Statistical and Clinical Evaluation of the Mattis Dementia Rating Scale–Spanish Adaptation: An Initial Investigation

Bill R. Arnold, Israel Cuellar, and Norma Guzman

The University of Texas–Pan American, Edinburg.

The usefulness of the Mattis Dementia Rating Scale (MDRS) for Spanish dominant elderly adults has seen little empirical documentation, though there is considerable reason to hypothesize its value with this diverse and growing population. A Spanish adaptation of the MDRS (MDRS-SA) was developed for the present study in a manner to facilitate linguistic equivalence. Reliability for the MDRS-SA showed high split half internal consistency across all scales and Total Score. Two groups of elderly adults, one neurologically impaired and one not, both with low Mexican American acculturation level, Spanish dominance, and little formal education were participants. Results of a MANCOVA analysis, controlling for age, education, acculturation, and gender, showed the MDRS-SA to be effective at discriminating between groups across all scales (p < .001) and Total Score (p < .05). After determining significant differences between original norms and normative data on the present Spanish group, clinical utility of these statistical findings was investigated by generating specificity, sensitivity, and accuracy rates for the MDRS-SA Total Score, using Spanish normative data. Results of this preliminary study supported the MDRS-SA for use within the Mexican American, Spanish dominant elderly population, though additional larger validation studies are needed, particularly with well-documented clinical samples.

The United States now has the fifth largest Latino population in the world estimated at over 22 million persons (Dana, 1998). Mexican Americans represent the largest portion of Latinos in the United States at 62%, followed by Puerto Ricans (13%), Cuban Americans (5%), Latinos from Central America and other Spanish speaking Latin American countries (12%), and Spain (8%) (Dana, 1998). A significant portion of this Latino population is Spanish speaking, particularly among the elderly adults who are most likely to be first or second generation and to have completed fewer than five years of formal schooling (Cuellar, 1990). The development of appropriate Spanish language instruments for assessing this population is far more complex than simple translation of English language instruments into Spanish. Simply translating an instrument from one language to another is not sufficient to demonstrate cross-cultural equivalence.

Cultural equivalence refers to the extent to which a word, concept, scale, or norm structure may be considered relevant and applicable to cultural groups other than the one in which these elements were developed (Marsella & Kameoka, 1989). Among the types of equivalence noted by Marsella and Kameoka in cross-cultural assessment are: 1) linguistic, 2) conceptual, 3) scale, and 4) norm equivalence. In recent years, a number of studies have examined for other than linguistic equivalence in Spanish translations of widely used neuropsychological instruments (Gurland, Wilder, Cross, Teresi, & Barrett, 1992; Mahurin, Espino, & Holifield, 1992; Melendez & Prado, 1981; Rey & Benton, 1991; Taussig, Henderson, & Mack, 1992). Many studies have reported important ethnocultural and sociocultural influences on prevalence rates for dementia (e.g., Gurland, 1981; Gurland et al., 1992). The influence of education and illiteracy in the assessment of dementia has been extensively examined (e.g., Ardila, Rosselli, & Rosas, 1989; Katzman, 1993; Reis, Guerreiro, & Castro-Caldas, 1994; Zhang et al., 1990). Cross-cultural methodological issues have been numerous and have posed complex challenges in the assessment of dementia (Gurland, 1981; Kittner et al., 1997). Objectives of the present study included translating the Mattis Dementia Rating Scale into Spanish (MDRS-SA; Arnold & Cuellar, 1992) and evaluating its norm equivalence in both a Spanish speaking normative and a Spanish speaking clinical sample of elderly Mexican American adults.

