

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

**Assessing Negative Response Bias with the Millon Clinical Multiaxial Inventory-IV (MCMI-IV): a Review of the Literature**

**This is the author's manuscript**

*Original Citation:*

*Availability:*

This version is available <http://hdl.handle.net/2318/1844069> since 2024-12-12T10:31:22Z

*Published version:*

DOI:10.1007/s12207-022-09442-4

*Terms of use:*

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

**Psychological Injury and Law (2022) 15:48–55****<https://doi.org/10.1007/s12207-022-09442-4>****James P. Choca & Claudia Pignolo****Assessing Negative Response Bias with the Millon Clinical Multiaxial Inventory–IV****(MCMI-IV): A Review of the Literature**

**Abstract**

The most current version of the Millon Clinical Multiaxial Inventory, the MCMI-IV, has practically no literature outside of the manual. The research on the embedded indicators of negative response bias is scarce at this time with regard to the MCMI-IV; however, these scales largely overlap with those of the MCMI-III, so that it is reasonable to hypothesized that a similar validity will generalize to the newer version. The present article reviews data from the manual and the scientific literature to offer a review of the simulation and known-group comparison studies conducted with the MCMI-III. The use of the Modifier Indices is discussed, as well as the scale patterns expected with records suggesting a negative response bias. As a conclusion, we offer a list of recommendation for future studies with the MCMI-IV.

*Keywords:* Negative response bias; MCMI; feigning

Outside of what is included in the test manual, there are only a few studies using the current version of the Millon Clinical Multiaxial Inventory (MCMI-IV; Millon et al., 2015). Of the 195 items of the MCMI-IV, however, 120 or 70% were taken from the MCMI-III (Choca, 2021). Not surprisingly, with the exception of four scales, the correlations of the respective scales of the two versions are above .75, according to the manual. This paper will review what is known about negative response bias on the MCMI-III (Millon et al., 2009) on the assumption that the findings apply to the current version as well. The MCMI-IV validity scales mostly date back to the MCMI-II (Millon, 1982) so, in some cases, we have also included studies done with that version.

### **Background and Conditions of Use**

The MCMI (Millon, 1979) was originally developed by Theodore Millon to reflect the personality theory he first described in 1969 (Millon, 1969). The important role that Millon played in the development of the third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III; American Psychiatric Association, 1980) contributed to the compatibility that the original MCMI had with the DSM-III. Since that time Millon's theory became increasingly more complex and strayed away from the DSM. Nevertheless, the MCMI-IV resembles the current DSM (DSM-5; American Psychiatric Association, 2013) more than most of the other broad-based questionnaires. Both the MCMI and the DSM are profile-based systems, systems where the attributes of an entity are described, and the clinician has to decide if the individual demonstrates enough of those attributes to merit being classified under that profile. Besides, ten of the personality profiles of the MCMI-IV, have relatively similar equivalents in the DSM-5 (i.e., schizoid, avoidant, dependent, histrionic, narcissistic, antisocial, compulsive,

## RUNNING HEAD: MCMI-IV and Negative Response Bias

schizotypal, borderline, and paranoid). In addition to these personality profiles, the MCMI-IV has five non-DSM personality scales (Melancholic, Negativistic, Turbulent, Sadistic, and Masochistic).

In spite of the similarities, a common issue with the MCMI is that the personality scales do not measure a personality disorder but a personality style. Millon's theory addresses personality styles and very few of the MCMI personality items are aimed at evaluating the examinee's personality dysfunction. In fact, some of the MCMI personality scales (e.g., the Histrionic, the Narcissistic, and the Compulsive Scales for the MCMI-III, and now the Turbulent Scale for the MCMI-IV) have been repeatedly shown to be signs of good health (Eastin, 2021; see Choca for more details, 2004).

The MCMI-IV consists of 12 Clinical Personality Pattern Scales, three Severe Personality Pathology scales,<sup>45</sup> Grossman Facet scales developed to highlight aspects of the different personality styles, seven Clinical Syndrome scales such as depression and drug abuse, and three Severe Clinical Syndrome scales. It contains two Random Response Indicators, (i.e., Invalidity [Scales V] and Inconsistency [Scale W]), and three Modifying Indices (i.e., Disclosure [Scales X], Desirability [Scale Y], and Debasement [Scale Z]) to evaluate whether the examinee was defensive and secretive, attempted to appear socially attractive, or more pathological, respectively. This self-report instrument has 195 true-false items, normed with 1,547 mostly outpatient psychiatric participants, in a sample said to resemble the population of people seeking therapeutic treatment in the United States. . It is the shortest of the available broad-based questionnaires, a feat accomplished by having a great number of items contributing to more than one scale. That system unfortunately led to high inter-scale correlations. Counting the Grossman Facet scales, the instrument has 71 scales, or 2.75 items per scale. According to the manual,

## RUNNING HEAD: MCMI-IV and Negative Response Bias

Cronbach's  $\alpha$  internal consistency average is .83; the test has reasonable convergence with parallel scales of the second edition of the Minnesota Multiphasic Personality Inventory (MMPI-2; Butcher et al., 1989) and the Brief Symptom Inventory (BSI; Derogatis, 1993). It can be used to measure emotional problems in a variety of places including clinics and correctional settings. The inventory works best with educated individuals with psychological sophistication who are open to revealing their self-perceptions.

