

HISTORICAL HIGHLIGHTS ON THE EMPIRICAL METHOD UNDERLYING THE MMPI/MMPI-2 AND MMPI-A¹ 3/29/20

James N. Butcher
Professor Emeritus
University of Minnesota

What is there about the MMPI scales that make them such valid and effective measures at detecting mental health and personality problems? What is the basis for their long-term survival value in personality assessment? The primary reason is that they were developed with the idea that personality scale items needed to be valid predictors before they were included on a measure. The original MMPI clinical valid scales were constructed following the *empirical method of scale construction*. The empirical method has evolved over the past 75 years. The following highlights describe some special contributions to the empirical scale construction method to assure that the test scales were appropriate, reliable, and valid in predicting symptoms and behavior relevant to psychological treatment. Hundreds of scales have been developed for the MMPI/MMPI-2 and MMPI-A. Several major research studies of the MMPI/MMPI-2/MMPI-A are summarized and both modification and verification of the empirical approach to scale development are highlighted.

1938 Paterson, Schneidler & Williamson did not accept the rational scale development method that was widely used in constructing psychological measures because they thought that some items could be predictive of relevant criteria without having an apparent or obvious content connection. They believed that items needed to have proven utility before they were incorporated in the scale. They considered the empirical method of item selection to be the most effective basis for scale construction.

¹ Citation: Butcher, J.N. (2020). Historical highlights on the empirical method underlying the MMPI/MMPI-2/MMPI-A. Retrieved from <http://www.umn.edu/mmpi>

- 1940 Hathaway & McKinley introduced the Minnesota Multiphasic Personality Inventory as an empirically based self-report instrument that could assess clinical symptoms by differentiating people with mental health problems from normal individuals. Scales were developed empirically: using items that actually discriminated the criterion group from the control sample. The basic clinical scales (Hypochondriasis or Hs, Depression or D, Hysteria or Hy, Psychopathic Deviate or Pd, Paranoia or Pa, Psychasthenia or Pt, Schizophrenia or Sc and Hypomania or Ma) have served the assessment field for decades as effective measures largely with the original item content (Hathaway and McKinley, 1940a,1940b; McKinley & Hathaway, 1940, 1944).
- 1945 The empirical method of scale construction was placed in theoretical context with Meehl's publication of the "Dynamics of structured personality tests." Meehl argued that a test response could be considered verbal behavior in its own right and concluded that nontest correlates could be discovered or established empirically. In his view, criterion keying permitted a wide range of practical problems to be addressed in assessment.
- 1946 One of the most effective and most useful scales that were developed for the MMPI is the Social Introversion/Extraversion or Si scale by Drake. This scale was empirically developed by using MMPI items that clearly discriminated college students who were high on the TSE (Thinking -Social Emotional Introversion Inventory) versus those who were low on the scale. Scores on the Si scale have shown high stability over long periods of time, a characteristic that is valuable to consider in personality assessment.
- 1951 Gough, McClosky and Meehl(1951,1952) developed two empirical scales that measure more "normal range" personality characteristics using MMPI items that have endured the test of time. Both measure assessment today to address personality adaptation. The Dominance scale (Do) was developed by selecting items that differentiated high school students who were rated by peers as being more or less dominant in social relationships. The Responsibility (Re scale included items that separated students who were more willing to accept the consequence soft behavior, having integrity, being trustworthy and having a sense of social responsibility.
- 1954 Meehl's article on clinical versus statistical prediction established actuarial prediction approach in psychology as a strength based scales. Empirically validated scale scores could be objectively used to predict or verify established patterns of behavior.

