Personality assessment in personnel selection using the MMPI-2: A cross-cultural comparison

Antonio Zapata-Sola (Puerto Rican Electric Power Authority, Puerto Rico),
Tony Kreuch (Sandia Laboratories Human Reliability Program, USA),
Richard N. Landers (University of Minnesota, USA),
Tim Hoyt (University of New Mexico, USA), and
James N. Butcher1 (University of Minnesota, USA)

ABSTRACT. The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) is the most widely used and researched clinical personality assessment instrument in personnel selection. This instrumental study addresses the comparability of the MMPI-2 in assessing Spanish-speaking employees to the performance of English-speaking employees. A sample of 332 Puerto Rican power plant repairers and installers, who were administered the test in Spanish, were compared with 327 English speaking employees from the U.S mainland. The overall MMPI-2 performance of both groups of participants was highly similar and well within the normal range with most scales within the standard error of measurement. A few scales, such as the Lie scale (L), showed small differences with Hispanic clients scoring higher than the Anglos, a finding that has been reported in other studies. The MMPI-2 symptom scales for Hispanics are generally similar to the Anglo scores and do not require adjustments in interpretive strategies. In personnel settings where test defensiveness is common, interpreters need to assure that the applicant’s L score is within the interpretive range.


RESUMEN. El Inventario de Personalidad Multifásico de Minnesota -2 (MMPI-2) es el más ampliamente utilizado y estudiado en el contexto de la evaluación de la persona-
lidad en selección de personal. Este estudio instrumental se ha ocupado de comparar la utilización del MMPI-2 en la evaluación de empleados hispanohablantes con la ejecución de los empleados anglosajones. Una muestra de 332 reparadores e instaladores de líneas eléctricas puertorriqueños a la que se administró el test en español fue comparada con otra muestra de de 327 empleados anglosajones en los Estados Unidos. La ejecución general en MMPI-2 de ambos grupos de participantes fue altamente parecida y se encontró dentro del rango normal con la mayoría de las escalas dentro del error estándar de la medida. Algunas escalas, tales como la escala de la Mentira (L) mostraron pequeñas diferencias, ya que los hispanos puntuaron más alto que los anglosajones, lo cual ya se había encontrado en otros estudios. Las puntuaciones en escalas de síntomas del MMPI-2 en los hispanos son generalmente parecidas a las de los anglosajones y no requieren ajustes en las estrategias de interpretación. En selección de personal, donde la defensividad en una prueba es común, los interpretadores necesitan asegurarse que la puntuación del solicitante en la escala L esté dentro del rango de interpretación.


Clinical assessment and the appraisal of emotional adjustment of applicants in personnel settings have a long history in psychology (Bernreuter, 1931; Woodworth, 1920). In many industrial organizational programs the personality assessments used typically involve evaluation of general personality attributes or dimensions, for example, conscientiousness or agreeableness, and minimize the use of personality testing that addresses clinical or personality problems (Butcher, Gucker, and Hellervik, 2009). The assessment of psychological problems and emotional adjustment has been an important component, however, in some types of personnel evaluations, particularly for positions that involve high responsibility and public safety (airline pilots applicants or nuclear power plant employees) and emotional maturity (police and fire department personnel). The MMPI-2 is the most widely used personality scale for screening employees when public safety is concerned (Butcher et al., 2009; Butcher, Ones, and Cullen, 2006). For a listing of MMPI/MMPI-2 personnel references see www.umn.edu/mmpi.

A substantial research base on MMPI/MMPI-2 scores of Hispanic clients has accrued across a broad range of clinical studies (Butcher, Cabiya, Lucio, and Garrido, 2007) with the publication of over 550 articles and books on this population. The effectiveness of the test with Spanish language populations is underscored by the large number of translations of the MMPI and MMPI-2 and its broad use in Spanish speaking countries around the world (Butcher, 1996). Hall, Bansal, and Lopez (1999) performed a meta-analysis comparing research on Hispanics and Anglos across a number of published studies and found relatively few differences emerging and concluded that differences between ethnic groups were trivial. Butcher et al. (2007) and Cabiya (1994) have pointed out that the American norms appear to work effectively in assessing Hispanic clients. Strong support for the use of the American norms to assess Latino clients living in the
United States has been provided by a number of studies. Construct validation was provided in several MMPI-2 studies in which Hispanics were administered the English language version of the MMPI-2 and scored on United States norms. One recent study reported that the Hispanic and White clients with Post Traumatic Stress Disorder (PTSD) produced comparable results (Mason, 1997). One study (Haskell, 1996) found essentially the same pattern of scale scores for Hispanic and Anglo clients with only a small difference on the Lie (L) scale.