A common problem is a lack of understanding about the weighted score that is used, the Base Rate Score (BRS). As will be seen, Millon's BRS is his own invention and, although presumably associated to the prevalence of the attribute, is not the same as the base rate. Millon noted the inaccuracy inherent in the use of a cutoff  $t$  score with questionnaires, typically set at 1.5 standard deviations above and below the mean, a cutoff that flags 7% of the population at each end. In the case of a common emotional problem, such as depression or anxiety, that cutoff may be too low; in the case of a severe disorder, such as schizophrenia, the cutoff may be too high. Millon's answer was to create a new measure. Elevations of Millon's BRS reflect the prevalence of the attribute under consideration. Thus, a BRS of 75 was arbitrarily set at the point where the attribute is present; a BRS of 85 was set at the point where the attribute is predominant. Technically, low BR scores are meaningless. The prevalence of the different profiles used was presumably what Millon found in the sample that was collected for the questionnaire. The procedure he followed to determine the prevalence or predominance was never made clear. Using current prevalence data available for some disorders would make the BR system more scientific and useful. The lack of transparency, however, was especially bothersome with regards to the Grossman Facet scales. How could the prevalence of the "temperamentally apathetic" or the "interpersonally paradoxical" be possibly established? The

## RUNNING HEAD: MCMI-IV and Negative Response Bias

measure he created has many problems, including the usual need for the clinician to consider whether the population being tested differs from the normative sample. Any of these considerations could of course become an issue in forensic work.

Although Millon used BR scores for the Modifier Indices or validity indicators, those scores were calculated differently from the clinical scales, and are more like *t* scores. Consider, for instance, the Disclosure Scale (Scale X). The raw score of that scale was mostly computed by adding the raw scores of the personality scales; every examinee possesses personality attributes, and the summation of all personality raw scores can be taken as an indication of how readily the examinee admits to having such attributes. As a measure of disclosure, that summation has a normal distribution: an individual at the high end is recognizing having many attributes while, at the low end, the examinee is recognizing very few attributes. Thus, low Disclosure Scale scores are not only interpretable but can be very meaningful, typically indicating a reticence to admit psychological attributes. The Desirability (Scale Y) and Debasement (Scale Z) scales were developed by asking college students to take the inventory while attempting to look their best or their worst respectively. Items that were most often endorsed in those experimentally induced distortions in self-report were then grouped to form normally distributed scales. As was the case with the Disclosure (X) Scale, extreme scores at both ends can then be interpreted.

A strong relationship has been reported between the MCMI-III personality scales and Modifier Indices, and the MMPI-2 clinical and validity scales (Morgan et al., 2002; Schoenberg et al., 2004). Oppositely, the Modifier Indices showed no relationship with performance validity tests (PVTs) such as the Test of Memory Malingering (TOMM; Tombaugh, 1996) and Reliable Digit Span (RDS; Greiffenstein et al., 1994), as reported by Ruocco et al. (2008) analyzing records of 105 patients referred for neuropsychological evaluations. One has to keep in mind that

## RUNNING HEAD: MCMI-IV and Negative Response Bias

these tests are different in nature: the SVTs evaluate the accuracy of self-reported problems, whereas the PVTs assess the validity and effort of task performances (Larrabee, 2012). While some studies have shown that there was not a relationship between SVTs and PVTs (Demakis et al., 2008; Nelson et al., 2007; Van Dyke et al., 2013), others have reported overlapping areas between the two categories (Gaasedelen et al., 2019; Gervais et al., 2007; Giromini et al., 2020; Grossi et al., 2017; Larrabee et al., 2017; Sabelli et al., 2021; Sumanti et al., 2006). As such, the relationship between SVTs and PVTs is still an open debate.

Millon used extremes scores obtained on the Modifier Indices to increase or decrease the BRS of clinical scales. Those alterations are not done in a clear and scientific manner and have been the source of criticism. Strack (1991), for instance, noted that the system is elaborate and confusing and worried about scales that are corrected up to four or five times; the fact that one cannot ascertain how the specific correction weights were derived, nor evaluate their clinical utility also concerned him. Nevertheless, the MCMI-IV manual provides empirical support for its validity (Millon et al., 2015; for an overview of the MCMI-III validity studies see Craig, 1999).

### **Response Bias**

According to the MCMI-IV Manual (Millon et al., 2015), a cut-off of BRS  $\geq 75$  on the Debasement (Z) scale would indicate the potential presence of negative response biases. Also the Disclosure (X) scale could help in identifying response bias: BRS between 61 and 114 and BRS  $\geq 115$  suggest that a negative distortion is possible and likely, respectively. As mentioned before, to date no study has evaluated the classification accuracy of the MCMI-IV Modifying Indices in detecting feigning of psychopathology. Therefore, in this section we have reviewed the published studies on the Modifier indices of the MCMI-III. For a meta-analysis of the results reviewed in this section, see Boccaccini and Hart (2018).



