The Psychometric Competence of Minnesota Multiphasic Personality Inventory-2 Reconstructed .(MMPI-2-RF) in Egyptian Sample

Prepared by

Omnia E. EL-Shenawy, Elham A. Khalil, Abdel -Moniem Shehata and Khaled Ziada
Professors of psychology, Dept. of psychology, Faculty of Arts - Menoufia University, Egypt

VOL . (1) NO. (1) July 2023
The Psychometric Competence of Minnesota Multiphasic Personality Inventory–2 Reconstructed (MMPI–2–RF) in Egyptian Sample

Prepared by
Omnia E. EL-Shenawy, Elham A. Khalil, Abdel -Moniem Shehata and
Khaled Ziada
Professors of psychology
Dept. of psychology, Faculty of Arts - Menoufia University, Egypt

Abstract
This study used data from 1624 normal participants (age= 31.52 ± 11 years) and 120 psychiatric patients (age= 32.36 ± 11.36 years), both are Egyptians, to investigate psychometric competence of Minnesota Multiphasic Personality Inventory–2 Reconstructed (MMPI–2–RF) in Egyptian society. Results revealed that MMPI-2 RF was reliable, internally consistent, except some subscales due to little number of items. In addition, there were significant differences between males and females, normal and patient, also between psychiatric patient categories on MMPI-2RF scales. Results were discussed considering literature and provided a strong evidence for the discriminative validity of the MMPI-2 RF scales.

Key words: MMPI2 RF, reliability, validity
Introduction

The Minnesota Multiphasic Personality Inventory (MMPI) instruments have been the most widely used psychological assessment tools for many decades (Marek et al., 2018), despite the wide use of the inventory, many researchers identified significant psychometric problems with the MMPI–2 Clinical Scales including excessive item overlap, insufficient resources for good cross-validation, and the inclusion of subtle items that Hathaway and McKinley thought would be immune to response bias, but more likely were the result of their scale construction strategy (Ben-Porath, 2012; Sellbom, 2019). Tellegen et al., 2003 (Marek et al., 2018) indicated that demoralization is the responsible factor of items overlap and intercorrelations and restructured the MMPI–2 Clinical Scales to remove common demoralization variance and create new scales that assess major distinctive core components of the original measures. Ben-Porath and Tellegen (2008/2011) completed the developing of MMPI-2RF by adding new measures of constructs that available with the MMPI–2 item pool. The final MMPI-2RF consists of nine Validity Scales (7 of which are revised versions of MMPI-2 validity scales); three Higher-Order (H-O) Scales, designed to measure broad dimensions of personality and psychopathology; the nine RC Scales, developed to assess major distinct core components of the original Clinical Scales; twenty-three Specific Problems (SP) Scales, designed to assess more narrowly focused constructs related to somatic and cognitive functioning, internalizing problems, externalizing behavior, and interpersonal functioning; two Interest Scales, derived from the original Clinical Scale 5; and five revised versions of the Personality Psychopathology Five (PSY-5) Scales that provide a dimensional assessment of personality disorder-related psychopathology (Detrick et al., 2016).
The MMPI-2RF offers several advantages, including scales that were developed with modern scale construction techniques, conceptual and empirical links to contemporary models and constructs in personality and psychopathology, a standard interpretation strategy, and a 40% reduction in length (Tarescavage et al., 2015, p. 188). A growing body of empirical research supports using the MMPI-2-RF scales in several setting such as medical setting: predicting treatment outcomes among patients with Chronic Low Back Pain (e.g., Tarescavage et al., 2018), psychological evaluation of bariatric surgery candidates (e.g., Tarescavage et al., 2013); legal setting: criminal forensic assessment (e.g., Sánchez et al., 2017; Sellbom, 2017; Sharf et al., 2017), Intimate partner violence intervention programs (IPVIPs) (e.g., Whitman et al., 2020), treatment success, and recidivism in offenders enrolled in a batterers’ intervention program (e.g., Sellbom, 2008); mental health settings: assessing personality disorders among clinical and forensic samples (e.g., Anderson et al., 2015). In addition, several studies indicated that MMPI-2 RF subscales had good psychometric properties in different samples. Archer et al. (2012) examined the MMPI-2-RF validity, RC, HO, SP, and PSY-5 scale characteristics of men and women undergoing evaluation within a child custody litigation. Results showed consistency between elevations previously found on MMPI-2 Validity scales L and K and elevations on MMPI-2 RF Validity scales L-r and K-r. Results also indicated that RC scale intercorrelation patterns and alpha coefficient estimations for MMPI-2-RF RC, H-O, SP, and PSY-5 scales were consistent with the alpha coefficients for the MMPI-2-RF scales as reported across different populations. Marek et al. (2018) indicated that MMPI–2 RF substantive scale scores were reliable and had good convergent and discriminant validity in a sample of spine surgery patients (n = 810) and in a sample of spinal cord stimulator patients.
Tarescavage et al. (2018) found that MMPI-2 RF scales have reliability estimates similar to the normative sample. And MMPI-2 RF scales were correlated with risk assessment instruments (the STATIC-99 and Level of Service Inventory–Revised (LSI-R) indicated that the scale had good construct validity.

