Stratified Norms for the Rivermead Behavioural Memory Test

H.G.G. van Balen¹,², P.S.H. Westzaan¹, and T. Mulder¹,³

¹Department of Research and Development, Sint Maartenskliniek, Nijmegen, The Netherlands
²Nijmegen Centre for Motor Behaviour and Cognition, Sint Maartenskliniek, Nijmegen, The Netherlands
³Neurological Institute, University of Nijmegen, The Netherlands

The purpose of this study is to investigate the value of stratified normative data for the Rivermead Behavioural Memory Test (RBMT). The RBMT has a demonstrated capacity to predict everyday memory problems and is recognised as a useful and ecologically relevant clinical tool. As the measurement of rate of change will often be the principal objective in neuropsychological rehabilitation, preferably in comparison to the group or the functional situation aimed at, the availability of stratified norms should enhance the adequate interpretation of test performance. To investigate this, 214 healthy, elderly individuals and 680 patients participated in this multicentre study using a clinical trial approach. Significant differences for test scores were expected for different groups according to age, aetiology, health care services, and some combined variables, for example, coma duration in traumatically brain-injured patients. Group effects in the expected directions were found for RBMT performance according to all stratification variables. Some implications and limitations of these results are described. Because the results clearly show the existence of homogeneous subgroups, taking stratified norms into account may improve the measurement of rate of change as well as decision making in clinical neuropsychological rehabilitation.

Requests for reprints should be sent to H.G.G. van Balen, Department of Research and Development, Sint Maartenskliniek, P.O. Box 9011, 6500 GM Nijmegen, The Netherlands, Tel.: + 31 24 3659157, Fax.: + 31 24 3659204.

This study would not have been possible without the co-operation of many colleagues working in health care services in The Netherlands and Belgium, nor without the support of Dr M.F.H.G. Wimmers, who unfortunately died in 1993. The authors also wish to thank Bart van der Feen for his contribution in the early stages of the project, and Patay Anderson, Hans Knoop, and Luciano Fasotti for their remarks on the final stages of this text. The study was made possible by the financial assistance of the Stichting Fonds Johannastichting.

© 1996 Psychology Press, an imprint of Erlbaum (UK) Taylor & Francis Ltd
INTRODUCTION

Memory deficits are among the most common cognitive sequelae of brain damage and may hinder everyday functioning to a considerable degree. Therefore, valid tests for memory function which provide a good estimate of memory problems in daily life activities are a useful clinical tool.

The Rivermead Behavioural Memory Test (RBMT; Wilson, Cockburn, & Baddeley, 1985) was chosen as the material for the present study. The test was developed for use in clinical practice to assess everyday memory problems and to complement traditional memory assessment procedures (Wilson, Cockburn, Baddeley, & Hiorns, 1989). It is one of the major memory tests likely to be utilised in most psychology departments (Mayes & Warburg, 1992), and it differs from the bulk of published memory tests in its attempt to sample memory behaviours characteristic of everyday life. Deelman (1990), therefore, considers the RBMT as a positive exception compared to other ecological memory assessment tools. According to Wade (1993) and Hodges (1994), the RBMT is indeed a distinguished instrument for the evaluation of memory abilities, and Wilson (1991) reports that the RBMT is one of the few tests that has a demonstrated capacity to predict everyday memory problems. Because of its attractiveness and apparent clinical relevance, the test was translated into Dutch (Van Balen & Groot Zwaartink, 1987). An additional validation study utilised the test results of 40 stroke patients admitted to a rehabilitation centre (Van der Feen, Van Balen, & Eling, 1990). The observations of everyday memory problems, as carried out by the patient, the patient's partner, and the rehabilitation staff were correlated with the RBMT Screening Score. The results indicated a significant correlation \((r = 0.75)\). Furthermore, the distribution of the screening score was almost similar to that in the validity and reliability study of the original version of the RBMT (Van Balen & Van der Feen, 1988). As these results confirmed the English data of Wilson et al. (1989), they warranted the study of stratified norms for the reasons given below.

Important components for adequate test interpretation in clinical neuropsychological assessment include the level of test scores, patterns of test performance, observation of test completion, and comments made by the patient. However, the fact that the level of test scores often receives more attention than the other factors can be questioned. Indeed, there are at least two weaknesses with respect to the understanding of test scores.