## RUNNING HEAD: MCMI-IV and Negative Response Bias

Daubert and Metzler (2000) administered the MCMI-III twice to 160 psychiatric outpatients (primary diagnosis of mood disorders or schizophrenia) within a simulation study design. Using a counterbalanced design, some patients were asked to fake good ( $n = 80$ ) or to fake bad ( $n = 80$ ) in a credible manner for one of the administrations, while taking it truthfully for the other. Participants did not receive any incentive for effortful simulation. Paired sample  $t$ -tests showed significant differences between the two administrations, with moderate effect size values ( $d = 1.13$  for Disclosure;  $d = 0.83$  for Desirability;  $d = 1.12$  for Debasement; see Rogers et al., 2003 for effect size characterizations in simulation studies). First, they evaluated the classification accuracy of the standard cut-off, which produced the following results: for Disclosure (X)  $BRS \geq 85$   $Se = 61\%$ ,  $Sp = 81\%$ , diagnostic power (DP) = 71%; for Desirability (Y)  $BRS \leq 35$   $Se = 58\%$ ,  $Sp = 76\%$ , DP = 67%; for Debasement (Z)  $BRS \geq 85$   $Se = 55\%$ ,  $Sp = 79\%$ , DP = 67%. Next, Daubert and Metzler (2000) suggested the best cut-off for each scale, choosing those that maximized the overall diagnostic power, which produced  $Se = 76\%$ ,  $Sp = 71\%$ , DP = 74% for Disclosure (X)  $BRS \geq 80$ ;  $Se = 64\%$ ,  $Sp = 78\%$ , DP = 69% for Desirability (Y)  $BRS \leq 39$ , and  $Se = 64\%$ ,  $Sp = 78\%$ , DP = 71% for Debasement (Z)  $BRS \geq 81$ . All in all, they concluded that the Modifier Indices were moderately effective in discerning response bias.

Evaluation of the MCMI-III on the detection of response bias can also be found in the work of Schoenberg and his colleagues. In their first study, Schoenberg et al. (2003, 2006) adopted a clinical comparison, simulation study design, comparing MCMI-III scores of 112 college students who were asked to take the test as a patient experiencing severe emotional and psychological problems and 111 college students and 181 psychiatric inpatients who took the test under standard instructions. Experimental feigners were given a list of symptoms, a vignette, a monetary incentive (\$200 lottery), and a caution about the fact that the test had validity scales

## RUNNING HEAD: MCMI-IV and Negative Response Bias

developed to identify exaggeration and random responding. They found significant differences between the fake-bad group and both the honest student ( $d = 2.59$  for Disclosure,  $d = -2.01$  for Desirability,  $d = 2.17$  for Debasement) and honest patient ( $d = 0.81$  for Disclosure,  $d = -0.66$  for Desirability,  $d = 0.59$  for Debasement) groups on the three Modifier Indices. The authors offered optimal cut-off scores that correctly classified the greatest number of true positives and negatives, and did not misclassify more than 50% of the experimental feigners, which are as follows: Disclosure (X)  $BRS \geq 85$ , Desirability (Y)  $BRS \leq 25$ , and Debasement (Z)  $BRS \geq 81$  that produced hit rate values of 65.2%, 64.8%, and 63.4%, respectively. Finally, ROC analyses concerning experimental feigners and psychiatric inpatients showed that Scale X (AUC = .719, SE = .030) best classified the experimental feigners, followed by Scale Y (AUC = .685, SE = .032) and Scale Z (AUC = .661, SE = .032). These researchers found the Modifier Indices to be of minimal clinical utility in discriminating between students feigning psychopathology and bona fide psychiatric inpatients. However, comparing all the MCMI-III scale scores, the experimental feigners produced greater scores on 12 MCMI-III scales. Schoenberg and colleagues (2006) went on to develop two discriminant analysis functions, using a multitude of MCMI scores, that were much more effective, detecting as much as 78% of the target cases. Function A used the MCMI-III scales as dichotomous level data with  $BRS > 84$  for Scales 8A Negativistic, C Borderline, P Paranoid, SS Thought disorder, PP Delusional disorder, along with continuous Scale X BR scores. Function B used the Scales X Disclosure, Y Desirability, 7 Compulsivity, PP Delusional disorder as continuous variables. It should be noted that the Debasement (Z) scale was excluded from the two functions. All in all, Schoenberg et al. (2003, 2006) found that using an algorithm that take into account the combination of multiple scales to classify experimental feigners from patients worked better than the single Modifier Indices.

## RUNNING HEAD: MCMI-IV and Negative Response Bias

Aguerrevere and his coworkers (2011) evaluated the accuracy of the MCMI-III to detect feigning with 107 traumatic brain injury (TBI) patients, who were divided into three groups. The partition of the groups was based on the Slick, Sherman, and Iverson (1999) criteria for malingered neurocognitive dysfunction (MND): MDN ( $n = 55$ ), not-MND ( $n = 26$ ), and Indeterminate ( $n = 26$ ). The authors performed a multivariate analysis of variance (MANOVA) to examine whether group belonging affected the MCMI-III Modifier Indices. They found a significant effect so that the MDN group scored higher on the Disclosure (X) and Debasement (Z) scales and lower on the Desirability (Y) scale compared to the other two groups, with moderate to large effect size values. Using receiver characteristic curve analysis at a 4% false positive rate, sensitivity was 47% for Disclosure (X), 51% for Desirability (Y), and 55% for Debasement (Z). Taking all three indices into account, the MCMI-III showed 54% sensitivity at cutoffs associated with 0% false positive error rate. Interestingly, Aguerrevere et al. (2011) found that the Debasement (Z) scale was the most sensitive of the three Modifier Indices. It should be noted that this is a criterion group design study evaluating the classification accuracy of the MCMI-III indices in a sample of real-life malingerers.