- 1954 Cook and Medley developed the Hostility (Ho) scale to predict rapport of teachers with students. The scale was developed by contrasting teachers who were rated either high or low on teacher pupil interactions. The items that discriminated these two group conditions were refined by content analysis. Subsequent research suggests the Ho scale measures cynicism and hostility and relates to general maladjustment and anger control. The Ho scale has been studied extensively as a prediction of personality factors that could predispose one to heart disease. The Ho scale was revised and updated for MMPI 2 in 1995 by Han, Weed, Calhoun and Butcher.
- 1955 Halbower provided the first empirical or behavioral correlation approach to objective test interpretation verifying Meehl' statistical (empirical) assessment strategy.
- 1955 Cronbach and Meehl provided a construct validity model that expanded the "blind empiricism" approach by advocating a construct validity approach rather than relying upon a strictly "blind empiricism" in test construction. This approach includes: articulating a set of theoretical concepts and their interrelations, developing ways to measure the hypothetical constructs proposed by the theory, and empirically verifying the hypothesized relations among constructs and their observable behaviors.
- 1963 Following Meehl's and Halbower's empirical validation strategy, Marks & Seeman's actuarial prediction research provided an interpretive resource, referred to as a "cookbook," for tying MMPI codes to behaviors of outpatients. This system provided valuable empirical correlates that served as an interpretive base for the MMPI measures for many years.
- 1965 Gilberstadt & Duker developed an MMPI code type cook-book interpretation system for VA inpatients that became widely used in clinical as well as computer-based interpretation.
- 1965 One of the most widely used and researched MMPI empirically developed "special scales" is the MacAndrew Substance Abuse Scale. This scale was developed by Craig MacAndrew using the empirical scale development approach from the MMPI items. The scale was successful in differentiating between alcoholic and non-alcoholic psychiatric outpatients. This 49 item scale was applied to the cross-validation samples. Cross-validation of empirical scales was an essential ingredient of empirical scale construction in order to eliminate random or non-replicable items. He reported that significant differences in MMPI responses do exist between these 2 classes of psychiatric patients.
- 1973 Gynther, Altman, & Sletten conducted a number of empirical validation studies for MMPI measures verifying the correlates of clinical and some validity scales.

- 1978 Clopton provided a summary and evaluation of the empirical methodology for developing MMPI personality scales. He provided recommendations concerning the need for cross-validation of developed measures, the importance of eliminating items that were obtained by chance, and the need to consider possible impact of defensiveness and whether K scale corrections are appropriate.³
- 1984 The scale developed by Keane, Malloy and Fairbank to assess post-combat related post-traumatic stress disorder has become one of the most widely used and effective empirical scales developed for the MMPI. This measure was developed by using items that discriminated between Vietnam War combat veterans who had been diagnosed with PTSD with veterans who had diagnoses other than PTSD.
- 1989 MMPI-2 was published to revise the MMPI (Butcher,Dahlstrom, etal.,1989). The empirically constructed clinical scales were maintained intact in order to provide continuity with the empirical scales developed by Hathaway and McKinley. These scales were re-normed with a more contemporary sample and further empirical validation was under taken. Personality based information was obtained on a subset of the normative sample to provide further behavioral correlate information. Moreover, a number of clinical studies were conducted to verify the empirical validity of the scales.
- 1989 Williams and Butcher conducted an empirical study of MMPI clinical scales with adolescents. They reported contemporary behavioral correlates for the empirically derived standard scales of the MMPI with adolescents.
- 1995 Clark and Watson broadened the scale development approach by also including the importance of cross-validation, determining internal consistency and establishing construct validity through external validation efforts.
- 1995 Archer, Griffin and Aiduk examined MMPI-2 clinical correlates found for 9 commonly occurring 2-point codes. Major findings indicated that descriptors identified for MMPI-2 based codes were generally similar to the established literature for corresponding codes derived from the original MMPI.
- 1995 Butcher and Han developed an empirical scale to assess test takers, such as job applicants, who proclaim extreme virtue and absence of psychopathology on the MMPI-2 item pool. The S-scale differentiated high virtue-claiming individuals (airline flight crew applicants) from a general nonclinical sample—the MMPI-2 normative sample. The items that differentiated a group of people with a clear, strong motivation to show themselves extremely well adjusted from people in general would reflect the tendency to portray oneself in a superlative manner. The goal was to develop a scale of superlative claim assertions that could summarize the attempt to proclaim overly positive self-presentation and would not be as susceptible to conscious distortion (as scale L). Further analyses assuring scale homogeneity were conducted. Subscales were developed using item factor analysis of the full scale to provide more specific information about high scoring clients.