First, test manuals often lack an explanation of how the test data may be interpreted with respect to the different purposes for which patients are referred to clinicians. It is, for example, meaningful to know whether the test only quantifies the deficit, or whether it can also be used to measure the rate of change in that deficit. A related topic is the availability of
stratified and population-based norms and how these norms should be used. It may, for example, be desirable to describe the level of everyday memory performance of a brain-injured patient, as measured with the RBMT, as “well above average” compared with a disease-related norm sample but as “far below average” in comparison with general population norms.

A second complication is that, for most tests, data concerning a systematic evaluation of the role of demographic and other potentially influential variables on test performance, such as brain damage, are unavailable (Randolph et al., 1994). Furthermore, few measures utilised in neuropsychology have been standardised on large samples. This is not only true for the RBMT, but also for such a well-known cognitive test as the Trail Making Test, which has been used for the past 30 years (Ruff & Crouch, 1991).

In conclusion, given the important contribution of test scores in neuropsychological evaluation, there is an urgent need for appropriate normative data for cognitive tests. The purpose of the present study is to discuss the need for stratified norms in clinical neuropsychological assessment and to provide (stratified) normative data1 for the Dutch version of the RBMT, in order to improve the decisions made in clinical neuropsychological rehabilitation.

According to Lezak (1995), population norms are most useful in measurement of deficit for functions that develop in the course of childhood but which are not closely tied to either education or general intellectual ability. These species-wide capacities do not vary much in intact people. Typically, they are represented by a rectangular or J-shaped frequency distribution in the general population. Everyday memory might be considered as such a species-wide capacity because it is generally accepted as being a relatively stable skill during adulthood and well within the capacity of all normal adults. In other words, everyday memory is not normally distributed in the adult population. Since across the age range 16–70 years, the distribution of RBMT scores is J-shaped (Wilson et al., 1989), the RBMT may be regarded as a measurement tool for this species-wide capacity.

If the purpose of a test is measurement of deficit, the results may be compared with norms of a normal population. However, the frame of

---

reference clinicians most often use is not limited to the general population. Assessment of brain-damaged patients to plan rehabilitation programmes and for return to daily life, school, and work, should provide information on the level at which memory functions. In such evaluations, the principal objective may be the measurement of rate of change, preferably in comparison to the group or the functional situation aimed at. According to Wilson (1987), the RBMT was developed also to monitor such a rate of change in the functioning of everyday memory. Measurement of rate of change may be accomplished by comparing the patient’s performance with norms of the normal population. However, rate of change measurement in comparison to the group or the functional situation aimed at is only possible if stratified norms are available. Obviously, if the referral question is rate of change with respect to an individual recovery process, repetitive test administrations are necessary. With respect to the RBMT, this is facilitated by the availability of four parallel versions of the test. In this context, some considerations with respect to the use of stratified norms which are specific for age, aetiology, and health care services are given in relation to memory assessment, specifically for the use of the RBMT.

Age stratification. As indicated by Cockburn and Smith (1989), the English results for people aged 70 and over show a normal-shaped frequency distribution of test scores. This departure from a species-wide capacity warranted the development of the Dutch stratified normative data according to age.

Stratification according to aetiology-specific characteristics. Norms based on aetiology-specific characteristics, such as coma duration in traumatically brain-injured patients, are needed to interpret the level of test performance in order to help form the prognosis. Although coma duration is not the best predictor of memory disturbance after traumatic brain injury (TBI), it is information which is usually available in most cases, whereas a better predictor, such as post-traumatic amnesia, often is not. Nevertheless, several studies report a significant relation between coma duration and memory deficits (Brooks, 1984). It is, therefore, hypothesised that groups of brain-injured patients based on different coma duration will differ in the frequency distribution of RBMT scores. As a result, the use of stratified norms based on coma duration for the same standardised RBMT scores of two patients with a different coma period might result in different interpretations, both in terms of prognosis and relative level of everyday memory skills.

Stratification according to combined characteristics. Since brain damage as well as older age are considered to have an impact on test
performance, norms for combined characteristics seem appropriate for certain cerebral diseases, such as stroke. If, for example, particular test performance suggests deterioration of everyday memory skills, the availability of age-related norms, especially for stroke patients older than 70, may justify more elaborate interpretations.