There are some limitations associated to the four reviewed studies that are worth noting. First of all, both Daubert and Metzler (2000) and Schoenberg et al. (2003, 2006) adopted a clinical comparison, simulation study design, with the latter having recruited college students as experimental feigners. In all simulation studies, the incentive offered to experimental feigners will never compare to what is at stake for real-life malingerers: as such, one cannot exclude the hypothesis that real-life malingerers may behave differently from experimental simulators. In addition, the demographic characteristics may be different between experimental feigners (especially college students) and real-life malingerers. Moreover, one cannot assume that

## RUNNING HEAD: MCMI-IV and Negative Response Bias

experimental feigners will comply to the instructions (e.g., Abeare et al., 2020). Finally, the only known-group design study (Aguerrevere et al., 2011) was conducted in a neuropsychological setting with patients with TBI, therefore it is unclear whether their results would apply also in other forensic settings. Taken together these studies suggest that the Modifier Indices alone showed only modest abilities to detect experimental feigners. Specifically, the Debasement (Z) scale showed promising performance in detecting malingering in real-life, neuropsychological evaluations, whereas it produced contradictory results in detecting feigning of psychopathology. Therefore, the first step would be to analyze the Modifier Indices as a group: the characteristic pattern for the exaggeration of psychopathology was elevations on both Disclosure (X) and Debasement (Z), and a low score on Desirability (Y), whereas the opposite pattern was seen with the minimization of psychopathology, with low scores on Disclosure (X) and Debasement (Z), and an elevation on Desirability (Y) (Bagby et al., 1991; Daubert & Metzler, 2000). The second step would be to use the algorithms described by Schoenberg et al. (2006), which showed increased ability to identify the experimental simulators. The third step would be to analyze the MCMI-III in its entirety. Typically, MCMI profiles suggesting an exaggeration of symptomatology have many scales with a BR score of 85, with the exception of the Histrionic, Narcissistic, and Compulsive scales that should exhibit low scores (Thomas-Peter et al., 2000). The clinical scales that tend to be elevated include the Schizoid, Avoidant, and Negativistic personality scales; the severe personality scales; and the Anxiety and Dysthymia scales (McNiel & Meyer, 1990; Retzlaff et al., 1991; Van Gorp & Meyer, 1986).

### ***Clinical exemplification***

Table 1 shows the MCMI-IV scores of 36-year-old Hispanic cook who slipped in the kitchen while at work. The patient complained of pain and had a great deal of trouble

## RUNNING HEAD: MCMI-IV and Negative Response Bias

ambulating, in spite of having only minor physical findings, according to his physicians. He was refusing to return to work and was involved in legal action against the employer. The examinee was married, had three children, and was the main breadwinner in the family.

Examining the results we note the Modifying Indices profile typical of an exaggerated record. An elevation on the Disclosure Scale indicates the examinee endorsed many more personality traits than normal; the elevation on Debasement showed excessive attribution of undesirable attributes, and that went with a low score on the socially laudable traits (Desirability Scale). The personality scales suggested an abundance of personality issues with an emphasis on the more dysfunctional scales. Finally, the clinical syndrome scales also indicate a high level of pathology.

A word of caution: MCMI-IV profiles that look like exaggerated profiles may be valid if obtained with chronic and very psychiatrically disturbed individuals. The above patient may have become very upset because of the accident he did have, and may believe he was experiencing all of the problems shown by the MCMI-IV. Depending on the whole file, a forensic assessor might attribute malingering, feigning, or case invalidity. As with the results of any one test, the MCMI should not be used alone to arrive at conclusions such as these.

### **Strengths and Weakness**

The MCMI is the shortest of the broad-based personality and emotional problem questionnaires. It is easy to read and complete. According to the manual, the current MCMI-IV version has good internal validity and concurrent validity. In terms of scale comparisons, it is highly correlated with the previous version, the MCMI-III. With the exception of the Emotional Assessment System (EAS-5; Choca, 2015), the MCMI-IV is more compatible with the DSM-5 than the other broad-based questionnaires, (e.g., the Minnesota Multiphasic Personality

RUNNING HEAD: MCMI-IV and Negative Response Bias

Inventory-3 [MMPI-3; Ben-Porath & Tellegen, 2020] or the Personality Assessment Inventory [PAI; Morey, 2007]).

The great advantage of the MCMI comes from its measurement of personality styles. The fact that one or more of the personality scales elevates with almost every examinee is very useful. For many years the first author administered the MCMI to the patients being admitted to psychiatric wards to determine if there was a prevalent personality style so to plan their treatment with the ward staff. In the case of an individual with a dependent personality style, for example, the supportive approach called for playing a paternal/maternal role, conveying the idea that the staff members were experts and would provide what was needed. For the individual with a compulsive personality style, an inspection of the DSM criteria was always gratifying, as well as the staff's being always meticulous and on time. The histrionic was in need of a great deal of attention and overt expressions of support. The MCMI revealed to be very useful with an antisocial personality individual who had left against medical advice on a previous occasion: he did not follow the rules of the ward, and did not accept the consequences. The staff discussed his personality style, pointing out its advantages but also the fact that his personality prevented him from being "a winner" at times by not being cunning enough to cope with the system. He was surprised as the staff looked at the times in his life when he had lost his edge because of impulsive and unproductive behavior. Other styles called for the staff's accepting interpersonal distance (schizoid), learning how to relax in their presence (avoidant), or treating the person with a great deal of admiration and respect (narcissist). The idea that an approach that was best for one individual may not be best for another was appealing, and the Millon system provided a finite number of personality style options that could be easily mastered.