In addition, several studies on the MMPI-2 RF subscales in different cultures such as validity scales (e.g., Gervais et al., 2010; Handel et al., 2010; Ingram & Temes, 2016; Purdon, et al., 2011; Ransom, 2012; Sánchez, et al., 2017), Higher Order Scales (e.g., Anestise et al., 2015; Haber & Baum, 2014; Kim et al., 2015; Rogers et al., 2017; Sellbom et al., 2012), Restructured Clinical Scales RCs (e.g. Arbisi et al., 2011; Hunter et al., 2014; Lijewski et al., 2013; Osberg et al., 2008; Sellbom & Ben-Porath, 2005; Rogers et al., 2006; Sellbom et al., 2008; Shkalim, 2015; Smith, 2010; Van der Heijden et al., 2008; Wolf et al., 2008), Specific Problems (SP) (e.g., Anestis et al., 2015; Sellbom et al., 2012), PSY-5 Scales (e.g., Anderson et al., 2013; Bagby et al., 2014; Harkness et al., 2013; Finn et al., 2014) have supported the psychometric competence of MMPI-2 RF Scales.

Despite this extensive body of research on the MMPI-2 RF in different cultures, little is known about the psychometric competence of the Inventory in Arabic culture. Only two studies attempted to examine the validation of the MMPI-2 RF in Syria. The first study conducted by Ahmed (2018) which aimed at examining the psychometric competence of the MMPI-2 RF among Syrian samples. The first sample consisted of 2482 male and female Damascus, and Tartous Universities students. Distributed over a validity sample (n = 712), and a reliability sample (n = 390), with mean age 20.5 year and the second sample consisted of 608 psychiatric patients. Results indicated the MMPI-2 RF had good reliability estimation α coefficients ranged from
0.37 to 0.87 for normal sample and from 0.45 to 0.90 in psychiatric patients. In addition, there were a statistical correlation between the MMPI-2RF scales scores and scores on Minnesota Multiphasic Personality Inventory-2 (MMPI-2), the Millon Index of Personality Style Revised (MIPS-r), and the Symptom Checklist (90-R). Results also indicated that there were significant differences between males and females, normal and patients on MMPI-2Rf scales. Exploratory Factors analysis for each group of Subscales items resulted in five interpretable factors for validity scales items, three interpretable factors for H-O scale items, six interpretable factors for RC scale items, seven interpretable factors for SP scale items, six interpretable factors for PSY-5 scale items. Using the same two samples, Ahmed (2019) conducted a discriminatory analysis to examine the diagnostic validity of MMPI-2RF Scales. Results indicated that there were differences between normal participants and psychiatric patients on all MMPI-2RF Scales. The classification accuracy rate was 73.4% for validity scales, 77.1% for H-O Scales, 77.1% for RCs, 76.7% for SP scales, and 62.9% for PSY-5 Scales. In addition, Kappa coefficient was statistically significant.