Most research on the MMPI/MMPI-2 with Hispanics has been conducted with mental health and medical populations, however, some empirical research on Hispanics in employment settings has been reported (Casullo, 1996; Casullo, Samartino, Brenella, Marquez, and Dupertuis, 1996) and some studies have examined work compensation for Latino workers (Clark, Velasquez, and Callahan, 1992, 1993). In addition, one study examined Hispanics and Anglos in a screening context showing minimal differences with Anglos (Wexler, 1996).

The present instrumental study (Carretero-Dios and Pérez, 2007; Montero and León, 2007) was designed to provide comparisons of a large sample of Spanish speaking applicants with a similar sample of English speaking applicants in order to evaluate the comparability of the MMPI-2 for diverse populations. The possibility that cultural factors might impact test performance in personnel screening needs to be evaluated to assure that the test fairly assesses employees from diverse backgrounds.

Method

Two similar employment settings that use the MMPI-2 in employee evaluations to assess mental health adjustment were studied. The Spanish-speaking group, powerplant repairers and installers from Puerto Rico, was compared with an English speaking sample from a Department of Energy group from the U.S Mainland.

Description of the Puerto Rico setting

The position of electrical power line installer and repairer is considered to be a high risk position by the Puerto Rico Electric Power Authority (PREPA). As part of the assessment process for employees, licensed psychologists conduct a psychological evaluation that includes the administration of the MMPI-2 and a structured interview for the job applicants. The tasks of the power line installer position contain elements of significant safety risk. Applicants are required to adhere to safety practices and procedures, such as checking equipment regularly and erecting barriers around work areas. They must open switches or attach grounding devices in order to remove electrical hazards from disturbed or fallen lines or to facilitate repairs. They may need to climb poles or use truck-mounted buckets to access equipment and place insulating or fireproofing materials over conductors and joints. They often install, maintain, and repair electrical distribution and transmission systems, including conduits, cables, wires, and related equipment such as transformers, circuit breakers, and switches. They must identify defective sectionalizing devices, circuit breakers, fuses, voltage regulators, transformers, switches, relays, or wiring, using wiring diagrams and electrical-testing...
instruments and drive vehicles equipped with tools and materials to job sites. It is important that they coordinate work assignment preparation and completion with other workers. They inspect and test power lines and auxiliary equipment to locate and identify problems, using reading and testing instruments. They may string wire conductors and cables between poles, towers, trenches, pylons, and buildings, setting lines in place and using winches to adjust tension, tasks that require some risk.

Participants
The Puerto Rican participants in this study were 332 male applicants for the electrical power line installer and repair position. Each applicant had been previously screened to verify the minimum qualifications including completion of high school, passing an ability test and a physical examination. The mean age of the sample was 30.82 and standard deviations of 6.82. The education level was as follows: the entire sample had a high school education and 15.60% had further vocational studies; 18.30% had attended university; and 15.60% of the sample had taken technical school courses.

Control study
The Mainland U.S. sample was comprised of 327 workers at a national nuclear facility who were participants in a Department of Energy safety and security program initiated in 2004, the Human Reliability Program (HRP). The purpose of the HRP is to continually evaluate individuals employed in certain safety-sensitive positions that afford access to nuclear materials at the national nuclear laboratories and identify those whose judgment or psychological functioning may be sufficiently impaired so as to represent a reliability or safety concern. As part of this program, participants undergo an annual physical and psychological evaluation, including a triennial administration of the MMPI-2. The subjects for the study included 327 males with a mean age of 44.49 (SD = 9.85) and a mean educational level of 14.58 years (SD = 3.53). All participants were employed in occupations that were considered to be «safety-sensitive» as follows: Scientists/Engineers (n = 104) Technologists (n = 19), Materials Handlers (n = 69), Security Officers (n = 163), Administrative Support (n = 5) and Miscellaneous (n= 9). Ages and educational levels for the major occupational groups were as follows: Scientists/Engineers: Mean Age 46.58 (10.22), Mean Education 17.76 (3.25); Technologists: Mean Age 47.01 (9.55), Mean Education 13.59 (3.86); Materials Handlers: Mean Age 45.58 (8.08), Mean Education 13.56 (2.66); Security Officers: Mean Age 41.70 (9.24), Mean Education 13.24 (2.35).

