

Impact of Using Raw versus T Scores in Minnesota Multiphasic Personality Inventory-2 Restructured Form Descriptive and Inferential Research

Jayne Luna and Katrina Conen, California State University, Monterey Bay
Supervised by Dr. Danielle Burchett, California State University, Monterey Bay
In Collaboration with Dr. David Glassmire, Patton State Hospital
and Dr. Anthony Tarescavage, Kent State University

Mental health clinicians use psychological assessment tools in order to determine client mental status and symptom severity. The Minnesota Multiphasic Personality Inventory-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011) is a self-report inventory that includes 51 scales used to assess for personality and psychopathology dysfunction as well as invalid response styles. Raw MMPI-2-RF scores are unaltered counts of the number of items an individual endorsed for a particular scale (e.g., a scale measuring depressive symptoms or anxiety symptoms). For many widely-used published tests, raw scores are converted to standardized scores to facilitate interpretation of assessment results. For instance, many widely used Intelligence Quotient (IQ) tests typically utilize a type of standardized score wherein 100 is representative of an average score in the general population and 115 represents a score that is one standard deviation higher than average ($M = 100, SD = 15$) (Groth-Marnat, 2009). Standardized scores help examiners understand how different an examinee's characteristics are, compared to the average person.

Translating Raw Scores to Uniform T Scores. Uniform T scoring is an alternative standardized scoring approach used to translate raw scores into standardized scores (Graham, 2012). The MMPI-2-RF Uniform T Score distribution is somewhat positively skewed—since experiencing a large number of psychiatric symptoms is relatively rare in the general population—with a mean of 50 and standard deviation of 10. Scores are anchored to the MMPI-2-RF normative sample, which was designed to be representative of the adult U.S. Census (Ben-Porath & Tellegen, 2008/2011). Thus, a score of 50T is consistent with the average self-reported

symptomatology of a person in the normative sample, whereas a score of 70T indicates the person reported psychiatric symptoms at a severity level that is quite rare (two standard deviations higher than the average person in the general population). Because they were designed to be uniform, Uniform T scores across scales measuring different types of symptoms can be interpreted to be at equivalent severity levels (Graham, 2012). Across scales, scores of 65T or higher are considered to be of clinical significance, indicating someone reported a relatively rare number of psychiatric symptoms. Clinicians use T scores to interpret test results, as they provide a standardized approach to understanding the extremity of self-reported symptoms as compared to the general population (Ben-Porath, 2012; Ben-Porath & Tellegen, 2008/2011; Graham, 2012). Despite clinicians interpreting Uniform T scores, many MMPI-2-RF researchers conduct analyses using the original raw scores (Tellegen & Ben-Porath, 2008/2011).

Current Study

Given the disconnect between the scale versions used by researchers and clinicians, we were interested in comparing the psychometric properties of MMPI-2-RF substantive scale raw and Uniform T scores within a sample of psychiatric forensic inpatients who completed the measure in a valid manner. To our knowledge, the impact of using raw versus Uniform T scores in MMPI-2-RF research has not been independently empirically examined. Given the change in distribution properties inherent in standardizing raw into T scores, we hypothesized that there would be meaningful differences in score distributions (e.g., skewness, kurtosis), scale intercorrelations, and correlations with relevant psychiatric diagnostic criteria, with raw scores exhibiting a more normal distribution and stronger correlations, as compared to Uniform T scores.

Method

Participants

We utilized data from a deidentified archival sample of 1,110 psychiatric inpatients at a large psychiatric state hospital who completed the MMPI-2-RF during the course of clinical or forensic evaluations. Of the original sample, 764 patients met criteria for having completed the test in a valid manner (using a standard procedure for identifying invalid protocols consistent with Ben-Porath & Tellegen, 2008/2011). In the final sample of 764 inpatients, 72.5% were male. Of the 694 with available ethnicity data, 58.2% identified as Caucasian, 23.6% as African American, 13.7% as Hispanic/Latino, 2.4% as Asian American, and 2.0% identified as being of another race. Their mean age was 40.63 years ($SD = 11.40$).