Stratification according to health care services. Delivery of health care services may depend on inclusion or exclusion criteria concerning level of memory performance. Patients with memory deficits require extensive cognitive support and repetitive presentation of information. The capacity of a rehabilitation unit or a nursing home to admit a patient with poor everyday memory performance could depend on the availability of treatment services with regard to the demands of patients already being treated in the unit. Setting-specific norms may help decision making. Within a specialised rehabilitation service, decisions related to such matters as the assignment of patients to memory group therapy for relatively lower or higher achievers may be supported by the use of service-specific norms.

In addition to the considerations above, stratified norms can only provide additional information if subgroups deviate significantly in statistical terms. In order to investigate this prerequisite, we compared the test scores for 18 Dutch stratified norm groups for the RBMT which were already available (Van Balen & Wimmers, 1993). Significant differences in Screening Score and Standard Profile Scores were expected for:

1. Different age groups in normal controls. People aged 70 or older were expected to perform more poorly.

2. Aetiology-specific variables. Stroke patients, TBI patients, and patients with alcohol-related diseases or dementia might differ in their performance on the RBMT. Expectations as to how they would differ were not clear. However, traumatically brain-damaged patients with a longer coma duration were expected to receive lower scores than those with a shorter coma period.

3. Combined characteristics. In stroke patients a combined impact of age and aetiology was expected which would blur the effect of age for patients beyond 70 years of age.

4. Service-specific norm groups. In-patients were expected to perform more poorly than those in a day treatment setting. Additionally, patients admitted to a general hospital were expected to obtain higher scores than those treated in a rehabilitation unit or a psychiatric hospital since the overwhelming majority of brain-damaged patients return home after hospital discharge (Van Balen, Mulder, & Keyser, in press).
METHOD

Subjects

Participants in this study were 214 healthy elderly people and 680 patients who were admitted within a one-year period to 35 health care services in The Netherlands and to six hospitals in the Dutch speaking part of Belgium. RBMT data for the non-patient sample were assigned to three age-related norm groups, one under 60 years of age (range 45–59, Mean = 55.3, SD = 3.2) one aged 60–69 (Mean = 64.9, SD = 2.7), and one over 69 (range 70–95, Mean = 76.2, SD = 4.3).

All subjects in the non-patient sample lived independently and voluntarily followed a course, organised by a foundation for well-being in the elderly, on how memory works. This course is explicitly developed for non brain-damaged people and is not a treatment method for memory deficits. The RBMT was obligatory. Participants with an ascertainable cerebral dysfunction were excluded from the sample.

The multicentre clinical sample included patients admitted to rehabilitation centres (60.3%), psychiatric institutions (22.6%), and general hospitals (12.7%). The remainder (4.3%) stayed in a nursing home or an equivalent form of care. The data of the patient sample were assigned to 15 strata organised according to age, aetiology, kind of health service, and combined variables. Thus, data for some patients are sometimes found in more than one stratified norm group. For example, the data of a 65-year-old stroke patient who attends an out-patient rehabilitation programme will be incorporated in the following norm groups: stroke patients; stroke patients aged 60–69; patients attending rehabilitation services; and patients attending rehabilitation out-patient services. Table 1 presents an overview of the different strata. In this table, the difference between the number of patients in rehabilitation (n = 431) and in inpatient and out-patient rehabilitation (n = 303 and 123, respectively) is explained by five missing values. For major demographic characteristics such as age, sex, level of education, level of employability, and health status for each subgroup see Van Balen and Wimmers (1993).