## RUNNING HEAD: MCMI-IV and Negative Response Bias

Although there is only a small amount of scientific literature available on the Modifier Indices of the MCMI-III and none of the MCMI-IV, we can draw some positive implications. For example, the step-by-step procedure suggested before can help clinicians and researchers in detecting negative response styles by using multiple source of information. Moreover, a study on the utility of the MCMI-IV to evaluate parental fitness was recently published by Eastin and colleagues (2021), indicating that forensic psychologists should use the MCMI-IV with parental fitness litigants. Hopefully, future studies will support its utility also within the field of the negative response bias. Finally, given that the MCMI-IV is a clinically relevant tool to understand personality styles, it could be used along with other instruments and sources of information.

The MCMI was designed to reflect Millon's theory of psychopathology. As the Millon theory evolved, it became more complex and unwieldy. The MCMI added personality entities that are not recognized by clinicians or the DSM. Consequently, the advantage of the compatibility with the official nosology was partially lost. Moreover, To maintain a small number of items, the MCMI-IV uses the same items in different scales and has very high interscale correlations. The result is that it gives clinicians the illusion of measuring a variety of constructs while, in fact, it is mostly measuring the extent of psychopathology. That approach belongs to a previous generation of instruments and does not follow the more current emphasis on having scales that cleanly measure one construct, limiting the "contamination" of that construct with other issues. Consider, for instance, the developments that have taken place with the MMPI, with the more focused scales that have been constructed for the MMPI-3. From that point of view, the MCMI-IV has not "seen the light" and is persisting on an antiquated system that is less useful than it could be.

## RUNNING HEAD: MCMI-IV and Negative Response Bias

The most salient weakness of the MCMI-IV is its lack of studies attempting to evaluate the classification accuracy of its Modifying Indices and to replicate previously published findings on the MCMI-III. Indeed, to date there are not yet studies analyzing the ability of the MCMI-IV to evaluate response styles and the credibility of symptoms presentations; as such, by presenting the scientific literature available on the Modifier Indices of the MCMI-III, we assumed we could generalize their conclusions to the MCMI-IV. The MCMI instruments have been the focus of a heated debate about whether they meet the *Daubert* criteria for admissibility into court (e.g., Dyer, 2005; Dyer & McCann, 2000; Rogers et al., 1999, 2000). Nevertheless, forensic practitioners reported to be using the MCMI-III in forensic settings and to believe that it does meet the Daubert standard (Bow et al., 2006, 2010; Schutte, 2001).

### **Future Perspectives**

Because there is a lack of available studies on the classification accuracy of the MCMI-IV in evaluating the credibility of reported symptomatology and response styles, the need for research on these topics should rank high on the research agenda of academics and practitioners. Despite the aforementioned criticisms, the MCMI-III was commonly used in forensic evaluations (Archer et al, 2006, Bow et al., 2005; Craig, 2006; Schutte, 2001). Because of that, researchers and clinicians should join forces to produce valid and sound research studies on the MCMI-IV so that practicing psychologists may confidently use of the MCMI-IV in forensic settings knowing its strengths and weakness. On detecting negative response styles with the MCMI-IV, three issues should be addressed by the academic community. The first one refers to the fact that different tests perform better in detecting some disorders compared to others. The studies reported in the literature do not describe the type of disorder participants were asked to feign. As such, practitioners and researchers still do not know whether the MCMI-IV performs better in



## RUNNING HEAD: MCMI-IV and Negative Response Bias

detecting one disorder or another. The second topic refers to the fact that it is difficult to determine in which context the MCMI-IV may perform better because, so far, only two simulation studies in psychiatric settings (Daubert & Metzler, 2000; Morgan et al., 2002; Schoenberg et al., 2003, 2004, 2006) and one known group comparison in a neuropsychological setting (Aguerrevere et al., 2011) were conducted. The Modifier Indices of the MCMI-III have acceptable to excellent classification accuracy in the known-group design in a neuropsychological setting (Aguerrevere et al., 2011), but poor discrimination ability in the clinical comparison, simulation study in a psychiatric setting (Schoenberg et al., 2003). Future studies should address whether these differences depend on the context (psychiatric *vs.* neuropsychological) or on the study design (simulation *vs.* known-group). The third topic refers to the fact that practitioners are at a loss when they have to choose the best cut-off scores. Indeed, the two simulation studies offered the analysis of the best cut-off. Depending on the setting and the aim of the assessment, one may favor sensitivity over specificity or vice versa. Future studies should identify cutoff for the Modifying Indices that would yield sensitivity and specificity values as close as possible to .80, .90, and .95.

RUNNING HEAD: MCMI-IV and Negative Response Bias

### **Compliance with Ethical Standards**

### **Conflict of Interest**

The authors declare that they have no conflict of interest.

### References

- Abeare, C., Erdodi, L., Messa, I., Terry, D. P., Panenka, W. J., Iverson, G. L., Silverberg, N. D. (2020). Development of embedded performance validity indicators in the NIH Toolbox Cognitive Battery. *Psychological Assessment*. Advance online publication. <https://doi.org/10.1037/pas0000958>
- Aguerrevere, L. E., Greve, K. W., Bianchini, K. J., & Ord J. S. (2011). Classification accuracy of the Millon Clinical Multiaxial Inventory—III modifier indices in the detection of malingering in traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 33, 497–504. <https://doi.org/10.1080/13803395.2010.535503>
- American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, D.C.: Author.
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological tests use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 85–95. [https://doi.org/10.1207/s15327752jpa8701\\_07](https://doi.org/10.1207/s15327752jpa8701_07)
- Bagby, R. M., Gillis, J. R., Toner, B. B., & Goldberg, J. (1991). Detecting fake-good and fake-bad responding on the Millon Clinical Multiaxial Inventory-II. *Psychological Assessment*, 3(3), 496-498. <https://doi.org/10.1037/1040-3590.3.3.496>
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Minnesota Multiphasic Personality Inventory—2. Manual for administration and scoring*. Minneapolis: University of Minnesota Press.

