MCMI-III Critique

From: law and psychology discussion list [mailto:PSYLAW-L@listserv.unl.edu] **On Behalf Of** Tyler Carpenter
Sent: 31 May, 2008 11:52
To: PSYLAW-L@listserv.unl.edu
Subject: Re: MCMI III

This is great, Barry. Is it all right for me to forward it to a forensic colleague? TIA.
Best,
Tyler

J. Tyler Carpenter, Ph.D., FAACP
PO Box 366063
Hyde Park, MA 02136-0019

Date: Fri, 30 May 2008 15:24:57 -0700
From: barryt@dslextreme.com
Subject: Re: MCMI III
To: PSYLAW-L@listserv.unl.edu

Here is my write-up from a recent case I did in which the opposing expert relied upon the MCMI-III:

There are five definitive articles on the forensic validity of the MCMI. Two are supportive, i.e., Dyer, Frank and Joseph McCann. "The Millon Clinical Inventories, Research Critical of Their Forensic Application, and Daubert Criteria." Law and Human Behavior 24.4 (2000): 487-97 and Retzlaff, Paul. "Comment on the Validity of the MCMI-III." Law and Human Behavior 24.4 (2000): 499-500, and three hold that the MCMI-III should not be used in forensic evaluations, in Hynan, Daniel J., "Unsupported Gender Differences on Some Personality Disorder Scales of the Millon Clinical Multiaxial Inventory-III." Professional Psychology: Research and Practice 35.1 (2004): 105-10.; Rogers, Richard, Randall Salekin, and Kenneth Sewell. "The MCMI-III and the Daubert Standard: Separating Rhetoric from Reality." Law and Human Behavior 24.4 (2000): 501-06.; and Rogers, Richard, Randall T. Salekin, and Kenneth W. Sewell. "Validation of the Millon Clinical Multiaxial Inventory for Axis II disorders: Does it meet the Daubert standard?" Law and Human Behavior 23.4 (1999): 425-43.

The most definitive findings for lack of forensic validity for the MCMI-III are found in Rogers et al (2000) and Hynan (2004). As Rogers (2000) notes:

> "[The] MCMI, MCMI-II, and MCMI-III are not [progressive] refinements, but fundamentally different measures, based on integral changes in items, scale composition, and scoring. As separate measures, extrapolations from the MCMI-III to the MCMI-III are inappropriate. Second, they [Dryer and McCann] do not quarrel with the Rogers et al. [1999] analysis of Daubert and its general applicability to MCMI-II and MCMI-III. Third, they allow to let stand specific limitations on MCMI-III content validity. Namely, the MCMI-III includes no mechanism to assess (a) the chronicity of symptoms from early adulthood or (b) distress or impairment arising from particular symptoms.
>
> > In evaluating construct validity, Rogers et al. (1999) simply asked that the convergent validities exceed the discriminant validities. Unfortunately, the MCMI-III did not meet this basic standard. With heavy reliance on Wiggins (1973), Dyer and McCann express bitter disappointment over the adequacy of other Axis II measures and the lack of method variance in the Rogers et al. (1999) meta-analysis.

> Unwittingly, their criticisms represent a devastating attack on the MCMI-III itself. Both the MCMI-III and the meta-analysis employ clinical interviews, the MMPI-2, and self-report measures. However, the meta-analysis goes beyond the MCMI-III to include MMPI-2 personality disorder scales, structured Axis II interviews (e.g., the PDE, SIDP, and SCID-II), and an Axis II questionnaire (i.e., the PDQ-R). If the meta-analysis is inadequate because of method variance, what can we conclude about the MCMI-III?

Dyer and McCann (2000) attempt to discount negative findings by finding fault with all the convergent measures utilized by independent researchers because the PPP [positive predictive power….of the MCMI-III is too low]. Dyer and McCann (2000) asserted that the content validity of DSM-IV Axis II symptoms was adequately described in Millon et al. (1997, pp. 24–26). While it is true that expert ratings were described on pages 24–26, these ratings are irrelevant to the discussion at hand. As noted, these ratings and procedures identified ''1,132 best-fit items'' (p. 26). These items were apparently ''searched'' by the senior author to ''adequately capture the official criteria'' (p. 26). This unstandardized method lacks the minimal scientific rigor for establishing content validity.

Dyer and McCann (2000) defend the content validity of the MCMI-III by asserting that Millon's membership on the DSM-IV task force for personality disorders accords him a special status in establishing content validity. Most members of the scientific community are likely to be unimpressed by this form of ipse dixits that is also evident in Millon et al. (1997).

Dyer and McCann (2000) tout the value of the MCMI-II for assessing response styles (i.e., malingering and defensiveness). Their silence on the MCMI-III is understandable in the absence of empirical research. The MCMI-III normative data (Millon et al., 1997) are very troubling for the Scale Z (Debasement). Given the apparently low prevalence of malingering in the normative group, the high correlations between Scale Z and clinical scales suggest serious confounds in the test development. Markedly high correlations (r's .80) are found on Scales C, D, SS, and CC; high correlations (r's .70) are observed for Scales 2B, 8A, 8B, S, A, H, and R (Millon et al., 1997, p. 66).