Materials

*Rivermead Behavioural Memory Test.* The test components of the RBMT can be seen as analogues of everyday memory situations that appear to be troublesome for brain-damaged patients. The 12 components include: remembering a first name; remembering a surname; remembering a hidden belonging; remembering an appointment; picture recognition; remembering a newspaper article (immediate and delayed
<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>SS Mean (SD)</th>
<th>SPS Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly controls (n = 214)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age in years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;60</td>
<td>26</td>
<td>9.5 (1.8)</td>
<td>20.5 (2.8)</td>
</tr>
<tr>
<td>60–69</td>
<td>99</td>
<td>9.4 (2.2)</td>
<td>20.5 (3.6)</td>
</tr>
<tr>
<td>&gt;69</td>
<td>89</td>
<td>8.5 (2.4)</td>
<td>19.0 (4.3)</td>
</tr>
<tr>
<td>Patients (n = 680)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aetiology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td>258</td>
<td>6.4 (3.4)</td>
<td>15.1 (6.2)</td>
</tr>
<tr>
<td>Dementia</td>
<td>32</td>
<td>4.5 (3.3)</td>
<td>11.1 (6.6)</td>
</tr>
<tr>
<td>Traumatic brain injury (TBI)</td>
<td>164</td>
<td>6.8 (3.0)</td>
<td>16.0 (5.5)</td>
</tr>
<tr>
<td>Alcohol-related disorders</td>
<td>77</td>
<td>5.1 (3.9)</td>
<td>12.4 (7.6)</td>
</tr>
<tr>
<td>Health service</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General hospital</td>
<td>87</td>
<td>7.6 (3.8)</td>
<td>16.4 (7.2)</td>
</tr>
<tr>
<td>Psychiatric hospital</td>
<td>154</td>
<td>6.2 (3.4)</td>
<td>14.3 (6.6)</td>
</tr>
<tr>
<td>Rehabilitation</td>
<td>431</td>
<td>6.7 (3.2)</td>
<td>15.6 (5.8)</td>
</tr>
<tr>
<td>Rehabilitation, in-patient</td>
<td>303</td>
<td>6.2 (3.2)</td>
<td>14.8 (6.0)</td>
</tr>
<tr>
<td>Rehabilitation, out-patient</td>
<td>123</td>
<td>7.6 (3.0)</td>
<td>17.3 (5.2)</td>
</tr>
<tr>
<td>Combined variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBI, coma &lt; 7 days</td>
<td>83</td>
<td>7.4 (2.9)</td>
<td>17.1 (5.1)</td>
</tr>
<tr>
<td>TBI, coma 7–28 days</td>
<td>31</td>
<td>7.1 (2.7)</td>
<td>16.8 (4.3)</td>
</tr>
<tr>
<td>TBI, coma &gt; 28 days</td>
<td>37</td>
<td>5.5 (2.9)</td>
<td>13.5 (6.0)</td>
</tr>
<tr>
<td>Stroke, age &lt; 60 years</td>
<td>117</td>
<td>7.0 (3.4)</td>
<td>15.9 (6.1)</td>
</tr>
<tr>
<td>Stroke, age 60–69 years</td>
<td>77</td>
<td>6.3 (3.1)</td>
<td>15.0 (5.5)</td>
</tr>
<tr>
<td>Stroke, age &gt; 69 years</td>
<td>57</td>
<td>5.6 (3.4)</td>
<td>13.4 (6.8)</td>
</tr>
</tbody>
</table>

recall); face recognition; remembering a new route (immediate and delayed); delivering a message; orientation; and date. Each subtest is adjusted to be on a level such that it would be passed by normal subjects but failed by those having everyday memory problems (De Wall, Wilson, & Baddeley, 1994). For each subtest, two scores are produced, a pass/fail screening score, and a standardised profile score with a possible score of 0–2 (0 points = abnormal; 1 point = borderline; 2 points = normal). Thus, each patient's evaluation results in two summarised scores, a Screening Score (SS) ranging from 0–12, and a Standardised Profile Score (SPS) ranging from 0–24. A detailed description of the test development and validation, as well as of the test items, is given in Wilson et al. (1989).
Procedure

Participating health care services were selected on the basis of an overview supplied by the Dutch test distributor of RBMT-equipped psychology departments. Participating patients were referred to these departments for various reasons. The administration of the RBMT was an integral part of the patient's clinical evaluation. Therefore, patient inclusion was defined by local circumstances and policies. The research project did not interfere with regular clinical procedures and only required a minimal time investment by the clinicians. Each subject was tested individually. Medical diagnosis, biographical data and RBMT data were centrally and anonymously collected and analysed.

RESULTS

Full normative data, including sample characteristics, SS quartile scores, SPS decile scores, and percentages of individual item SPS for all norm groups are described in Van Balen and Wimmers (1993; see also Footnote 1). Table 1 presents an overview of the standard deviations and means for the Screening Score (SS) and for the Standardised Profile Score (SPS) for the different norm groups.

Preliminary exploration of the data indicated that the t-test could not be used to compare different norm groups. English RBMT data for 118 control subjects (Wilson et al., 1989) and data from the Dutch sample of healthy elderly controls \((n = 214)\) show a similar J-shaped pattern. Both data sets confirm the lack of a normal distribution. Therefore, non-parametric analyses have been used. Kruskal-Wallis median tests were done on the SS and on the SPS in order to determine overall effects of stratification according to age, aetiology, combined characteristics, and health service specificity. In addition, if significant group differences were found within a stratification variable, repetitive Mann-Whitney \(U\) tests for independent samples were conducted to reveal significant differences in median SS and SPS for independent stratified norm groups. An alpha level of 0.05 was used.