Mahurin, Espino, and Holifield (1992) note, in their review, the dearth of validated Spanish language translations of commonly used neuropsychological test batteries for Latino elders. The diversity of the Spanish speaking population in the United States makes the creation, translation, and validation of Spanish language psychological tests a highly complex and technical process. A few studies have reported relatively higher rates of cognitive impairment among older minority populations (Bird, Canino, Stupec, & Shrout, 1987; Escobar, Burman, & Wilkstrom, 1986; Gurland et al., 1992; Holtzer, Tischler, Leaf, & Meyers, 1984; Lopes-Aqueres et al., 1987). One of these studies, Taussig, Henderson, and Mack (1992), found that a Spanish version of the Mattis Dementia Rating Scale (MDRS; Mattis, 1988) differentiated between mild to mod-
erately impaired Spanish speaking (SS) Alzheimer’s disease patients and a comparison group of nonimpaired Spanish speaking elders, though the sample size was very small \((n = 18)\). Gurland and colleagues (1992) and Mahurin, Espino, and Holifield (1992) suggested that such translated tests may require adjustment of cutoff scores for impairment due to higher false positive rates. Mahurin and colleagues suggested a tendency to underestimate cognitive functioning, with recent empirical evidence supporting this hypothesis (Arnold, Montgomery, Castaneda, & Longoria, 1994). Evaluation of translated tests therefore becomes essential to ensure reliable and valid results.

In one area, evaluation of the equivalence of the MDRS-SA in the present investigation came from a study of its ability to demonstrate different performances for neurologically impaired individuals when compared with nonimpaired individuals. The MDRS has been used with other neurologically impaired samples where its ability to demonstrate discriminative validity was supported (Kovener, Lazar, Lesser, Perecman, Kaplan, Hainline, & Napolitano, 1992; Sadek, Johnson, Paulsen, Salmon, Swenson, & Butters, 1993).

Still another approach to evaluate the equivalence of the two versions in the present investigation came from a determination of the reliability, or internal consistency, of the MDRS-SA. One prior study by Gardner, Oliver-Munoz, Fisher, and Empting (1981) showed split-half reliability for the MDRS at \(r = .90\) for a sample of subjects with diffuse neurological correlates, suggesting a useful test for discriminating level of impairment. These results are similar to those of other related MDRS studies, in that they also yielded similar reliability coefficients (Coblentz, Mattis, Zingesser, Kasoff, Wisniewski, & Katzman, 1973; Vitaliano et al., 1984).

With this in mind, the current study, after determining satisfactory reliability of the MDRS-SA, investigated: (1) the ability of the MDRS-SA to discriminate between neurologically impaired and nonimpaired individuals, and (2) the sensitivity, specificity, and accuracy rates for the MDRS-SA in the present samples.

**Methods**

**Participants**

Participants for the study were drawn from two groups. One hundred and two participants (33 men; 69 women) were selected from a normal elderly population of Spanish dominant Mexican Americans residing in the Rio Grande Valley along the U.S./Mexico border in southern Texas. These individuals were attending four government supported nutritional and social outreach centers located in three adjacent south Texas communities. Attendees drove or were transported by van to each facility for a day of meals and social activities. Participants selected were volunteers who were solicited following a brief overview of the present study presented by a research assistant. This group had a mean age of 71.33 \((SD = 7.81)\) and educational level of 4.62 \((SD = 3.59)\). The Acculturation Rating Scale for Mexican Americans (ARSMA; Cuellar, Harris, & Jasso, 1980) showed a Mexican cultural affiliation, with a mean of 1.99 \((SD = .69)\) for the sample. All participants were retired and on fixed incomes. All were paid \$3 as a token of appreciation for their participation in the study. Individuals were screened out if they reported a prior history of neurological impairment or diagnosed psychiatric disorder on a health questionnaire, which could potentially affect MDRS-SA performance. From these 102 participants, a group of 20 individuals was randomly selected (7 men; 13 women) in order to create two groups of equal numbers for the present investigation. The 20 individuals in this nonimpaired group had an average age of 72.45 \((SD = 6.36)\), educational level of 5.40 \((SD = 3.56)\), and ARSMA of 1.76 \((SD = .58)\).