Dyer and McCann (2000) contended that the MCMI-III can be used in forensic practice and scoff at the notion that ''its applicability to forensic issues remains virtually untested'' (Rogers et al., 1999, p. 439).

The key questions remain: Has the MCMI-III been tested with forensic populations? Does it produce reliable differences that are useful in addressing forensic issues? Dyer and McCann (2000) cite a total of six studies to vouch for the MCMI-III and its applicability to forensic practice. These studies are provided as evidence that ''the MCMI-III provides very useful data that can inform consideration of forensically related issues'' (Dyer & McCann, 2000). Is this an accurate reflection of these studies? For example, Craig and Olson (1997) utilize non-forensic samples to compare combat-related PTSD to substance abuse. Craig (1997, p. 1388) studied what he termed ''primary substance Misusers'' in treatment at a VA medical center.

The citation of Flynn and McMahon (1997) to address ''forensically related issues'' with the MCMI-III is even more baffling. The chapter provides a summary of substance abuse on the MCMI and MCMI-II. Not only does the chapter not address forensic issues, it does not even consider MCMI-III research beyond Millon (1994) because ''studies using the MCMI-III substance dependence scales have not yet appeared in the literature'' (Flynn & McMahon, 1997, p. 181). Similarly, Hyer, Brandsma, and Boyd (1997) is a non-forensic discussion of PTSD based almost entirely on the MCMI and the MCMI-II. Despite their protestations, Dyer and McCann (2000) were unable to marshal sufficient evidence supporting the use of the MCMI-III in evaluating forensic issues. Readers of this journal are ill-served by any attempts to obfuscate this critical point.

An oversight of Rogers et al. (1999) was the slight modifications in the MCMI-III test manual (Millon et al., 1997). The revised edition reported the identical MCMI-III scale intercorrelations (see Table 3.6, Millon et al., 1997; and Appendix G, Millon, 1994).

Therefore, no changes in the conclusions regarding construct validity are indicated. For establishing diagnostic efficiency, however, they entirely eliminated the original sample of 998 patients and substituted a comparatively smaller sample of 321 cases (see Table 4.1, Millon et al., 1997). This new sample is constrained in its representation of primary Axis II disorders. In particular, Cluster A diagnoses have very modest groups (M 13.7 patients per diagnosis).

For DSM-IV diagnoses, Millon et al. (1997) reported an impressive improvement over the original validation study. In the original study, the mean PPP was .18. In the new study, the average PPP soared to .70, an increase in accuracy of 388.9%. We were originally puzzled how a similar study would achieve this dramatic increase. A close inspection of the methodology provides two compelling explanations. First, the original study (see Appendix J-2, Millon, 1994) used the standard criterion of BR scores 85 for establishing MCMI-III disorders. The revised study did not employ any standard criterion for BR scores. Instead, the concordance with clinician-based diagnoses was simply achieved by calculating the "percentage of patients having a particular MCMI-III personality scale as the highest in their profile" (Millon, 1997, p. 98). This lax criterion is unacceptable. Conceivably, it would allow a BR score below 50 to be counted as a "correct" diagnosis as long as other Axis II scores were lower. Second, the original study masked psychologists to the MCMI-III test interpretation, while the revised study allowed criterion contamination."

Barry

Barry T. Hirsch, Ph. D.
2001 South Barrington Avenue, Ste 202
Los Angeles, California 90025-5385

e: barryt@dslextreme.com
phone: 310-444-1439
fax: 310-473-4633

From: law and psychology discussion list [mailto:PSYLAW-L@listserv.unl.edu] **On Behalf Of** Allan Cooperstein
Sent: Friday, May 30, 2008 8:38 AM
To: PSYLAW-L@listserv.unl.edu
Subject: [PsyLaw-L] MCMI III

Does anyone have references on the shortcomings of the MCMI III as the only test given in a forensic evaluation?

Allan
M. Allan Cooperstein, Ph.D.
Licensed Psychologist, Clinical
(215) 546 6808
allanco@verizon.net
www.allanpsych.com

From: law and psychology discussion list [mailto:PSYLAW-L@listserv.unl.edu] **On Behalf Of** David Clark
Sent: 31 May, 2008 07:29
To: PSYLAW-L@listserv.unl.edu

Subject: Re: MCMI III

Since I've decided from his many erudite posts that Barry probably is the smartest person I've never met, I'm especially thrilled to point out that there is another very important critique of the MCMI-III he did not mention.

Halon, R L. (2001). The Millon Multiaxial Clinical Inventory III: The normal quartet in child custody cases. American Journal of Forensic Psychology, 19, 57-75.

Halon does a step-by-step analysis to show how high but meaningless scale scores on MCMI-III Histrionic, Narcissistic and Compulsive scales are simply hard-wired into the test when non-patients complete the test in non-clinical settings where they have reason to respond in a socially desirable fashion (like custody evaluations). I stopped using the MCMI-III immediately after I read the article.

David

David B. Clark, Ph.D. Clinical Psychologist 16216 Baxter Road, Suite 323 Chesterfield, MO 63017 tel: 636-537-8222 fax: 636-537-8223 dbc_phd@msn.com