## RUNNING HEAD: MCMI-IV and Negative Response Bias

Ben-Porath, Y., & Tellegen, A. (2020). *The Minnesota Multiphasic Personality Inventory – 3*.

Minneapolis, MN: Pearson Assessments.

Blood, L. (2008). The use of the MCMI-III in completing parenting capacity assessments.

*Journal of Forensic Psychology Practice*, 8, 24-38.

<https://doi.org/10.1080/15228930801947286>

Boccaccini, M. T., & Hart, J. R. (2018). Response Style on the Personality Assessment Inventory and Other Multiscale Inventories. In R. Rogers & S. D. Bender (Eds.), *Clinical Assessment of Malingering and Deception* (4th ed.), pp. 280-300. New York, NY: The Guilford Press.

Bow, J. N., Flens, J. R., & Gould, J. W. (2010). MMPI-2 and MCMI-III in forensic evaluations:

A survey of psychologists. *Journal of Forensic Psychology Practice*, 10, 37-52.

<https://doi.org/10.1080/15228930903173021>

Bow, J. N., Flens, J. R., Gould, J. W., & Greenhut, D. (2005). An analysis of administration, scoring, and interpretation of the MMPI-2 and MCMI-II/III in child custody evaluations.

*Journal of Child Custody*, 2(4), 1–22. [https://doi.org/10.1300/J190v02n04\\_01](https://doi.org/10.1300/J190v02n04_01)

Bow, J. N., Gould, J. W., Flens, J. R., & Greenhut, D. (2006). Testing in child custody evaluations- Selection, usage, and Daubert admissibility: A survey of psychologists.

*Journal of Forensic Psychology Practice*, 6, 17–38.

[https://doi.org/10.1300/J158v06n02\\_02](https://doi.org/10.1300/J158v06n02_02)

Choca, J. P. (2015). *The Emotional Assessment System: A DSM-5 diagnostic questionnaire*.

Chicago: Choca Assessments.

RUNNING HEAD: MCMI-IV and Negative Response Bias

- Choca, J. P. (2021). Millon Clinical Multiaxial Inventory (MCMI-IV). In I. B. Weiner and J. H. Kleiger (Eds.), *Psychological Assessment of Disordered Thinking and Perception*. Washington, DC: American Psychological Association.
- Choca, J. P. (2004). *Interpretative guide for the MCMI (3rd ed.)*. Washington, DC: American Psychological Association.
- Craig, R. J. (1999). Overview and Current Status of the Millon Clinical Multiaxial Inventory. *Journal of Personality Assessment, 72*(3), 390-406.  
<https://doi.org/10.1207/S15327752JP720305>
- Craig, R. J. (2006). The Millon Clinical Multiaxial Inventory-III. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 121–145). Mahwah, NJ: Lawrence Erlbaum.
- Daubert, S. & Metzler, A. (2000). The detection of fake-bad and fake-good responding on the Millon Clinical Multiaxial Inventory III. *Psychological Assessment, 12*, 418-424.  
<https://doi.org/10.1037/1040-3590.12.4.418>
- Demakis, G. J., Gervais, R. O., & Rohling, M. L. (2008). The effect of failure on cognitive and psychological symptom validity tests in litigants with symptoms of post-traumatic stress disorder. *The Clinical Neuropsychologist, 22*(5), 879–895.  
<https://doi.org/10.1080/13854040701564482>
- Derogatis, L.R. (1993) *The Brief Symptom Inventory (BSI): Administration, Scoring and Procedures Manual*. Minneapolis, MN: National Computer Systems..
- Dyer, F. J. (2005). Forensic applications of the MCMI-III in light of current controversies. In R. J. Craig (Ed.), *New Directions in Interpreting the Millon Clinical Multiaxial Inventory-III* (pp. 201–223). New York: John Wiley & Sons.

## RUNNING HEAD: MCMI-IV and Negative Response Bias

- Dyer, F. J., & McCann, J. T. (2000). The Millon clinical inventories, research critical of their forensic application, and Daubert criteria. *Law and Human Behavior, 24*, 487–497.  
<https://doi.org/10.1023/A:1005500615111>
- Eastin, S. M., Brandwein, D., Safran, R., & Giordano, K. (2021). The MCMI-IV: A new normal quartet for parental fitness litigants. *Journal of Personality Assessment*, published online.  
<https://doi.org/10.1080/00223891.2021.1912058>
- Gaasedelen, O. J., Whiteside, D. M., Altmaier, E., Welch, C., & Basso, M. R. (2019). The construction and the initial validation of the Cognitive Bias Scale for the Personality Assessment Inventory. *The Clinical Neuropsychologist, 33*(8), 1467–1484.  
<https://doi.org/10.1080/13854046.2019.1612947>
- Gervais, R. O., Ben-Porath, W., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment, 14*(2), 196–208.  
<https://doi.org/10.1177/1073191106295861>
- Giromini, L., Barbosa, F., Coga, G., Azeredo, A., Viglione, D. J., & Zennaro, A. (2020). Using the inventory of problems – 29 (IOP-29) with the Test of Memory Malingering (TOMM) in symptom validity assessment: A study with a Portuguese sample of experimental feigners. Advanced online publication. *Applied Neuropsychology: Adult, 27*(6), 504-516. <https://doi.org/10.1080/23279095.2019.1570929>
- Greiffenstein, M., Baker, W.J., & Gola, T. (1994). Validation of malingered amnesia measures in a large clinical sample. *Psychological Assessment, 6*(3), 218-224.  
<https://doi.org/10.1037/1040-3590.6.3.218>