Procedure
The MMPI-2 was administered in Spanish by booklet and scored by electronic scanner using software from Pearson Assessments. The applicants were tested in groups as a component of their medical and psychological screening in the pre-employment phase. For purposes of the assessment evaluation, any candidates that invalidated the test (L > 81 T) were administered the test again. The initial MMPI-2s were included in the study in order to evaluate the impact of high L on the population. The trained test administrator or a psychologist described the purpose and procedures involved in the testing program and asked the applicant to sign a release form. They were provided the
answer sheet and MMPI-2 booklet, and a glossary sheet for some words and the meaning for any items they might not understand. This procedure is commonly used in Mexico and Puerto Rico to assure that the items are understood by the individual. The MMPI-2 answer sheets were optically scanned using the Pearson Assessments Q Local Software System and the responses were stored in a computer file and processed in an SPPS (V16.0) file for the statistical analysis.

All participants in the U.S. Mainland sample were from Sandia National Laboratories Human Reliability Program (HRP). Participants were administered the MMPI-2 as a standard part of their annual HRP medical suitability evaluation. The participants were individually administered the MMPI-2 during their first HRP assessment cycle between 2004 and 2006. They took the MMPI-2 as part of an examination day that included a thorough physical examination, standardized drug testing and a psychological interview by a licensed psychologist. In addition to the MMPI-2, the participants were also administered a Mini Mental Status Exam (Folstein, Folstein and Fanjiang, 1975), the OQ 45 Symptom Checklist (Umphress, Lambert, Smart, Barlow, and Clouse, 1997), the Alcohol Use Disorders Identification Test (AUDIT) (Babor, Higgins-Biddle, Saunders, and Monteiro, 2001) and the Life Change Index (Bartone, 1987) and all instruments in addition to a clinical interview, job task analysis, supervisor review checklist and an Occupational Medicine records review were incorporated into the eventual psychological examination findings. All participants signed a release form for research participation. The MMPI-2 was administered to all participants in English in accordance with the standard MMPI-2 test administration instructions by trained administrative personnel through the Occupational Medicine Department EAP at Sandia National Laboratories, Albuquerque, New Mexico with oversight by the HRP Site Designated Psychologist. The booklet form of the MMPI-2 was utilized and the answer sheets were subsequently optically scanned for scoring with the Adult Extended Score Report generated and stored and maintained on a Sandia Occupational Medicine computer using the Pearson Assessments Q Local Software System database. The MMPI-2 report was also included in the participants HRP Psychological File for review by the examining psychologist.

Results

The results of the study show that the Puerto Rican participants, taking the MMPI-2 in Spanish, and the US Mainland employees, taking the test in English, produce results that are highly comparable to the test scores of the MMPI-2 normative sample (see Figures 1 and 2). An examination of Figures 1 and 2 indicates that the majority of the MMPI-2 scales fall near the MMPI-2 normative sample mean ($T = 50$). About 1/3 of the clinical and content scales for Puerto Rican applicants fall below those of the U.S. mainland sample. Thus, there is no general tendency for the MMPI-2 to overpathologize Hispanic clients that is, showing consistently more symptoms among this group. Both employee groups produced test scores on the clinical and validity scales that were in the normal range, most falling within the Standard Error of Measurement for the MMPI-2 scales.
FIGURE 1. MMPI-2 validity and clinical scale profiles for the Puerto Rican sample and the U.S. mainland.

**MMPI-2**

![Graph showing MMPI-2 scale profiles for Puerto Rican and U.S. mainland samples.]


FIGURE 2. MMPI-2 Content Scale profiles for the Puerto Rican sample and the U.S. mainland Sample.