Age-based norm groups

With respect to age effects in the normal population, a Kruskal-Wallis median test revealed significant differences for SS and SPS group medians: \(\chi^2(2, n = 214) = 8.75, p = 0.0126\), and \(\chi^2(2, n = 214) = 9.00, p = 0.0111\), respectively. Mann-Whitney \(U\) tests for independent samples were performed to determine the source of this effect. SS and SPS group medians were significantly higher for individuals younger than 60 or aged
60–69 than for persons aged 70 or older: \([U(26, 99) = 908.0, p = 0.0207]\) and \([U(26, 99) = 855.5, p = 0.0469]\), and \([U(89, 99) = 3329.5, p = 0.0033]\) and \([U(89, 99) = 3406.5, p = 0.0018]\), respectively. SS and SPS group medians did not differ for the two younger age groups.

Aetiology-specific norm groups

First, differences between four aetiological categories (stroke, TBI, dementia, alcohol-related diseases) were investigated. Differences were significant for SS as well as SPS group medians: \([\chi^2(3, n = 531) = 17.15, p = 0.0007]\), and \([\chi^2(3, n = 531) = 18.85, p = 0.0003]\), respectively. Subsequent analysis showed significant differences in overall RBMT performance between stroke patients and traumatically brain-injured patients on the one hand, and patients with alcohol and dementia-related diseases on the other. However, no differences were found between stroke patients and patients with TBI, or between patients with alcohol and dementia-related disorders. See Table 2 for an overview.

Second, with respect to TBI, three groups were formed, based on coma duration: less than 7 days, 7–28 days and more than 28 days. Kruskal-Wallis median tests revealed a significant effect on both SS and SPS: \([\chi^2(2, n = 151) = 11.99, p = 0.0025]\), \([\chi^2(2, n = 151) = 12.31, p = 0.0021]\), respectively. Mann-Whitney \(U\) tests showed significant differences in SS and SPS group medians between patient groups with a coma duration of less than 7 days and 7–28 days on the one hand, and more than 28 days on the other: \([U(83, 37) = 929.0, p = 0.0005]\) and \([U(83, 37) = 943.0, p = 0.0007]\), and \([U(31, 37) = 1968.0, p = 0.0153]\) and \([U(31, 37) = 1948.0, p = 0.0143]\), respectively. The SS and SPS group medians did not differ

<table>
<thead>
<tr>
<th>Aetiology</th>
<th>Mann–Whitney (U)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stroke</td>
<td>Dementia</td>
</tr>
<tr>
<td>SS</td>
<td>2889.5**</td>
</tr>
<tr>
<td>SPS</td>
<td>2793.5**</td>
</tr>
<tr>
<td>Dementia</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>1553.0***</td>
</tr>
<tr>
<td>SPS</td>
<td></td>
</tr>
<tr>
<td>TBI</td>
<td>SS</td>
</tr>
<tr>
<td>SPS</td>
<td></td>
</tr>
</tbody>
</table>

\(*p < 0.05, \text{two-tailed.} \ **p < 0.01, \text{two-tailed.} \ ***p < 0.001, \text{two-tailed.}\)
between the patient groups with a coma duration of less than 7 days and 7–28 days.

Norm groups based on combined characteristics

In order to determine age effects in stroke patients, three stroke patient groups were formed, age up to 60, 60–69, and age 70 and over. Kruskal-Wallis median tests revealed significant SS and SPS group differences: \( \chi^2(2, \ n = 251) = 6.045, \ p = 0.0487 \) and \( \chi^2(2, \ n = 251) = 6.745, \ p = 0.0343 \), respectively. However, additional comparisons showed only significant differences for SS as well as SPS group medians between stroke patients aged younger than 60 and those aged 70 or older: \( U(117, 57) = 2578.5, \ p = 0.0075 \) and \( U(117, 57) = 2625.5, \ p = 0.0114 \), respectively. And although the SPS group median of the stroke patients aged up to 60 and 60–69 tend to differ, \( U(117, 77) = 3918.0, \ p = 0.0623 \), only the SS group median was significantly higher for stroke patients aged younger than 60 \( U(117, 77) = 3872.0, \ p = 0.0485 \). The SS and SPS group medians did not differ for the two older groups.