The second group consisted of 20 Spanish dominant persons (12 men; 8 women) with various neurological and neuropsychiatric disorders, drawn primarily from a private psychiatric hospital, and from the same nutritional and social outreach centers described above, having indicated a CNS level neurological impairment on the history questionnaire described above. Disorders included cerebrovascular accident, senile dementia, and multi-infarct dementia, which had been diagnosed by their attending physician, or was documented by the authors based on clinical presentation (e.g. hemiparesis), subject or family report on a health questionnaire, or by medical records. All displayed mild to moderate levels of functional impairment in adaptive behavior. Mild level of impairment was defined by any indication that the subject had the ability to perform independent adaptive behaviors with minimal assistance, whereas moderate level of impairment was suggested by an ability to perform some independent adaptive behaviors with only some assistance. Information was gathered via observation and questioning of the participant or a relative. Mean age for this group was 72.75 \((SD = 8.33)\), with mean grade completed at 5.65 \((SD = 5.16)\). These individuals were also retired and on fixed incomes. Mean ARSMA rating for this group was 2.08 \((SD = .73)\), indicating a Mexican cultural affiliation.

**Instruments**

**Acculturation Rating Scale for Mexican Americans.**—The Acculturation Rating Scale for Mexican Americans (Cuellar, Harris, & Jasso, 1980) is a bilingual 20-item behavioral rating scale that measures acculturation based on subject responses to questions relating to several content areas, including ethnic identification, language use and preference, peer relationships, food preferences, and generation removed from Mexico. An average rating is obtained for each person, placing them on a unidimensional continuum ranging from “1” (monolingual Spanish, monocultural Mexican) to “3” (bilingual English/Spanish, bicultural Mexican American) to “5” (monolingual English, monocultural Anglo).

**Mattis Dementia Rating Scale.**—The MDRS was developed as a clinical instrument designed to measure general cognitive ability of individuals displaying neurological dysfunction (Mattis, 1988). Published norms are based on a sample of 85 healthy, elderly, English speaking community...
dwellers ranging in age from 65 to 81. Items are grouped into five scales measuring attention, initiation and perseveration, construction, conceptualization, and memory. A Total Score is also generated summarizing performances across scales. Items on MDRS scales are presented in a fixed order, with the most difficult items administered first. As a result, if the difficult item is passed, the examiner provides credit for a correct performance and proceeds to the next scale, significantly decreasing testing time. Using MDRS Total Score, a prior study of diagnostic accuracy of this instrument showed a sensitivity, or correct classification rate, of 62% for demented patients, and 36% for focal brain damaged patients (Montgomery and Costa, as cited in Mattis, 1988).

Spanish adaptation of the Mattis Dementia Rating Scale.
—Developing a useful translated version of a testing instrument involves many steps (Brislin, 1976, 1980). Ensuring a reliable translation of the MDRS, in terms of performance requests, was an initial step in the process of developing the current adaptation. The first phase of this process involved use of a professional translation consultant to translate the instrument from English to Spanish, attempting to generate a “generic” translation. The consultant was asked to generate a basic, generic vocabulary to address low educational level issues, as well as regional linguistic variations. At this point, a team of bilingual mental health professionals with linguistic backgrounds from Florida, Puerto Rico, Cuba, Mexico, Texas, Central America, and South America were asked to review the Spanish translation independently and critically, with the English MDRS available for comparison, and to suggest revisions. Subsequently, two bilingual psychologists reviewed the panel members’ suggestions and incorporated revisions judged appropriate to ensure equivalent performances were being requested for the Spanish adaptation. The original translator then reviewed the revised translation for correct grammar, spelling, and syntax. Field trials were conducted with two elderly, monolingual, Spanish speaking psychiatric inpatients who had provided informed consent. The field trials were conducted to evaluate their comprehension of performance requests on the MDRS-SA. Afterwards, minor wording changes were made and the instrument was then ready for research use in the present study.

Participants in each group were selected independently to determine the internal consistency of the MDRS-SA. Internal consistencies for Total Score and for each scale were determined for each of the two groups and scale. Results of statistical analyses provided Spearman-Brown coefficients for the nonimpaired group ranging from $r = .60$ for the Construction subscale, to $r = .96$ for the Initiation/Perseveration subscale. Consistencies were generally higher for the impaired group, with all subscales and Total Score ranging from $r = .85$ to $r = .96$. Table 1 provides internal consistencies for each group and scale.