Measure and Diagnostic Variables

Minnesota Multiphasic Personality Inventory-2 Restructured Form. We used the MMPI-2 Restructured Form (MMPI-2-RF; Ben-Porath & Tellegen, 2008/2011), a 338-item self-report measure with 51 scales measuring personality constructs, protocol validity, and psychopathology. Extensive analyses on the reliability and validity of MMPI-2-RF scales are available in the technical manual (Tellegen & Ben-Porath, 2008/2011). MMPI-2-RF internalizing scales—designed to measure depressive, anxious, and phobic symptoms—include Emotional/Internalizing Dysfunction (EID), Demoralization (RCd), Low Positive Emotions (RC2), and Dysfunctional Negative Emotions (RC7). MMPI-2-RF thought dysfunction scales—designed to measure severe psychiatric conditions including schizophrenia spectrum disorders and psychotic symptoms—include Thought Dysfunction (THD), Ideas of Persecution (RC6), and Aberrant Experiences (RC8). MMPI-2-RF externalizing scales—designed to measure impulsivity, antisocial behavior, and substance abuse—included Behavioral/Externalizing

Dysfunction (BXD), Antisocial Behavior (RC4), and Hypomanic Activation (RC9) (Romero, Toorabally, Burchett, Tarescavage, & Glassmire, 2016).

Psychiatric Diagnoses. Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association [APA], 2000) psychiatric diagnoses from the date of testing (uninfluenced by test results) were recorded. Using a diagnostic coding scheme developed by Romero et al. (2016) to align with modern models of psychopathology (e.g., Kreuger, 1998), we used three dichotomized diagnostic variables to examine presence of (1) internalizing dysfunction (e.g., depression, anxiety), (2) thought dysfunction (e.g., schizophrenia, manic episode), and (3) externalizing dysfunction (e.g., substance use disorder, antisocial personality disorder) diagnoses.

Procedure

IRB approval was obtained from the State of California Committee for the Protection of Human Subjects (CPHS) by our collaborator, David Glassmire. Our mentor, Danielle Burchett, used a de-identified dataset to conduct analyses examining patient archival test scores, demographics, and clinical diagnosis. We interpreted results to compare Higher-Order and Restructured Clinical Scale raw and Uniform T score means, standard deviations, kurtosis values, skewness values, and clinical diagnoses. We also examined raw and Uniform T score scale intercorrelations and associations between scales and the presence of conceptually relevant psychiatric disorders (i.e., internalizing disorder, thought disorder, externalizing disorder).

Results

When raw and Uniform T scale intercorrelations for the same scale (e.g., correlating EID raw scores with EID Uniform T scores) were examined, we found that raw and Uniform T scores were extremely highly associated (r s ranged from 0.97 to 1.00). Despite their similarities, an

examination of raw and Uniform T score means, standard deviations, skewness values, kurtosis values, and their associations with relevant clinical diagnostic variables indicated there were some notable differences between raw and Uniform T scores in spread, kurtosis, and skewness (Table 1). The raw score data were slightly more skewed and leptokurtic than the Uniform T score data for internalizing and thought dysfunction scales, but the pattern for externalizing scales was less clear in regards to skewness. Externalizing scale raw scores were slightly more platykurtic, as compared to Uniform T scores. Point-biserial correlations between scores and clinical diagnosis were similar (Table 1). Of note, four of twelve scale intercorrelations were slightly stronger for raw than for Uniform T scores (Table 2).

Discussion

We investigated the impact of using raw versus Uniform T scores on MMPI-2-RF scale psychometric properties. Many raw scores had slightly higher skewness and kurtosis values, indicating greater non-normality of distributions. Associations with extra-test diagnostic criteria were similar, and scale intercorrelations were only very slightly stronger for raw as compared to T scores. These modest differences suggest that researchers should consider using Uniform T scores rather than raw scores, but that literature using raw scores is likely to meaningfully generalize to clinical settings where Uniform T scores are interpreted. Future studies should examine whether the results replicate across settings and with a wider variety of external criteria.

Appendix: References, Funding Disclosure, and Tables

- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*, Fourth Edition, Text Revision. Washington, DC.
- Ben-Porath, Y. S. (2003). Assessing personality and psychopathology with self-report inventories. In J. Graham & J. Naglieri (Eds.), *Handbook of psychology, Volume 10: Assessment psychology* (pp. 553-577). Hoboken, NJ: John Wiley & Sons, Inc.
- Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF*. Minneapolis: University of Minnesota Press.
- Ben-Porath, Y. S., & Tellegen, A. (2008/2011). *MMPI-2-RF: Manual for administration, scoring, and interpretation*. Minneapolis: University of Minnesota Press.
- Graham, J. R. (2012). *MMPI-2: Assessing personality and psychopathology* (5th ed.). New York: Oxford University Press.
- Kreuger, R.F., Caspi, A., Moffitt, T. E., & Silva, P. A. (1998). The structure and stability of common mental disorders (DSM-III-R): A longitudinal-epidemiological study. *Journal of Abnormal Psychology, 107*(2), 216-227.
- Romero, I. E., Toorabally, N., Burchett, D., Tarescavage, A. M., & Glassmire, D. M. (2016). Mapping the MMPI-2-RF substantive scales onto internalizing, externalizing, and thought dysfunction dimensions in a forensic inpatient setting. *Journal of Personality Assessment*. Advance online publication. DOI: 10.1080/00223891.2016.1223681
- Tellegen, A., & Ben-Porath, Y. S. (2008/2011). *MMPI-2-RF technical manual*. Minneapolis: University of Minnesota Press.