**Current investigation**

Given the importance usage of MMPI-2 RF in several settings (Mental health, treatment, Forensic, addiction, etc.), as discussed above, in addition to the rare of Arab research that examined the psychometric competence of the MMPI-2RF in local environment. The current investigation seeks to extend past findings on the psychometric competence of the substantive and validity scales of the MMPI–2–RF among normal participants and psychiatric patients in different cultures using Egyptian samples. In this investigation we calculated means and standard deviations for all MMPI–2–RF scales T scores separately for Egyptian community and psychiatric samples. We also
examined reliability of scale scores with mean inter-item correlations, internal consistency, and standard error of measurement estimates. We investigated the construct (diagnostic) validity of MMPI-2-RF scales by assess the capacity of the MMPI-2-RF scales to discriminate between normal participants and psychiatric patients and to distinguish patients’ groups with different diagnosis. In addition, we assess differences between Egyptian females and males mean T scores on all MMPI-2RF. We hypothesized that the descriptive findings and reliability estimates in the current investigation would yield similar findings to those found in past research. We also hypothesized that MMPI-2 RF scales would demonstrate good construct validity, with comparable patterns across normal and patients, patients’ groups, and gender.

**Method**

**Participants**

Community sample: consisted of 1665 participants (1067 females, 581 males, and 17 did not mention their gender), aged between 18-80 years (31.52 ± 11 years).

Psychiatric sample: consisted of 97 inpatients from Al-Abbasiya Mental Health Hospital, Abu Al-Azayem Psychiatric Hospital, and Bedaya Psychiatric Center in Shebin El-Kom (Egypt), with an average age 32.36 years, and a standard deviation 11.36. In addition to 23 patients were taken from the community sample who reported that they were psychiatric patients. Therefore, the final sample consisted of 120 patients.

**Measure**

**Minnesota Multiphasic Personality Inventory–2–Restructured Form**
The MMPI-2-RF (Tellegen & Ben-Porath, 2008) is a self-report instrument that assesses personality and psychopathology across 42 substantive scales. The test also assesses protocol validity (inconsistent responding; fixed responding; overreporting of psychological, somatic, or cognitive symptoms; and underreporting response styles) across nine validity scales. The Arabic version of MMPI-2 RF translated by Abdullah Soliman was used after taking Minnesota University permission to standardize the MMPI-2 RF in the Egyptian environment and using it for scientific research purposes.

Results

Descriptive statistics

Means and standard deviations for the 51 MMPI-2-RF scales T scores are presented separately in Table 1 for both community and psychiatric samples, along with scale abbreviations. The values for both samples were in normal range, as they ranged from 48.84 ± 10.02 (BRF) to 51.56 ± 9.75 (FBS-r) for the psychiatric sample, and from 49.89 ± 9.75 (MLS) to 50.4 ± 10.02 (MEC), 50.04 ± 10 (SHY) for the normal sample.

Reliability

Mean inter-item correlations, Cronbach’s alpha internal consistency, and standard error of measurement (SEM) estimates for MMPI-2RF scales are provided in table 2 for both the community (normative) Egyptian sample and the Egyptian psychiatric sample. American normative sample estimates for substantive scales are also included for comparison.

Most of mean inter-item correlation estimates for MMPI-2 RF substantive scales in the current research fall within the adequate range as
indicated by Briggs and Cheek (1986) mean inter-item correlation estimates ≥ .10 are adequate and ≥ .20 are optimal, except for RC1, IPP and AGGR-r scales, which had appreciably lower (.05) mean interitem correlations. In general, the mean inter-item correlation estimates of Egyptian normative sample are lower compared to the American normative sample. Alpha coefficient estimates for the MMPI-2RF substantive scales for normative Egyptian sample are lower than those for the American normative sample, except for, EID, THD, BXD, RCD, RC4, RC6, RC7, RC8, RC9, GIC, SUB, AES, MEC, PSYC-r, DISC-r all of which approximated those from the American normative sample, and for SUI and AXY, both of which had meaningfully higher estimates (≥|.10|). Differences in internal consistency greater than .10 were considered as clinically meaningful in the current research, depending on past investigation of reliability comparability of MMPI-2-RF scores among samples (Tarescavage e a., 2013). SEMs (expressed in T score terms) for the substantive scales in the Egyptian normative sample were lower than those reported for the American normative sample (see Table2).