Procedure
Informed consent was obtained from all participants. A release of confidential information was also obtained from the individuals in the impairment group, or their guardians, in order for researchers to have access to medical records for more precise medical histories and accurate diagnostic information. Demographic data were obtained from the participant, a family member, or the medical record. Subsequently, each participant was administered the ARSMA and the MDRS–SA by a trained graduate research assistant. Data collection typically lasted about an hour and a half.

RESULTS

Analysis of Group Differences
Covariance procedures were used to control for possible differences between the two groups of 20 on independent variables. A multivariate analysis of covariance (MANCOVA) was used to analyze the data for the impaired and nonimpaired groups using MDRS–SA scale raw scores as the dependent variables and age, education, gender, and ARSMA scores as the covariates. One case from the clinical sample had missing ARSMA data, but was included in analyses. The MANCOVA results showed a significant effect for group differences with a Wilks’ Lambda (30, 118) = .001 with $p < .001$ obtained. An analysis of variance (ANOVA) for subscale means showed significant group differences between impaired and nonimpaired groups on all MDRS–SA scales ($p < .001$): Attention, $F(6, 33) = 860.98$; Initiation/Perseveration, $F(6, 33) = 119.42$; Construction, $F(6, 33) = 136.82$; Conceptualization, $F(6, 33) = 119.90$; and Memory, $F(6, 33) = 96.30$.

Means and standard deviations for the two groups for each scale and total MDRS–SA had been computed for

<table>
<thead>
<tr>
<th>Group</th>
<th>ATTN</th>
<th>I/P</th>
<th>CONST</th>
<th>CONCEP</th>
<th>MEM</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonimpaired</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=102)</td>
<td>$r = .71$</td>
<td>$r = .96$</td>
<td>$r = .60$</td>
<td>$r = .79$</td>
<td>$r = .79$</td>
<td>$r = .91$</td>
</tr>
<tr>
<td>Impaired</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=20)</td>
<td>$r = .89$</td>
<td>$r = .89$</td>
<td>$r = .87$</td>
<td>$r = .85$</td>
<td>$r = .94$</td>
<td>$r = .96$</td>
</tr>
</tbody>
</table>

$ATTN =$ Attention
$I/P =$ Initiation/Perseveration
$CONST =$ Construction
$CONCEP =$ Conceptualization
$MEM =$ Memory
$TOT =$ Total Score

Table 1. Internal Consistency Coefficients for the MDRS–SA
these analyses, and are presented in Table 2. Table 2 shows the Attention scale impaired mean at 32.30 (SD = 3.48) and the nonimpaired mean at 33.55 (SD = 2.04). For Initiation/Perseveration, impaired mean was 24.75 (SD = 7.77) and nonimpaired mean was 33.60 (SD = 4.15). Construction and Conceptualization showed impaired means of 5.10 and 26.45, with SDs of 1.55 and 7.00, respectively. Nonimpaired means for these two scales were 5.53 (SD = .86) and 26.51 (SD = 5.22), respectively. The Memory Scale mean for the impaired group was 15.85 (SD = 5.83) and for the nonimpaired group was 20.20 (SD = 2.57).

An additional analysis of covariance (ANCOVA) was performed for the MDRS-SA Total Score with age, education, gender, and ARSMA scores as covariates. These results, also presented in Table 2, show F(1,33) = 5.28 with p < .05. MDRS-SA Total Score mean for impaired group subjects was 104.45 (SD = 21.47), with nonimpaired equal to 118.73 (SD = 10.37).