Service-specific norm groups

First, differences within one setting, i.e. rehabilitation, were investigated. Both, the SS and SPS group medians were significantly lower for in-patients \( (n = 303) \) than for those following a day treatment programme: \( (n = 123) \), \( U(303, 123) = 8073.0, \ p = 0.0003 \), and \( U(303, 123) = 7987.0, \ p = 0.0002 \), respectively. These patient groups included only stroke and TBI patients. In agreement with the above findings concerning aetiology-specific norms, additional analysis showed no significant differences between in-patient stroke and TBI patients \( (n = 147 \ and \ 92, \ respectively) \), or stroke patients and TBI patients in day treatment \( (n = 55 \ and \ 36, \ respectively) \).

Second, three groups were formed, based on where patients were being treated: rehabilitation centres, psychiatric hospitals, or general hospitals. Kruskal-Wallis median tests were done for SS and SPS group medians: \( \chi^2(2, \ n = 651) = 9.89, \ p = 0.0071 \) and \( \chi^2(2, \ n = 651) = 12.57, \ p = 0.0019 \), respectively. SS and SPS group medians were significantly higher for patients assessed in general hospitals than for those treated in rehabilitation centres or psychiatric hospitals: \( U(87, 431) = 14279.0, \ p = 0.0033 \) and \( U(87, 431) = 15327.0, \ p = 0.0389 \) and \( U(87, 154) = 4968.5, \ p = 0.0008 \) and \( U(87, 154) = 5169.0, \ p = 0.0032 \), respectively. Furthermore, the SPS group median was significantly higher for those treated in rehabilitation centres compared with psychiatric hospitals \( U(431, 154) = 28132.5, \ p = 0.0458 \).
DISCUSSION

The results of this study illustrate the importance of stratified norms in the clinical neuropsychological assessment of everyday memory abilities. Group effects were found on RBMT performance according to all stratification variables. Subsequent analysis showed differences in the expected directions. As presumed, healthy controls aged 70 or older perform more poorly than those aged under 70, and patient groups show lower scores than healthy elderly people. Patient group effects were established for different aetiologies, i.e. stroke, dementia, TBI, and alcohol-related disorders. Expectations as to how these groups should differ were not clear, but the results show that TBI and stroke patients (as defined by the sample characteristics; see Footnote 1) perform better than patients with dementia or alcohol-related disorders. Furthermore, TBI patients with a coma duration of more than 29 days perform less well than those with a coma period of less than 29 days, out-patient rehabilitation patients have higher scores than in-patients, and patients admitted to general hospitals have higher scores than those treated at a rehabilitation unit or a psychiatric hospital. However, caution is still necessary.

With respect to age, the results are in accordance with Cockburn and Smith (1989), who concluded in their study on norms for elderly people \((n = 119)\), that for those aged 70 or older a close relationship exists between age and current memory skills. Next, no differences were found between people younger than 60 and those aged 60–69. This is consistent with Wilson et al. (1989), who describe no effect of age on test performance across the range 16–70 years for controls \((n = 118)\). Thus, age-based norm groups of the RBMT for healthy control subjects only seem justified for two groups: those younger than 70 and those 70 or older.

When other sample characteristics (see Footnote 1) are taken into account, several other implications and limitations may be drawn from the findings. First, the differences found in aetiology-specific everyday memory deficits between the stroke and TBI patient groups in comparison to patients with alcohol-related disorders or dementia could be subject to sample biases. In this study, 82 and 84% of the stroke and TBI patients, respectively, were in a rehabilitation setting, whereas 75% of the patients with alcohol-related disorders and 63% of the patients with dementia were in psychiatric institutions or nursing homes (Van Balen & Wimmers, 1993). In neurological rehabilitation settings, many disabled people are referred for neuropsychological assessment, predominantly TBI and stroke patients. However, such policy is not yet common in other health care settings. In general hospitals, for example, those patients who are thought to have memory problems will typically be referred for neuropsychological
evaluation. Therefore, just a small percentage of all brain-damaged patients admitted to a general hospital will be assessed. Thus, the aetiology-specific norm groups in this study may only be representative for the kind of health services which mainly provided the protocols used.