**MMPI-2**

![Graph showing MMPI-2 content scale profiles for Puerto Rican and U.S. mainland samples.]

Notes. ANX = Anxiety, FRS = Fears, OBS = Obsessiveness, DEP = Depression, HEA = Health, BIZ = Bizarre Mentation, ANG = Anger, CYN = Cynicism, ASP = Antisocial Personality, TPA = Type A, LSE = Low Self Esteem, SOD = Social Discomfort, FAM = Family Problems, WRK = Work Adjustment, TRT = Treatment Planning.
Even though the mean scores for both industrial samples were within the normal profile range we conducted further statistical analyses to investigate the equivalence of the MMPI-2 between the Puerto Rican and U.S. Mainland samples, including a MANOVA to investigate the appropriateness of further tests. Finding that the overall differences between the samples on all of the scales were unlikely to have occurred by chance (Multivariate $F_{(28, 626)} = 37.90, p < .01$), we conducted individual ANOVAs on each of the 28 scales. Because the investigation was exploratory in nature, a Bonferroni correction was applied to tests of statistical significance, such that alpha was divided by 28 for each $F$-test. Considering the large sample size, effect sizes ($\eta^2$) were also calculated to capture the percentage of variance in each of the scales explained by sample membership. Results appear in Table 1. Although several statistically significant differences were found, the effect sizes of sample membership varied from small to moderate.

It was noted that there was a sizable average age difference between the two samples ($t_{(656)} = 20.54, p < .01, \eta^2 = .39, M_{US} – M_{PR} = 13.59$). Because some evidence has suggested that there may be some modest relationships between some of the subscales and age (Spiro, Butcher, Levenson, Aldwin, and Bosse, 2000), correlations between the scales and age were computed across samples (see Table 1) to gauge the appropriateness of a covariate in the analyses. As several moderate correlations were found, a MANCOVA was conducted. Again finding that the overall differences between the samples on all of the scales were unlikely to have occurred by chance (Multivariate $F_{(28, 625)} = 23.80, p < .01$) and that the same could be said for the age covariate (Multivariate $F_{(28, 625)} = 3.53, p < .01$), we conducted ANCOVAs on each of the 28 scales. A Bonferroni correction was used here as well, and effect sizes were also calculated. Results also appear in Table 1. The pattern of statistical significance found was similar to those of the MANOVA, but the effect sizes of sample membership decreased. All differences found were small, and within the typical range expected for the MMPI-2, similar to the scores of the normative sample. Across all scales, an average of only 4% of the variance in the MMPI-2 could be explained by differences between the U.S- Mainland and Puerto Rican samples.

**TABLE 1.** Descriptives, mean differences, and comparison statistics for Puerto Rico ($n = 328$) and U.S. Mainland ($n = 327$) samples.

<table>
<thead>
<tr>
<th></th>
<th>Puerto Rico</th>
<th>U.S. Mainland</th>
<th>Mean difference</th>
<th>$r_{age}$</th>
<th>ANOVA $F$ (df = 1, 653)</th>
<th>Sample effect ($\eta^2$)</th>
<th>ANCOVA $F$ (df = 1, 652)</th>
<th>Sample effect ($\eta^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>7.75</td>
<td>2.95</td>
<td>5.88</td>
<td>2.74</td>
<td>1.87</td>
<td>-.26**</td>
<td>68.80**</td>
<td>.10</td>
</tr>
<tr>
<td>F</td>
<td>3.46</td>
<td>2.96</td>
<td>2.35</td>
<td>1.69</td>
<td>1.11</td>
<td>-.09*</td>
<td>34.68**</td>
<td>.05</td>
</tr>
<tr>
<td>K</td>
<td>19.50</td>
<td>4.40</td>
<td>20.76</td>
<td>4.06</td>
<td>-1.26</td>
<td>.03</td>
<td>14.93**</td>
<td>.02</td>
</tr>
<tr>
<td>Hs</td>
<td>3.78</td>
<td>2.98</td>
<td>2.87</td>
<td>2.56</td>
<td>.91</td>
<td>.02</td>
<td>18.16**</td>
<td>.03</td>
</tr>
<tr>
<td>D</td>
<td>19.63</td>
<td>3.38</td>
<td>17.42</td>
<td>3.01</td>
<td>2.21</td>
<td>-.11**</td>
<td>78.29**</td>
<td>.11</td>
</tr>
<tr>
<td>Hy</td>
<td>21.93</td>
<td>4.20</td>
<td>22.05</td>
<td>3.41</td>
<td>.12</td>
<td>.12**</td>
<td>.19</td>
<td>.00</td>
</tr>
<tr>
<td>Pd</td>
<td>14.01</td>
<td>3.50</td>
<td>14.62</td>
<td>3.10</td>
<td>-.61</td>
<td>.07</td>
<td>5.79</td>
<td>.01</td>
</tr>
<tr>
<td>Mf</td>
<td>21.73</td>
<td>3.66</td>
<td>21.73</td>
<td>3.95</td>
<td>0</td>
<td>.13**</td>
<td>12.64*</td>
<td>.02</td>
</tr>
<tr>
<td>Pa</td>
<td>9.23</td>
<td>2.53</td>
<td>9.89</td>
<td>2.24</td>
<td>-.66</td>
<td>.11**</td>
<td>12.64*</td>
<td>.02</td>
</tr>
</tbody>
</table>
TABLE 1. Descriptives, mean differences, and comparison statistics for Puerto Rico (n = 328) and U.S. Mainland (n = 327) samples (cont.).