Regarding the Psychiatric sample, most of mean inter-item correlation estimates for MMPI-2 RF were appropriate considering the criterion approved by Briggs and Cheek (1986) (except for RCI, RC4, RC9, NUC, ANP, MSF, SUB, ACT, FML, SAV, AGGR-r, DISC-r all of which were low. (see Table:2). Alpha coefficient estimates for the validity scales ranged from 0.27 for (TRIN-r) to 0.77 for (VRIN-r). And for substantive scales from 0.08 for SUB scale to 0.84 for RC6. SEMs (expressed in T score terms) for the validity scales ranged from 1.01 for Fs scale to 2.55 for FBS-r scale. And for substantive scales ranged from 3.59 for EID scale to 0.84 for SFD scale.
Diagnostic validity

A series of one-way analyses of variance (ANOVAs) were conducted to compare mean scale T score differences between the diagnostic groups (Schizophrenia group (SC) N= 27; mood disorders (MD) group N=17; anxiety disorders group (AN) N= 16, and substance abuse group (SUB) N= 35). A post hoc pairwise comparison differences between groups were assessed using the Tuckey procedure, with alpha set at p ≤ .01. For H-O scales, the ANOVA was significant for THD and BXD scales F(3, 91) = 4.38, and 4.04 p < .01 respectively, Table (4). Pairwise comparison results indicated that patients with SC scored higher on the THD scale compared to patient with either MD or AN, also patients with SC scored higher on the BXD scale than patient with AN. For the RC scales, the ANOVA was significant for three scales: RC4, RC6, and RC8 scales F (3, 91) = 3.65, 3.41, and 3.11, respectively. And these scales yielded significant, pairwise differences, as expected patients with SUB scored higher on RC4 scale compared with patient with SC and patients with SC scored higher on RC6 scales compared with patients with AN. For Specific Problems scales, the ANOVA was significant for only five scales: NUC, BRF, JCP, SUB, and FML. All five scales yielded significant group, pairwise differences. Patients with SC scored higher on NUC and FML scales than patients with SUB. Patients with SC also scored higher on BRF and JCP scales than patients with AN, and patients with SUB scored higher on JCP and SUB scales compared with patients with AN. For Interests scales, ANOVA was significant for MEC scale F (3, 91) = 7.93. pairwise comparisons indicated that patients with SC scored higher on MEC scale compared with patients with AN or MD, and patients with SUB scored higher on MEC scale than patients with AN. For PSY-5 scale, ANOVA results were significant only for DISC-r scale F(3,
DISC-r scale yielded significant group, pairwise differences. Patients with SUB scored higher on DISC-r scale than patients with AN. The means, standard deviations, and statistical tests of difference for the ANOVA analyses are displayed in Table 3.

**Here Table (3)**

**Group differences**

Validity also assessed in the current research by calculating mean T scores for MMPI-2RF scales between psychiatric patients and normal participants, using t-test for independent samples analysis. Normal participants sample was drawn from the community sample and was equivalent to psychiatric patients sample in terms of gender and age (n = 120). Gender differences were also calculated using t-test for independent samples analysis. Table 4 shows the results of this analysis. All statistical tests were supplemented with effect size estimates (Cohen’s d), with d values of .20, .50, and .80 reflecting small, medium, and large effect sizes (Schafer and Schwarz, 2019). Results indicated that there were significant differences between patients and normal participants only on few scales, as psychiatric patients scored higher than normal participants on THD, RCd, RC3, RC8, NUC, NFC, and PSYC-r scales.

Males and females produced very similar scores. However, there were a few small gender differences that were consistent with what is reported in previous studies. Females producing higher over-reporting scales, somatic/cognitive and internalizing scale scores and men producing higher under-reporting scales, externalizing scale T-score means when significant differences are found.

**Here Table (4)**

**Discussion**
The results of the current research are an initial step in calculating the psychometric competence of MMPI-2 RF in the Egyptian environment. Mean T scores of the psychiatric patients’ sample and the normative sample in the current research were in the normal range. This result can be explained through two points of view: The first is concerned with the difference between statistical significance and clinical significance meaning. The statistical significance is the possibility that the results of the study were due to chance and not to the effect of the real treatment. The traditional limit for an acceptable significance value is 0.05 (or 5%). While clinical significance refers to the size of the actual treatment effect (that is, the difference between the two intervention and the control groups, also known as "treatment effect size") (Ranganathan, et al., 2015). From this standpoint, the absence of significant statistical differences between mean T scores for normal participants and Psychiatric patients can be attributed to the treatment effect as it produced a clinical significance among patients and caused a real change in patients’ lives, which in turn reduced the statistical difference between the two groups. Another explanation comes from simulation studies’ results, such as the study of Odland, et al. (2011) which used the "Monte Carlo" methodology. Authors found that 36.8% of normal adult show at least one high score on clinical scales of 65T. Odland et al. (2015) results showed that when all 40 MMPI-2-RF scales are simultaneously considered, 70% of normal adults are likely to have at least one scale elevation at or above 65T and 20% will have five or more elevated scales. When the Restructured Clinical (RC) scales are under consideration, 34% of normal adults have at least one elevated score. The scales for specific problems and the PSY-5 which were revised resulted in higher-than-expected rates of significant scores. This increases the likelihood of increasing the normative sample.
scores to approximate those of the patients’ sample, resulting in no statistically significant differences.