A second reason for caution can be illustrated by the difference found in SPS test scores between rehabilitation and psychiatric hospital patients (Mean = 15.6, SD = 5.8, and Mean = 14.3, SD = 6.6, respectively). Because these results may be explained by the high percentage of brain-damage (31%) and alcohol-related disorders (32%) in the psychiatric sample, it may not be justified to presume a lower mean SPS for psychiatric patients in general. The results do indeed underscore the importance of taking into account the referral policy for neuropsychological assessment within a particular setting when using setting-specific norms. Even so, clinicians (particularly Dutch ones) who are aware of these limitations may find aetiology and setting-specific norms useful when deciding whether to assign a stroke patient in a rehabilitation setting to group therapy for relatively lower or higher achievers, or when considering a referral of a patient with Korsakoff’s syndrome to a specialised psychiatric “Korsakoff unit”. See also Van Balen and Wimmers (1993) for a detailed description of the percentages and aetiology of brain damage in the distinguished norm groups according to health care services.

Obviously, the results and interpretations should not automatically be generalised beyond populations taking the Dutch translation of the test. Procedures for hospitalisation and referring for neuropsychological assessment may be substantially different in other countries, and may result in markedly different group distributions of scores. Thus, even though in this study the distributions of the SS of TBI and stroke patients treated in rehabilitation were similar to those in the validity and reliability study of the English version of the RBMT (Van Balen & Wimmers, 1993), and even though the data may be valid for sub-populations in other countries, this cannot be assumed to be true in all cases. Indeed, it should be acknowledged that the normative data in this study may only be applicable to decisions made in Dutch clinical neuropsychological rehabilitation.

Third, because group effects were established for different aetiologies, one should be aware of the limitations in interpreting test scores in relation to rate of change measurement if norm groups which are unspecified with respect to aetiology are used. The clinical relevance of stratified norms on the other hand can be demonstrated by RBMT data for subgroups of TBI patients. The comparison of an individual in relation to similar patients could facilitate classification, since a TBI patient with a coma history of 40 days and an SS of 8 would be a relatively good scorer, whereas a TBI patient with a coma duration of one day and an SS of 8 may be judged as
scoring only moderately, even though the actual everyday memory performance might be similar for both these patients. In clinical practice, such nuances are important for prognosis. However, it should be realised that some aetiology-based differences in test performance may only be found by further analysis of subtest results. For example, Wilson et al. (1989) showed that the performance of left and right stroke groups \((n = 34 \text{ and } 42, \text{ respectively})\) on the RBMT differed significantly for the memory of names and for delayed story recall. Such analysis is beyond the scope of this study. Nevertheless, the comparison of subtest scores may be helpful, especially if performance is to be compared across different clinical subgroups. In order to compare subtest results, one may want to consult Van Balen and Wimmers (1993), who also describe separate norm groups for left and right stroke patients \((n = 88 \text{ and } 106, \text{ respectively})\). Future research should reveal whether there are profiling differences between groups and if, for example, the different combinations have equivalent potential for subsequent change.

An additional limitation of this study is the lack of sufficient incorporation of the variable “length of time post-onset at the time of testing” (TPO). It is well recognised that the rate of change and the prevalence of circumscribed long-term sequelae after brain damage are partially dependent on the time period passed since onset. In spite of this, there is hardly any test which provides for data addressing to this variable. Nevertheless, TPO was precoded in this study, but only 23% of the patient protocols included a TPO. Most TPO data were provided by rehabilitation centres and concerned TBI and stroke patients. The available TPO data did not significantly correlate with the SPS \((p = 0.730)\). Subsequent pathology-specific calculations for in-patient and out-patient TBI patients again revealed \(p\) values with a significance of 0.906 and 0.859, respectively. At least two reasons may account for these effects. First, TBI and stroke patients treated in rehabilitation settings comprise only a minor portion of all patients with this pathology (Van Balen, Mulder, & Keyser, in press). Second, reference to rehabilitation settings depends more upon the actual (and estimated) level of functioning than on TPO. Although these variables interact, a linear relation is far from expected. Nevertheless, a thorough investigation of the impact of TPO on test scores of patients with different aetiologies of brain damage and at different moments during their recovery process seems to be a promising field to enrich assessment and decision making in neuropsychological rehabilitation, but this will also be a Sisyphean labour.

In conclusion, it is recommended to use the differences found in the age and aetiology-based strata and those found in norm groups derived from distinct health services according to the implications and limitations described above.
REFERENCES


Manuscript received 12 May 1995
Revised manuscript received 3 November 1995