DRS MEM subscale. WMS-R Verbal Memory has been shown to be a valid measure of
short-term verbal memory function (Butters et al., 1988; Lezak, 1995). The index
consists of tasks of short-term narrative learning/recall and of verbal paired-associate
learning. The DRS MEM subscale consists of tasks of short-term recall of two
sentences, recognition memory for words and geometric designs, and orientation for
time (day, date, month, year), place (hospital, city) and current events (president,
governor, mayor). We felt that WMS-R Verbal Memory, and not WMS-R General
Memory, was the better criterion match for DRS MEM, as DRS MEM loads very
heavily on short term verbal memory and only addresses spatial memory briefly in a
recognition format.

We wish to emphasize that the neuropsychologic tests selected represented the most
appropriate criterion measures available from our dementia battery. Given the diverse item
content of the DRS subscales, it would have been interesting and useful to have had two or
more criterion measures for each DRS subscale. However, as discussed above, the size of our
dementia battery limited us in most cases to only one matched criterion measure. As
expected, matched criterion measure tasks were similar but not identical to the DRS subscale
tasks, and were generally more difficult than the DRS tasks.

Statistical Analyses

All data was analyzed using the raw scores of each test, with the exception of WAIS and
WMS-R variables. Age corrected scaled scores were used for WAIS Block Design and
TABLE 1

Means, SDs, and Score Ranges of DRS Subscales and Criterion Variables for the AD Sample (N = 50)

<table>
<thead>
<tr>
<th>DRS Subscale</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
<th>Matched Criterion</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATT</td>
<td>32.2</td>
<td>4.1</td>
<td>17-37</td>
<td>WMS-R Attention</td>
<td>69.5</td>
<td>21.0</td>
<td>50-115</td>
</tr>
<tr>
<td>IP</td>
<td>25.0</td>
<td>8.0</td>
<td>2-37</td>
<td>CFL</td>
<td>24.0</td>
<td>10.5</td>
<td>5-48</td>
</tr>
<tr>
<td>CN</td>
<td>4.5</td>
<td>1.9</td>
<td>0-6</td>
<td>WAIS Block Design</td>
<td>5.0</td>
<td>4.3</td>
<td>0-14</td>
</tr>
<tr>
<td>CON</td>
<td>29.0</td>
<td>6.8</td>
<td>13-39</td>
<td>WAIS Similarities</td>
<td>8.3</td>
<td>2.3</td>
<td>0-12</td>
</tr>
<tr>
<td>MEM</td>
<td>11.4</td>
<td>4.9</td>
<td>0-24</td>
<td>WMS-R Verbal Memory</td>
<td>60.7</td>
<td>11.5</td>
<td>50-82</td>
</tr>
</tbody>
</table>

Similarities, and Index scores were used for WMS-R Attention/Concentration and Verbal Memory.

The five DRS subscales were first correlated with the selected criterion variables using Pearson r (SAS PEARSON) (SAS Institute, 1985). Stepwise multiple regression analyses were then performed using the DRS subscales as predictors for each of the dependent criterion measures (SAS STEPWISE) (SAS Institute, 1985). We accepted the Step 1 and Step 2 multivariate solutions as the identified predictors. A p value of .05 and a subject to predictor variable ratio of 10:1 were used for the regression analyses.

RESULTS

Table 1 shows the means, standard deviations, and score ranges of the DRS subscales and criterion measures for our AD patient sample.

Table 2 shows the univariate correlations between the DRS subscales and criterion variables for our AD patient sample. Each DRS subscale correlated significantly and most strongly with its assigned criterion variable. Moderately strong correlations occasionally also emerged between DRS subscales and non-assigned criterion variables.

Table 3 shows the results of the stepwise multiple regression analyses for our AD patient sample. Each DRS subscale emerged as the Step 1 predictor of its assigned criterion measure, with the exception of DRS CN, which was the Step 2 predictor of its criterion measure Block Design.

DISCUSSION

The increasing prevalence of dementia in our aging population (Evans et al., 1989), as well as the changing health care environment (Patterson & Sharfstein, 1992), heighten the need for brief, informative, and inexpensive measures of cognitive function in the elderly. While the search for new and better geriatric assessment instruments continues, another response is to enhance the sensitivity and validity of existing dementia instruments. The DRS is a well-accepted, valid, reliable, and relatively brief measure of overall cognitive function in dementia. It is also currently one of the few existing dementia instruments that provides subscale information for discrete cognitive functions. Although the DRS subscales and their constituent test items appear to be face valid, and in clinical practice have been assumed to be valid, this property has not until now been empirically investigated.

The results of this study suggest that the DRS subscales are valid measures of their respective constructs within a mild and moderately impaired AD patient population. Univariate correlation analyses demonstrated that each DRS subscale correlated significantly and most strongly with its assigned criterion variable (Table 2). Multivariate analyses further supported the criterion validity of four of the five DRS subscales: DRS ATT was the Step 1
Subscale Validity of the DRS

TABLE 2
Univariate Correlation Matrix: DRS Subscales and Criterion Variables*

<table>
<thead>
<tr>
<th>Criterion Variables</th>
<th>ATT</th>
<th>IP</th>
<th>CN</th>
<th>CON</th>
<th>MEM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WMS-R Attention</strong></td>
<td>0.70</td>
<td>0.39</td>
<td>0.53</td>
<td>0.26</td>
<td>0.43</td>
</tr>
<tr>
<td><strong>CFL</strong></td>
<td>0.56</td>
<td>0.63</td>
<td>0.39</td>
<td>0.47</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>WAIS Block Design</strong></td>
<td>0.62</td>
<td>0.59</td>
<td>0.56</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td><strong>WAIS Similarities</strong></td>
<td>0.30</td>
<td>0.53</td>
<td>0.17</td>
<td>0.56</td>
<td>0.41</td>
</tr>
<tr>
<td><strong>WMS-R Verbal Memory</strong></td>
<td>0.43</td>
<td>0.57</td>
<td>0.20</td>
<td>0.57</td>
<td>0.69</td>
</tr>
</tbody>
</table>