RUNNING HEAD: MCMI-IV and Negative Response Bias

- Grossi, L. M., Green, D., Einzig, S., & Belfi, B. (2017). Evaluation of the Response Bias Scale and Improbable Failure Scale in assessing feigned cognitive impairment. *Psychological Assessment, 29*(5), 531-541. <https://doi.org/10.1037/pas0000364>
- Lampel, A.K. (1999). Use of the Millon Clinical Multiaxial Inventory-III in evaluating child custody litigants. *American Journal of Forensic Psychology, 17*, 19–31.
- Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of International Neuropsychological Society, 18*(4), 625–630. <https://doi.org/10.1017/s1355617712000240>
- Larrabee, G. J., Bianchini, K. J., Boone, K. B., & Rohling, M. L. (2017). The validity of the MMPI-2/MMPI-2-RF Symptom Validity Scale (FBS/FBS-r) is established: Reply to Nichols (2017). *The Clinical Neuropsychologist, 31*(8), 1401–1405. <https://doi.org/10.1080/13854046.2017.1363293>
- McCann, J.T., Flens, J.R., Campagna, V., Collman, P., Lazzaro, T., & Connor, E. (2001). The MCMI-III in child custody evaluations: A normative study. *Journal of Forensic Psychology Practice, 1*(2), 27–44. [https://doi.org/10.1300/J158v01n02\\_02](https://doi.org/10.1300/J158v01n02_02)
- McNiel, K., & Meyer, R. (1990). Detection of deception on the Millon Clinical Multiaxial Inventory (MCMI). *Journal of Clinical Psychology, 46*(6), 755-764. [http://dx.doi.org/10.1002/1097-4679\(199011\)46:6<755::AID-JCLP2270460609>3.0.CO;2-J](http://dx.doi.org/10.1002/1097-4679(199011)46:6<755::AID-JCLP2270460609>3.0.CO;2-J)
- Millon, T. (1969). *Modern psychopathology: A biosocial approach to maladaptive learning and functioning*. Philadelphia: W. B. Saunders.
- Millon, T. (1977). *Millon Clinical Multiaxial Inventory*. Minneapolis, MN: National Computer Systems.

RUNNING HEAD: MCMI-IV and Negative Response Bias

- Millon, T. (1982). *Millon Clinical Multiaxial Inventory manual (2nd ed.)*. Minneapolis, MN: National Computer Systems.
- Millon, T., Davis, R., Millon, C., & Grossman, S. (2009). *Millon Clinical Multiaxial Inventory-III, 4th ed. (MCMI-III)*. Minneapolis, MN: Pearson Assessments.
- Millon, T., Grossman, S., & Millon, C. (2015). *Millon Clinical Multiaxial Inventory-IV (MCMI-IV)*. Minneapolis, MN: Pearson Assessments.
- Morey, L. C. (2007). *Personality assessment inventory (PAI): Professional manual*. Lutz, FL: Psychological Assessment Resources.
- Morgan, C.D., Schoenber, M.R., Dorr, D., & Burke, M.J. (2002). Overreport on the MCMI-III: Concurrent validation with the MMPI-2 using a psychiatric inpatient sample. *Journal of Personality Assessment, 78*(2), 288-300.  
[http://dx.doi.org/10.1207/S15327752JPA7802\\_05](http://dx.doi.org/10.1207/S15327752JPA7802_05)
- Nelson, N. W., Sweet, J. J., Berry, D. T., Bryant, F. B., & Granacher, R. P. (2007). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society, 13*(3), 440-449. <https://doi.org/10.1017/S1355617707070373>
- Retzlaff, P. D., Sheehan, E. P., & Fiel, A. (1991). MCMI-II report style and bias: Profile and validity scales analyses. *Journal of Personality Assessment, 56*(3), 466-477. [http://dx.doi.org/10.1207/s15327752jpa5603\\_8](http://dx.doi.org/10.1207/s15327752jpa5603_8)
- Rogers, R., Salekin R. T., & Sewell, K. W. (2000). The MCMI-III and the Daubert standard: Separating rhetoric from reality. *Law and Human Behavior, 24*(4), 501–506.  
<https://doi.org/10.1023/A:1005504716020>