Grant Funding

This research project was made possible by a grant from the University of Minnesota Press, which supported data collection and from the California State University, Monterey Bay Undergraduate Research Opportunities Center (UROC) which provided financial and logistical support.

Table 1

Descriptive results for MMPI-2-RF Higher-Order and Restructured Clinical Scale raw and rounded truncated Uniform T scores in a forensic inpatient sample ($n = 764$)

	<i>M</i>	<i>SD</i>	Min	Max	Skewness	Kurtosis	r_{pb} w/ dx.
Internalizing Dysfunction Scales							
EID raw	10.33	8.02	0	38	1.03	0.46	.19
EID T	49.84	11.93	30T	89T	0.71	0.16	.18
RCd raw	5.83	5.72	0	23	1.00	0.00	.17
RCd T	52.08	11.48	37	85	0.62	-0.29	.17
RC2 raw	4.35	3.28	0	17	0.98	0.81	.11
RC2 T	51.11	12.61	34	99	0.91	0.68	.11
RC7 raw	5.40	4.81	0	21	0.92	0.07	.18
RC7 T	48.05	11.04	34	86	0.88	0.41	.18
Thought Dysfunction Scales							
THD raw	3.53	3.83	0	20	1.52	0.09	.07
THD T	56.51	14.72	39T	100T	0.95	0.09	.08
RC6 raw	2.48	2.89	0	16	1.46	1.80	.03
RC6 T	60.39	15.64	43	100	0.71	-0.20	.04
RC8 raw	2.96	3.02	0	15	1.28	1.30	.06
RC8 T	53.82	12.21	39	96	0.75	0.18	.07
Externalizing Dysfunction Scales							
BXD raw	8.20	4.35	0	22	0.38	-0.38	.20
BXD T	55.69	11.06	32T	92T	0.38	-0.20	.21
RC4 raw	7.78	4.29	0	20	0.37	-0.53	.21
RC4 T	58.95	11.94	34	93	0.34	-0.42	.21
RC9 raw	9.79	5.37	0	27	0.48	-0.43	.06
RC9 T	46.55	10.79	25	88	0.75	0.52	.06

Note. r_{pb} w/ dx. (point-biserial correlation with diagnostic category); internalizing dysfunction scales were associated with presence of internalizing diagnosis; thought dysfunction scales were associated with presence of thought dysfunction diagnosis; externalizing scales were associated with presence of externalizing diagnosis). See text for full scale names.

Table 2

MMPI-2-RF scale raw and T score intercorrelations in a forensic inpatient sample ($n = 764$)

	Internalizing Scales			Thought Dysfunction Scales			Externalizing Dysfunction Scales			
	EID	RCd	RC2	RC7	THD	RC6	RC8	BXD	RC4	RC9
EID	---	.91**	.70**	.75**	.45**	.44**	.46**	.33**	.42**	.31**
RCd	.92**	---	.53**	.74**	.49**	.46**	.52**	.39**	.48**	.39**
RC2	.70**	.55**	---	.25**	.14**	.17**	.12**	-0.01	.15**	-.19**
RC7	.75**	.74**	.25**	---	.59**	.54**	.62**	.48**	.47**	.62**
THD	.42**	.46**	.13**	.58**	---	.85**	.87**	.31**	.29**	.51**
RC6	.42**	.44**	.18**	.52**	.88**	---	.62**	.24**	.23**	.43**
RC8	.45**	.51**	.12**	.63**	.87**	.65**	---	.41**	.38**	.57**
BXD	.32**	.38**	-0.009	.48**	.30**	.23**	.40**	---	.88**	.70**
RC4	.42**	.47**	.15**	.47**	.28**	.22**	.37**	.88**	---	.49**
RC9	.30**	.38**	-.18**	.61**	.49**	.42**	.56**	.70**	.49**	---

Note. * $p < .05$; ** $p < .01$. Raw score intercorrelations are presented below the diagonal. Uniform T score intercorrelations are presented above the diagonal. Grey areas denote within-domain intercorrelations. See text for full scale names. Shading indicates correlations in the same domain of psychopathology.