*Correlations >0.55 are significant at p < 0.001.

predictor of WMS-R Attention, as was DRS IP for CFL, DRS CON for WAIS Similarities, and DRS MEM for WMS-R Verbal Memory (Table 3). Although the r and R²'s were of moderate strength, we felt that stronger correlations with the assigned criterion variables would have been hard to obtain given the diverse item content of many of the DRS subscales and the disparity in difficulty level between the DRS subscales and the assigned criterion variables. Accordingly, we feel that overall the results strongly supported the criterion validity of the DRS subscales.

Only DRS CN failed to correlate most strongly with its criterion Block Design on a multivariate level. DRS ATT was the Step 1 predictor of WAIS Block Design, with DRS CN emerging as the Step 2 predictor (Table 3). There are several possible explanations for this finding. First, DRS CN and Block Design are different construction tasks, and probably measure somewhat different cognitive processes (Benton, 1985). DRS CN is a two-dimensional drawing task involving simple graphomotor praxis, while Block Design is a three-dimensional block assembly task involving more complex visuomotor and visuconstructive functions. Block Design is also a timed task with points assigned for speed, whereas DRS CN is untimed. Dissociation between drawing tasks and block assembly tasks has frequently been noted clinically and has also been demonstrated empirically (Dee, 1970). Accordingly, Block Design may not be a fully satisfactory criterion match for DRS CN.

Secondly, given the complexity and timed format of Block Design, it is understandable that attentional measures might emerge as very strong correlates of performance. DRS ATT, with its broad score range (0–37 points) and specific spatial attention items (11 points), may be more sensitive than DRS CN (range 0–6) to disturbances in performance on Block Design.

Finally, the failure of DRS CN to be the strongest correlate of Block Design probably also reflects psychometric limitations of both measures. As noted above, DRS CN has a very limited score range, while Block Design had a restricted range of scores and attenuated variance in our sample due to the difficulty of this task for an AD population (Table 1). Thus, the limited number of items in DRS CN combined with floor effects on Block Design likely decreased their covariance.

Accordingly, while the present results provide limited support for DRS CN as a valid measure of spatial construction, they also indicate the need for further investigation with different criterion variables.

The criterion validity of the DRS subscales has important clinical and research implications. Clinically, the DRS subscales can serve as brief and useful supplements to existing specific measures of cognitive function. In appropriate circumstances, the DRS subscales may also be substituted for more comprehensive and lengthy evaluations. The clinical utility of the DRS subscales, however, currently remains limited due to the relative absence of well
TABLE 3
Stepwise Multiple Regression Analysis: DRS Subscales* and Criterion Variables

<table>
<thead>
<tr>
<th>Criterion Variables</th>
<th>Step 1 Predictor</th>
<th>Partial $R^2$</th>
<th>$p$</th>
<th>Step 2 Predictor</th>
<th>Partial $R^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WMS-R Attention</td>
<td>DRS ATT</td>
<td>(0.49)</td>
<td>0.0001</td>
<td>DRS CN</td>
<td>(0.08)</td>
<td>0.04</td>
</tr>
<tr>
<td>CFL</td>
<td>DRS IP</td>
<td>(0.40)</td>
<td>0.0001</td>
<td>DRS ATT</td>
<td>(0.08)</td>
<td>0.01</td>
</tr>
<tr>
<td>WAIS Block Design</td>
<td>DRS ATT</td>
<td>(0.38)</td>
<td>0.0001</td>
<td>DRS CN</td>
<td>(0.08)</td>
<td>0.01</td>
</tr>
<tr>
<td>WAIS Similarities</td>
<td>DRS CON</td>
<td>(0.32)</td>
<td>0.0001</td>
<td>DRS IP</td>
<td>(0.08)</td>
<td>0.02</td>
</tr>
<tr>
<td>WMS-R Verbal Memory</td>
<td>DRS MEM</td>
<td>(0.47)</td>
<td>0.0001</td>
<td>DRS CON</td>
<td>(0.07)</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*Bolded DRS subscales reflect original matches with criterion values.

standardized subscale performance norms for controls and different diagnostic groups (Schmidt et al., 1994; Shay et al., 1991).

From a research standpoint, validation of the DRS subscales offers a number of advantages. First, investigators will be able to track specific cognitive changes in dementia longitudinally with greater confidence. Secondly, use of the DRS subscales in correlational and related studies can offer a clearer understanding of the contribution of specific cognitive changes to functional and behavioral change in dementia. Finally, as discussed above, the DRS subscales also offer the benefit of brevity of administration, an advantage when doing neuropsychological research with a demented population.

Several limitations of our study should be noted, as well as directions for further research. First, our sample size was relatively small and did not include severely impaired patients (CDR = 3.0) or patients with very early dementia (CDR = 0.5). Establishing the validity and sensitivity of the DRS subscales in very early AD would have great diagnostic value. Secondly, we had available only a limited selection of criterion variables from our dementia battery, and in the case of DRS CN a completely satisfactory criterion match was unavailable. The findings of the present study should be confirmed by other studies using additional, and where possible, multiple criterion variables. Finally, the value of the DRS subscales would be increased by establishing their criterion validity in other neurological and psychiatric populations.

In summary, the present investigation provides initial support for the validity of the DRS subscales as clinical and research measures in mild and moderate dementia.

REFERENCES


Subscale Validity of the DRS 275