## RUNNING HEAD: MCMI-IV and Negative Response Bias

- Rogers, R., Salekin R. T., & Sewell, K. W. (1999). Validation of the Millon Clinical Multiaxial Inventory for Axis II disorders: Does it meet the Daubert standard? *Law and Human Behavior, 23*(4), 425–443. <https://doi.org/10.1023/a:1022360031491>
- Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment, 10*, 160-177. <https://doi.org/10.1177/1073191103010002007>
- Ruocco, A.C., Swirsky-Sacchetti, T., Chute, D.L., Mandel, S., Platek, S.M., & Zillmer, E.A. (2008). Distinguishing between neuropsychological malingering and exaggerated psychiatric symptoms in a neuropsychological setting. *The Clinical Neuropsychologist, 22*(3), 547-564. <http://dx.doi.org/10.1080/13854040701336444>
- Sabelli, A. G., Messa, I., Giromini, L., Lichtenstein, J. D., May, N., & Erdodi, L. A. (2021). Symptom Versus Performance Validity in Patients with Mild TBI: Independent Sources of Non-credible Responding. *Psychological Injury and Law, 14*, 17–36. <https://doi.org/10.1007/s12207-021-09400-6>
- Schoenberg, M., Dorr, D., & Morgan, D. (2003). The ability of the Millon Clinical Multiaxial Inventory--Third Edition to detect malingering. *Psychological Assessment, 15*(2), 198-204. <https://doi.org/10.1037/1040-3590.15.2.198>
- Schoenberg, M., Dorr, D., & Morgan, D. (2004). A comparison of the MCMI-III personality disorder and modifier indices with the MMPI-2 clinical and validity scales. *Journal of Personality Assessment, 82*(3), 273-280. [https://doi.org/10.1207/s15327752jpa8203\\_03](https://doi.org/10.1207/s15327752jpa8203_03)
- Schoenberg, M., Dorr, D., & Morgan, D. (2006). Development of discriminant functions to detect dissimulation for the Millon Clinical Multiaxial Inventory(3rd edition). *The*

RUNNING HEAD: MCMI-IV and Negative Response Bias

*Journal of Forensic Psychiatry & Psychology*, 17(3), 405-

416. <https://doi.org/10.1080/14789940600761394>

Schutte, J. W. (2001). Using the MCMI-III in forensic evaluations. *American Journal of*

*Forensic Psychology*, 19(2), 5–20.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered

neurocognitive dysfunction: Proposed standards for clinical practice and research. *The*

*Clinical Neuropsychologist*, 13(4), 545–561. <https://doi.org/10.1076/1385->

[4046\(199911\)13:04;1-Y;FT545](https://doi.org/10.1076/1385-4046(199911)13:04;1-Y;FT545)

Strack, S. (1991, August). Response bias and the MCMI-II: Clinical and research issues. Paper

presented at the 99th Annual Convention of the American Psychological Association, San

Francisco, CA.

Sumanti, M., Boone, K. B., Savodnik, I., & Gorsuch, R. (2006). Noncredible psychiatric and

cognitive symptoms in a workers' compensation "stress" claim sample. *The Clinical*

*Neuropsychologist*, 20, 754–765. <https://doi.org/10.1080/13854040500428467>

Thomas-Peter, B.A., Jones, J., Campbell, S., & Oliver, C. (2000). Debasement and faking bad on

the Millon Clinical Multiaxial Inventory III: An examination of characteristics,

circumstances and motives of forensic patients. *Legal and Criminological Psychology*,

5(Part 1), 71-81. <https://doi.org/10.1348/135532500167985>

Tombaugh, T.N. (1996). *Test of Memory Malingering (TOMM)*. Toronto, ON: Multi-Health

Systems.

Van Dyke, S. A., Millis, S. R., Axelrod, B. N., & Hanks, R. A. (2013). Assessing effort:

Differentiating performance and symptom validity. *The Clinical Neuropsychologist*,

27(8), 1234–1246. <https://doi.org/10.1080/13854046.2013.835447>

RUNNING HEAD: MCMI-IV and Negative Response Bias

Van Gorp, W. G., & Meyer, R. G. (1986). The detection of faking on the Millon Clinical  
Multiaxial Inventory (MCMI). *Journal of Clinical Psychology, 42*(5),742-747.

[https://doi.org/10.1002/1097-4679\(198609\)42:5<742::AID-JCLP2270420510>3.0.CO;2-](https://doi.org/10.1002/1097-4679(198609)42:5<742::AID-JCLP2270420510>3.0.CO;2-)

3

## RUNNING HEAD: MCMI-IV and Negative Response Bias

Table 1

*The MCMI-IV scales and their base rates of the Clinical Exemplification*

Modifier Indices:	Scores=>	Raw	Base Rate
X - Disclosure.....	=	172	100 ***
Y - Desirability.....	=	4	20
Z - Debasement.....	=	30	95 **
V - Validity.....	=	0	
W - Inconsistency.....	=	3	
 Personality Style Scales:			
1 - Introversive (Schizoid).....	=	20	88 **
2A - Inhibited (Avoidant).....	=	22	90 **
2B - Depressive.....	=	19	87 **
3 - Cooperative (Dependent).....	=	20	90 **
4A - Dramatic (Histrionic).....	=	6	6
4B - Turbulent.....	=	2	9
5 - Confident (Narcissistic).....	=	2	0
6A - Competitive (Antisocial).....	=	11	53
6B - Aggressive/Sadistic.....	=	17	58
7 - Disciplined (Compulsive).....	=	11	21
8A - Negativistic (Passive-Aggressive/Explosive).=		23	89 **
8B - Self-Defeating.....	=	22	87 **
 Severe Personality Scales:			
S - Schizotypal.....	=	23	90 **
C - Borderline.....	=	20	79 *
P - Paranoid.....	=	20	80 *
 Clinical Symptom Scales:			
A - Anxiety.....	=	14	89 **