<table>
<thead>
<tr>
<th></th>
<th>Puerto Rico</th>
<th>U.S. Mainland</th>
<th>Mean difference</th>
<th>r&lt;sub&gt;n&lt;/sub&gt;, ANOVA F (df = 1, 653)</th>
<th>ANOVA F (df = 1, 652)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>Sample effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(df = 1, 653)</td>
</tr>
<tr>
<td>Pt</td>
<td>6.39</td>
<td>4.84</td>
<td>4.99</td>
<td>4.03</td>
<td>1.40</td>
</tr>
<tr>
<td>Sc</td>
<td>7.62</td>
<td>5.73</td>
<td>8.50</td>
<td>3.88</td>
<td>2.82</td>
</tr>
<tr>
<td>Ma</td>
<td>18.07</td>
<td>3.88</td>
<td>14.13</td>
<td>3.38</td>
<td>3.94</td>
</tr>
<tr>
<td>Si</td>
<td>20.17</td>
<td>5.94</td>
<td>20.78</td>
<td>7.24</td>
<td>-6.1</td>
</tr>
<tr>
<td>ANX</td>
<td>4.27</td>
<td>3.09</td>
<td>2.13</td>
<td>2.45</td>
<td>2.14</td>
</tr>
<tr>
<td>FRS</td>
<td>3.46</td>
<td>2.86</td>
<td>2.40</td>
<td>.60</td>
<td>.05</td>
</tr>
<tr>
<td>OBS</td>
<td>3.03</td>
<td>2.80</td>
<td>1.94</td>
<td>2.23</td>
<td>1.09</td>
</tr>
<tr>
<td>DEP</td>
<td>2.05</td>
<td>2.49</td>
<td>1.26</td>
<td>1.86</td>
<td>.79</td>
</tr>
<tr>
<td>HEA</td>
<td>4.74</td>
<td>3.22</td>
<td>2.73</td>
<td>2.14</td>
<td>.02</td>
</tr>
<tr>
<td>BIZ</td>
<td>2.99</td>
<td>2.57</td>
<td>2.93</td>
<td>1.45</td>
<td>2.07</td>
</tr>
<tr>
<td>ANG</td>
<td>2.67</td>
<td>2.44</td>
<td>2.33</td>
<td>2.04</td>
<td>.34</td>
</tr>
<tr>
<td>CYN</td>
<td>8.87</td>
<td>5.07</td>
<td>4.80</td>
<td>4.25</td>
<td>4.07</td>
</tr>
<tr>
<td>ASP</td>
<td>5.99</td>
<td>3.81</td>
<td>4.13</td>
<td>3.02</td>
<td>1.86</td>
</tr>
<tr>
<td>TPA</td>
<td>5.05</td>
<td>3.58</td>
<td>4.31</td>
<td>2.90</td>
<td>.74</td>
</tr>
<tr>
<td>LSE</td>
<td>1.66</td>
<td>2.28</td>
<td>1.78</td>
<td>2.36</td>
<td>-.12</td>
</tr>
<tr>
<td>SOD</td>
<td>4.71</td>
<td>3.35</td>
<td>6.14</td>
<td>4.39</td>
<td>1.43</td>
</tr>
<tr>
<td>FAM</td>
<td>2.65</td>
<td>2.80</td>
<td>2.56</td>
<td>2.19</td>
<td>.08</td>
</tr>
<tr>
<td>WRK</td>
<td>2.75</td>
<td>2.91</td>
<td>2.70</td>
<td>2.90</td>
<td>.05</td>
</tr>
<tr>
<td>TRT</td>
<td>2.61</td>
<td>2.68</td>
<td>1.83</td>
<td>2.11</td>
<td>.77</td>
</tr>
</tbody>
</table>