**Norm Group Comparisons**

A recurrent question in the literature for translated instruments has been whether existing normative data are satisfactory. Marsella and Kameoka (1989) suggested the need for establishing norm equivalence for translated instruments used across cultures. In Table 3, means and SDs for the original Mattis (1988) normative group of 85 nonimpaired English speaking elders and data from the 102 Spanish speaking counterparts from the present study are provided. On each of the MDRS scales, the mean is lower and the measure of variability is greater for the nonimpaired elderly Spanish group. One-sample t tests were conducted to determine if these differences were significant. Significant differences were found between the two normative samples for each of the MDRS scales and Total Score. Table 3 also contains these data. These differences suggest the need for Spanish norms for the MDRS-SA.

**Specificity, Sensitivity, and Accuracy Rates**

To determine the clinical usefulness of these results, further analyses of data provided sensitivity, specificity, and accuracy rates for the MDRS-SA, using normative data generated from the 102 nonimpaired Spanish speaking older adults described earlier in this study (Kaplan & Sacuzzio, 1997). A Total Score of 98 at 2 SD below the mean was used as the cutoff to approximate the typically acceptable classification rate of 85%, as had been done by Mattis (1988). Using MDRS-SA Total Score data, Table 4 shows classification rates by frequency for the impaired and nonimpaired groups in the present study. The detection rate for impaired participants, or sensitivity, was 40%. That is, 40% of the impaired participants were correctly classified as impaired. On the other hand, 97% of the nonimpaired participants were correctly classified, reflecting the specificity of the test. As a summary of clinical usefulness of the test, the accuracy rate of 75% for the MDRS-SA represents the combined sensitivity and specificity of the test using the Spanish norms. Additionally, the MDRS-SA sensitivity rate can be improved further without dramatically affecting the

### Table 3. One-Sample t Test Scale Comparisons of Nonimpaired MDRS English Norm Means and Nonimpaired MDRS-SA Spanish Norm Means From Current Research

<table>
<thead>
<tr>
<th>Group</th>
<th>ATTN</th>
<th>I/P</th>
<th>CONST</th>
<th>CONCEP</th>
<th>MEM</th>
<th>TOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDRS</td>
<td>35.5</td>
<td>35.5</td>
<td>5.8</td>
<td>37.2</td>
<td>23.3</td>
<td>137.3</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>(1.6)</td>
<td>(3.0)</td>
<td>(.6)</td>
<td>(2.6)</td>
<td>(2.1)</td>
<td>(6.9)</td>
</tr>
<tr>
<td>MDRS-SA</td>
<td>33.88</td>
<td>32.52</td>
<td>5.53</td>
<td>26.51</td>
<td>20.29</td>
<td>118.73</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>(1.9)</td>
<td>(4.0)</td>
<td>(.9)</td>
<td>(5.0)</td>
<td>(2.6)</td>
<td>(10.3)</td>
</tr>
<tr>
<td>t values</td>
<td>-18.09*</td>
<td>-7.56*</td>
<td>-3.16</td>
<td>-20.70*</td>
<td>-11.79*</td>
<td>-8.51*</td>
</tr>
<tr>
<td>p values</td>
<td>&lt;.001*</td>
<td>p &lt;.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*ATTN = Attention  
*I/P = Initiation/Perseveration  
*CONST = Construction  
*CONCEP = Conceptualization  
*MEM = Memory  
*TOT = Total Score

### Table 4. MDRS-SA Total Score Classification Rates Using Spanish Norms at 2 SD Below the Mean

<table>
<thead>
<tr>
<th>Actual</th>
<th>Impaired</th>
<th>Nonimpaired</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurological Deficit</td>
<td>A</td>
<td>B</td>
<td>20</td>
</tr>
<tr>
<td>Normal</td>
<td>C</td>
<td>D</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>111</td>
<td>122</td>
</tr>
</tbody>
</table>

A = Hit (true positive)  
B = False negative  
C = False positive  
D = Hit (true negative)  
Sensitivity = A/(A + B) = 40%  
Specificity = D/(C + D) = 97%  
Accuracy Rate = (A + D)/(A + B + C + D) = 75.3%
specificity of the test. Frequency data from the two Spanish samples in the present study can be used to identify the cutoff score to maximize correct classifications for both. Raising the MDRS–SA raw score cutoff for Total Score to 105 improves correct detection to 55% of the impaired group, and still correctly identifies 92% of those nonimpaired.