External correlate data for the S scale were also provided.

- 1996 Grove and Meehl reaffirmed that objective assessment procedures are equal or superior to subjective interpretation methods. Given a data set about an individual or a group such as interviewer ratings, life history or demographic facts, test results, or self-descriptions, there are two methods of combining data for a predictive or diagnostic purposes. The clinical method relies on human judgment whereas the mechanical method involves a formal, algorithmic, objective procedure (e.g., equation) to reach the decision. Empirical comparisons of the accuracy of the two methods show that the mechanical method is almost always equal to or superior to the clinical method.
- 2000 Butcher, Rouse and Perry conducted an empirical validation study of MMPI-2 clinical scales and developed an empirically-based correlate net measuring psychopathology in therapy clients including well validated correlates of MMPI-2 scales.
- 2000 Butcher provided an overview of the status of personality assessment from an empirical scale-development perspective drawn from recent work on the MMPI-2. Meehl's heuristic defense of empirically based personality-scale construction is reexamined and the lasting influences of these views highlighted. Meehl's early conceptualization of the relative unimportance of item content in personality-test construction and several alternative views are summarized for his modified position. The role that test-taking attitudes can play in personality assessment was discussed in Meehl's original article, and Meehl's views on the need for appraisal of invalidating conditions have been reaffirmed in contemporary test development. Finally, the "dynamics" of a structured personality item response are discussed from a contemporary perspective.
- 2018 Erard, Nichols and Friedman (2018) provided a comprehensive chapter on evaluating psychopathology with personality assessment instruments.
- 2019 Lee examined whether there is equivalence found in MMPI-2 codetype 1-3/3-1 and the scales provided in the MMPI-2 RF. The existence of subtypes within the 1-3/3-1 sample led to vastly different scale elevations on the MMPI-2-RF and, therefore, caution is warranted in assuming linearity of the MMPI-2-RF with the MMPI-2 in interpretation. Instead, the measures need to be seen as two separate instruments with their own unique strengths and weaknesses and research basis.

References

- Archer, R. P., Griffin, R. & Aiduk, R. (1995). Clinical correlates for ten common code types. *Journal of Personality Assessment*, 65, 391-408.
- Butcher, J. N. (2000). Dynamics of personality test responses: The empiricist's

manifesto revisited. *Journal of Clinical Psychology*, 56(3), 375-386.

Butcher, J. N., Dahlstrom, W.G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI-2. An administrative and interpretive guide.* Minneapolis, MN: University

Butcher, J. N. & Han, K. (1995). Development of an MMPI-2 scale to assess the presentation of self in a superlative manner: The S Scale. In J.N. Butcher and C.D. Spielberger (Eds.) *Advances in personality assessment*, Volume 10. (pp 25-50). Hillsdale, N.J: LEA Press.

Butcher, J. N., Rouse, S. V., & Perry, J. N. (2000). Empirical description of psychopathology in therapy clients: Correlates of MMPI-2 scales. (pp 487-500). Chapter in J. N. Butcher (Ed) (2000). *Basic sources on the MMPI-2.* Minneapolis, MN: University of Minnesota Press.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.

Clopton, J. R. (1978). MMPI scale development methodology. *Journal of Personality Assessment*, 42(2), 148-151.

Cook, W. N. & Medley, D. M. (1954). Proposed hostility and pharisaic-virtue scales for the MMPI. *Journal of Applied Psychology*, 38, 414-418.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological test. *Psychological Bulletin*, 52, 281-302.

Drake, L. E. (1946). A social I.E. scale for the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology*, 30, 51-54.