The reliability of the MMPI-2 RF scales was estimated in the current research by calculating the mean-inter item correlations, alpha coefficients ($\alpha$), and the values of the standard error of the measurement (SEM) estimates. The results showed that most mean-inter item correlations (Table 2) were in the good range indicated by "Briggs & Cheek" except for RC1, IPP, and AGGR-r scales for the normative sample, and RCI, RC4, RC9, NUC, ANP, MSF, SUB, ACT, FML, SAV, AGGR-r, DISC-r for the psychiatric sample which were low. Also, the Cronbach alpha coefficient estimates in the current research were not appropriate for some validity scales, especially the (TRIN-r). This result is not surprising, because this scale is designed to be content-free indicator and has greater restriction of range than other validity indicators. Thus, the restriction of the range due to the sample characteristics and the nature of the validity scales themselves reduce the estimates of reliability coefficients to some extent (Ben-Porath, 2012). While the alpha coefficients were appropriate for the higher-order scales and the RCs, except for the alpha estimates for RC1, RC2 and RC3 which were below 0.70; While alpha coefficients estimates were low for specific problems scales, this is due to the shortness of the items that compose these scales. This result is consistent with previous research results (e.g., Archer, e a., 2012; Tarescavage, e a., 2013; Rogers, e a., 2017. Which indicated a decrease in alpha coefficients estimates for many specific problems’ scales, and this is consistent with what was reported in the technical manual of the MMPI-2 RF of low estimates of alpha coefficients for some of these scales (Marke, e a., 2018). The SEMs estimates in the current research were less than 3 T scores
The Psychometric Competence of Minnesota Multiphasic Personality Inventory–2 Reconstructed (MMPI–2–RF) in Egyptian Sample

for all MMPI-2RF scales except for, the COG scale for the normative sample, and the EID scale for the clinical sample, that was, respectively, 3.42, 3.59.

Regarding diagnostic validity of the MMPI-2 RF, patients with SC scored higher on THD scale than patients with MD and AN. And patients with SC scored higher on RC6 scale compared to patients with AN. It should be noted that functional impairment of thinking and thoughts of persecution are the main features of schizophrenia patients. These results are consistent with those of the Sellbom, e a. (2012) whose results indicated that the THD and RC6 scales distinguished schizophrenia, major depressive patients and patients with bipolar disorder as schizophrenic patients had higher averages on these scales. In the current research, SC patients also had higher mean scores on the BXD and BRF scales than patients with AN. These results are consistent with research findings that have indicated an association between psychosis and fear (Russell, ea., 2007). The mean scores for SC patients on the FML and NUC scales were more higher compared to patients with SUB. This is consistent with what was stated in the Fifth Diagnostic and Statistical Manual of Mental and Mental Disorders that a schizophrenic patient suffers from a deterioration in one or more important areas in individual life, such as his relationship with family members, friends, or work (Khalil, 2015). The higher mean T scores of patients with SC NUC scale compared to patients with SUB are attributed to the therapeutic effect on SC patients. the results also indicated that patients with SUB scored higher on JCP, SUB, and DISC-r scales compared with patients with AN, these results are consistent with what was reported in the MMPI-2 RF Technical Manual (Haber & Baum, 2014) that psychiatric substance abuse disorders are considered in the context of the MMPI family of Externalizing disorders scales. In addition, the current results are consistent with research results that indicated that there as
appositive correlation between psychopathic personality disorders and substance abuse (e.g., Anderson, e a. 2015; Mariani, e a., 2008; Uzun, e.a. 2006), as the experience of illegally substances use is an exciting experience for those individuals; It makes the world a fast-paced and reinforces their self-image (Uzun, ea., 2006).