Sample effect refers to effect of sample membership (Puerto Rico vs. Mainland) on means.

**p < .01 *p < .05 (Bonferroni-corrected for ANOVA and ANCOVA).

There was no general tendency for Hispanics to be more defensive on the testing than the Anglo sample. The mean Defensiveness or K scale scores (assessing test defensiveness) was actually lower for Hispanics than the U.S. Mainland sample. Although the differences between the two samples on the MMPI-2 L scale were small (1.8 raw score points between groups) we decided to further examine the group’s performance on this scale because applicant’s often attempt to present themselves in a favorable way on the test. The raw score frequency distributions for L scores are shown in Figure 3. The samples were found to have similar distributions with the Puerto Ricans ranging slightly higher (see Figure 3). Given that both assessment programs studied limit their interpretations of the test to applicants with L scores equal to or less than 80 we decided to do a further analysis (see Figure 4). We next eliminated all cases from both samples with an L score ≥ 81 resulting in the raw score distributions shown in Figure 4.
FIGURE 3. Distribution of Lie (L) scale scores for the Puerto Rican sample and the U.S. mainland sample.
FIGURE 4. Distribution of Lie (L) scale scores for the Puerto Rican applicants (n = 332) and the U.S. Mainland applicants (n = 327) compared with the MMPI-2 normative sample (N=1,138).

The results show that both personnel applicant groups have a small number of participants who score at higher elevations on the Lie (L) scale as compared to the MMPI-2 normative sample (see Figure 4). Although the Hispanic applicants tend to have slightly higher L scores (which could be the result of an artifact of the scale), they appear to be generally comparable to the U.S. Mainland applicants.

Discussion

The results of this study indicate that the performance of Hispanic powerplant repairers and installers from Puerto Rico showed few profile symptomatic differences compared with a personnel sample from the Mainland U.S. Their MMPI-2 performance was highly comparable with less than 4% of the variance accounted for by group differences. As a group, the Hispanic applicants tended to obtain slightly higher L scores. This possibility was given consideration to determine if the two groups differ in this defensiveness variable in order to determine if different interpretive strategies for the MMPI-2 profiles need to be taken in personnel settings. Given that this finding has been reported in other studies comparing Hispanic and Anglo groups the difference found in this study was evaluated further.
Although, people taking the MMPI-2 in personnel applications tend to be somewhat defensive in order to present a low problem picture on the test, there was no tendency for the Hispanic applicants to be more defensive on the test than the Mainland U.S. sample as measured by the K scale although both personnel samples were more defensive than the normative MMPI-2 sample as expected. In order to reduce the impact of test defensiveness in personnel settings many assessment programs employ a procedure of retesting defensive applicants using instructions encouraging them to be more open (less defensive) the second time (see discussions by Butcher, Morfitt,., Rouse, and Holden, 1997; Cigrang and Staal, 2001; Gucker and McNulty, 2004; Walfish, 2007). This procedure results in a less defensive response approach for the majority of employees.

The small difference on the L scale between the Hispanic and the U.S. Mainland sample suggests that some Hispanic applicants could obtain a slightly higher score on L than Anglos but this difference is not likely suggestive of a higher defensiveness on the test.

References


Developments in the Use of the MMPI-2/MMPI-A Workshop and Symposia. Minneapolis, MN, USA.