Erard, R. E., Nichols, D. and Friedman, A. (2018). Evaluating Psychopathology with Personality Assessment Instruments. In Butcher, J. N. (Editor-in-Chief), & Hooley, J. M. (Ed.) (2018). *APA Handbook of Psychopathology: Vol 1. (Pp. 169-199). Understanding, assessing, and treating adult mental disorders.* Washington, DC: American Psychological Association.

Halbower, C. C. (1955). A comparison of actuarial versus clinical prediction to classes discriminated by the Minnesota Multiphasic Personality Inventory. *Dissertation Abstracts International*, 15, 1115.

Gilberstadt, H. D. J. (1965). *A handbook for clinical and actuarial MMPI interpretation.* Philadelphia: W.B. Saunders.

Gough, H. G., McClosky, H. & Meehl, P. E. (1951). A personality scale for dominance. *Journal of Abnormal and Social Psychology*, 46, 360-366

- Gough, H. G., McClosky, H. & Meehl, P. E. (1952). A personality scale for social responsibility. *Journal of Abnormal and Social Psychology*, 47, 73-80.
- Grove, W. M. & Meehl, P. E. (1996). Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: The clinical–statistical controversy. *Psychology, Public Policy, and Law*, 2(2), 293-323.
- Gynther, M. D., Altman, H., & Sletten, I. W. (1973). Replicated correlates of MMPI two - point code types: The Missouri actuarial system. *Journal of Clinical Psychology*, 29(3), 263-289.
- Halbower, C. C. (1955). A comparison of actuarial versus clinical prediction to classes discriminated by the Minnesota Multiphasic Personality Inventory. *Dissertation Abstracts International*, 15, 1115.
- Han, K., Weed, N., Calhoun, R., & Butcher, J. N. (1995). Psychometric characteristics of the MMPI-2 Cook-Medley Hostility Scale. *Journal of Personality Assessment*, 65, 567-586.
- Hathaway, S. R., & McKinley, J. C. (1940a). A multiphasic personality schedule (Minnesota): I. Construction of the schedule. *Journal of Psychology*, 10, 249 - 254.
- Hathaway, S. R., and McKinley, J. C. (1940b). A multiphasic personality schedule (Minnesota) III. The measurement of symptomatic depression. *Journal of Psychology*, 14, 73-84
- Keane, T. M., Malloy, P. F. & Fairbank, J. A. (1984). Empirical development of a subscale for the assessment of combat-related posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology*, 52, 888-891.
- Lee, N. (2019). *Is there an equivalent of MMPI-2 codetype and 1-3/3-1 in the MMPI-2 RF scales?* (Doctoral dissertation). Retrieved from ProQuest Information and Learning. (52510-019).
- MacAndrew, C. (1965). The differentiation of male alcoholic outpatients from nonalcoholic psychiatric outpatients by means of the MMPI. *Quarterly Journal of Studies of Alcohol*, 26, 238-246.
- Marks, P. A. Seeman, W. (1963). *The actuarial description of personality: an atlas for use with the MMPI*. Baltimore: Williams and Wilkins.
- McKinley, J. C. & Hathaway, S. R. (1944). The MMPI: V. Hysteria, hypomania, and

psychopathic deviate. *Journal of Applied Psychology*, 28, 153-174.

McKinley, J. C. Hathaway, S. R. (1940). A multiphasic personality schedule (Minnesota): II. A differential study of hypochondriasis. *Journal of Psychology*, 10: 255-268.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and are view of the evidence*. Minneapolis: University of Minnesota Press.

Meehl, P. E. (1945). The dynamics of "structured" personality tests. *Journal of Clinical Psychology*, 1, 296-303, of-303.

Paterson, D. G., Schneidler, G. G. & Williamson, E. G. (1938). *Student guidance techniques: A handbook for counselors in high schools and colleges*. New York: McGraw-Hill Book Company.

Williams, C. L., & Butcher, J. N. (1989). An MMPI study of adolescents: I. Empirical validity of the standard scales. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 251-259.