The results of the current research also showed statistically significant differences between psychiatric patients and normal sample on THD, RCd, RC3, RC8, NUC, NFC, PSYC-r, and NEGE-r scales as patients had higher scores on these scales than normal participants. Although the size of the effect of these differences was less than the average level, they confirm the discriminatory validity of the MMPI-2 RF.

Gender differences in mean T scores for MMPI-2RF scales showed that there were small significant statistical differences between males and females and did not exceed 3: 4 T-scores except for, the mean T scores for DISC-r, MEC, MSF and BRF scales, as differences ranged from 5: 8 T-score, considering a T score difference of 5 T or more is indicative of clinically significant group differences as used by the extensive MMPI research. The results indicated that females scored higher on (FBS-r) and (RBS) scales than males, and these results were consistent with the results of the Ahmed' study (2018) which indicated that females scored higher on over-reporting scales. The results of the current study indicated that females are higher on malingering than males and reporting an increasing in memory complaints that may be associated with some emotional factors such as depression that may affect the symptoms of memory complaints (Friedman, ea. 2015), this is confirmed by the females higher scores on (RCd) and RC7 scales compared to males. As females have a greater sense of demoralization, pessimism, low
The Psychometric Competence of Minnesota Multiphasic Personality Inventory—2 Reconstructed (MMPI–2–RF) in Egyptian Sample

self-esteem, and excessive oppression. In addition to misery, helplessness, rumination, feeling carefree, sensitivity, and guilt.

The results of the current research also indicated that males scored higher on Behavioral / Externalizing Dysfunction scales such as (BXD, RC4, JCP, MEC, AGGr-r, DISC-r), while females scored higher on Somatic /Cognitive Scales and Internalizing Scales such as (RCd, RC7, MLS, GIC, AXY, BRF, DSF, NEGE-r). These results are consistent with those of Ahmed’s (2018), Archer, e a. (2012), Tarescavage, e a. (2013), and Maza studies, e a. (2019). In general, these results indicated to the consistency of differences between males and females on the MMPI-2 RF scales across different samples and cultures. Males higher scores on the BXD scale indicates that they suffer from dysfunctions that appear through a wide range of problems, with poor ability to control their impulses, a history of criminal behavior, and violent and abusive behavior. This is evidenced by their high scores on RC4, DISC-r and AGGR-r (Friedman, e a., 2015). In addition, females’ higher scores on the RCd and RC7 scales reflect a sense of general dissatisfaction with significant negative emotional experiences such as anxiety, fear, unsteadiness, feelings of sadness and unhappiness, targeting for feelings of guilt, the presence of self-criticism tendencies, and the expectation of failure. But they believe that they have already failed and lack the ability to cope with the current situation or situation. This was confirmed by their high scores on GIC, HPC and AXY scales. In general, these results support the diagnostic validity of the MMPI-2 RF in the Egyptian culture.

Conclusion

This research was conducted after approval by the University of Minnesota to use MMPI-2 RF and standardize it in Egyptian culture. This is considered the first research in Egyptian culture and the second at the level of
Arab cultures after two studies of Ahmed (2018-2019). This research results have proven the psychometric competence of MMPI-2 RF in the Egyptian culture using several methods for calculating reliability (he average correlations between items, internal consistency values (Cronbach's alpha), and standard error measurement (SEM) values), and also for calculating validity (differences between the mean scores of the T scores for males and females and Known Groups Validity). However, MMPI-2 RF more studies are needed in Egyptian culture, using larger clinical groups.
References

Ahmed, S. (2019). Efficiency of the revised Minnesota Multiphasic Personality Inventory. Arts and Humanities Series. 3 (6), 27-45 (in Arabic).


The Psychometric Competence of Minnesota Multiphasic Personality Inventory–
.2 Reconstructed (MMPI–2–RF) in Egyptian Sample


The Psychometric Competence of Minnesota Multiphasic Personality Inventory–
.2 Reconstructed (MMPI–2–RF) in Egyptian Sample


Tarescavage, A. M., Cappo, B. M., Ben-Porath, Y. S. (2018). Assessment of sex offenders with the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF). Sex Abuse, 30,413–437.