**Discussion**

The current findings of this preliminary investigation suggest that the MDRS–SA may prove to be a reliable and clinically useful measure of cognitive abilities when attempting to discriminate between impaired and nonimpaired Spanish speaking elders, or track cognitive decline. These results appear to be comparable to those found by Taussig and colleagues (1992) who showed that the MDRS was able to differentiate between elderly Spanish speaking groups. However, this earlier translation did not report the translation process, reliability data, or classification rate data, and consisted of homogeneous groups of Alzheimer’s patients. The sensitivity rate of 40% to a maximized rate of 55% in the present study appears similar to the rate ranging from 36% to 62% reported in the study by Montgomery and Costa (as cited in Mattis, 1988). Even with a classification rate comparable to other studies, it should be noted that the maximized sensitivity of 55% also indicates an incorrect, false negative classification of 45% of the sample as not impaired, when in fact they were part of the impaired group. For maximum correct detection of impairment, the collaborative use of other sources of patient data appears essential for the MDRS–SA, just as it is for the MDRS, as well as other neuropsychological measures.

Sample size for the current study was small, limiting the generalizability and potential usefulness of findings in the present investigation, making it a preliminary investigation of these issues. Large sample clinical studies are essential to further determination of the reliability with which data from this MDRS–SA can be used to make important clinical decisions. Sample size may contribute to the relatively low detection rates found in studies of neurologically impaired samples. In addition to sample size, an additional limitation of the present study involved the selection of participants for the impaired group. No external psychometric or functional criterion documenting cognitive impairment was used, with reliance primarily on reported medical, historical, and observational data. Future studies may strengthen this methodological weakness by including other measures of cognitive functioning which have been previously validated with Spanish elders, and by using medical data frequently correlated with cognitive deficits, such as MRI or PET scan information.

The present study demonstrated the linguistic equivalence of the MDRS–SA. Perhaps more importantly, it demonstrated the importance of determining normative equivalence of translated instruments used across cultures. Because of the significant differences found between the two norm groups, a high 56% rate of “false positives” (incorrect classification of nonimpaired Spanish speaking elders as neurologically impaired) would have been generated if original English norms had been used. This finding appears consistent with warnings provided by Arnold and colleagues (1994), Gurlin and associates (1992), and Mahurin and colleagues (1992), though the proposal by these researchers for adjusting cutting scores to correct this problem may not prove a satisfactory solution in all cases.

Further research is needed to clarify each of these issues due to the documented needs of the elderly Hispanic population. Valid and reliable assessment tools are needed in order to improve diagnostic accuracy, track functional cognitive abilities, and provide adequate services to this growing population of patients, with the MDRS–SA appearing to be one of few instruments with preliminary, limited empirical support in this regard.

**Acknowledgments**

Acknowledgment and appreciation for assistance with this study is provided to Amigos del Valle, Charter Palms Behavioral Health Systems, and Psychological Assessment Resources, Inc. Technical assistance provided by Roberto E. Maldonado also was greatly appreciated. This project was supported in part by Grant # GM080838-21 from the Minority Biomedical Research Program—National Institute for Deafness and Communication Disorders, National Institutes of Health, Washington, DC. This study is based in part on papers presented at the annual conference of the Texas Psychological Association, 1994, and the annual conference of the National Academy of Neuropsychology, 1997. The MDRS–SA may be obtained from the Research Division of Psychological Assessment Resources, Inc., P. O. Box 998, Odessa, FL 33556-9901.

Address correspondence to Dr. Bill R. Arnold, Department of Psychology and Anthropology, University of Texas–Pan American, Edinburg, TX 78539. E-mail: barnold@panam.edu

**References**


MATTIS DEMENTIA RATING SCALE

P369


Received May 30, 1997
Accepted June 11, 1998