Minnesota Multiphasic Personality Inventory (MMPI)

Tayla T. C. Lee
Ball State University, Muncie, IN, USA

Synonyms
MMPI

The Minnesota Multiphasic Personality Inventory (MMPI, Hathaway and McKinley 1943) is a self-report inventory consisting of 550 true/false items historically used to assess a test-taker’s personality, as well as their personal and social adjustment. Between the time it was developed by Starke Hathaway and J. Charnley McKinley in the 1940s and its revision by James Butcher and colleagues in the 1980s, the MMPI was the most widely used psychological test assessing personality and psychopathology (e.g., Harrison et al. 1988). The MMPI and its subsequent revisions, the MMPI-2 (Butcher et al. 2001) and the Minnesota Multiphasic Personality Inventory – 2 – Restructured Form (MMPI-2-RF; Ben-Porath and Tellegen 2008/2011; Tellegen and Ben-Porath 2008/2011), have had a lasting impact on the field of psychological assessment due to the use of empirical scale development methodologies, the inclusion of scales that assess a test-taker’s response style, and having interpretation guided by research conducted using the empirical correlates approach.

The MMPI was developed during a period of great skepticism toward self-report inventories but eventually stood out due to the empirical methods used to develop and validate the scale scores (Dahlstrom 1992). In developing the MMPI, Hathaway and McKinley’s goal was to develop a self-report instrument that would provide a more efficient and reliable method of determining psychiatric diagnoses than the intensive and unreliable battery of interviews, observations, and projective and self-report inventories that were standardly used in hospital settings (Hathaway and McKinley 1940). Starting with a pool of well over 1000 potential items drawn from diagnostic interviews and other self-report instruments, Hathaway and McKinley developed eight scales intended to assess psychiatric diagnoses of the time period. These became known as the “Clinical Scales” and included Hypochondriasis (1 – Hs), Depression (2 – D), Hysteria (3 – Hy), Psychopathic Deviance (4 – Pd), Paranoia (6 Pa), Psychasthenia (7 – Pt), Schizophrenia (8 – Sc), and Hypomania (9 – Ma) (Dahlstrom and Dahlstrom 1980). Two additional scales, Masculinity/Femininity (Mf) and Social Introversion (Si) were later developed and added to the existing MMPI Clinical Scales.

The Clinical Scales were developed using an empirical keying approach, in which individuals’ responses to items are contrasted statistically to determine which items best discriminate between
MMPI scales were developed by comparing the responses of individuals in varying diagnostic patient groups with those provided by “Minnesota Normals,” a group that consisted of visitors and relatives who came to the hospital (Hathaway and McKinley 1940). Using the empirical keying approach to develop scales for the MMPI was a departure from the typical methods used in the 1940s (Dahlstrom 1992). Most instruments at that time were developed using a logical keying approach where items and scoring were based on the test creator’s subjective judgment. The adequacy of instruments developed using the logical keying approach had been questioned due to research demonstrating an inconsistent pattern of scale score differences between purportedly distinct groups of individuals. Over time, the MMPI item pool proved to be a rich source of content for additional scale development efforts. Dahlstrom, Welsh, and Dahlstrom noted in 1975 that over 450 additional scales had been developed from the MMPI item pool using varying methods. Although these scales varied greatly in purpose, as well as in the rigor with which they were developed, many of these scales represented a continuation of the MMPI being on the cutting edge of psychometric considerations.

In addition to using the empirical keying approach, Hathaway and McKinley designed scales for the MMPI that were intended to assess the test-taker’s response style, which later became known as “Validity Scales.” It was increasingly recognized in the 1940s that a test-taker could falsify or distort their responses to test items resulting in scale scores that did not accurately reflect that individual. As such, the original MMPI included two scales intended to detect individuals whose approach to answering the test items differed from what they were instructed (Dahlstrom and Dahlstrom 1980; Hathaway and McKinley 1943). The first was the Cannot Say (?) Scale, which is a count of the number of items a test-taker did not provide a response to or answered as both True and False. High ? scores called into question the interpretability of the substantive scales, as the omission of a large number of items suppressed scale scores. The second was the Infrequency (F) scale, which consisted of items that were rarely endorsed by the normative sample. High F scores were originally believed to be indicative of the test-taker answering items without closely considering their content. Later research would suggest that high F scores could be indicative of severe psychopathology or an over-reporting response style where the test-taker attempted to exaggerate or fabricate problems and difficulties. Rational scale development methods were used to create a scale intended to detect individuals who engaged in unsophisticated attempts to present in a favorable light. Named the Lie (L) scale, it consisted of items that indicated the test-taker was denying minor faults and shortcomings. A fourth Validity Scale was developed to detect clinical defensiveness. This scale, called the Correction (K) Scale, was intended to detect those who under-reported their difficulties. Scores on this scale were also used to adjust scores on substantive scales to provide a profile that was believed to represent what the test-taker’s true functioning would be if a defensive test taking style had not been used. The inclusion of these scales on an instrument used in clinical practice situated the MMPI uniquely to garner attention and popularity after it was published (Dahlstrom 1992). Their inclusion also later influenced the development of additional validity scales detecting other specific types of response distortion for the MMPI family of instruments, as well scales assessing test taking styles on other self-report inventories.

The last strength of the MMPI was the large research base supporting interpretation of the instruments’ scale scores. Although MMPI scale scores were intended to provide a reliable method of diagnosing patients, it quickly became clear that the Clinical Scales were insufficient for this purpose (Dahlstrom and Dahlstrom 1980). This conclusion was reached because patients with a specific diagnosis often achieved high scores on purportedly unrelated Clinical Scales and because individuals with no observable psychopathology sometimes achieved scores indicative of a problem on the Clinical Scales. MMPI proponents
argued that despite these difficulties, it was clear that MMPI scale scores were detecting something, as research had demonstrated that there were reliable score differences between groups of individuals known to differ in important ways. They advocated for an approach to understanding what each of the MMPI scales was measuring that became known as the empirical correlates approach. In this approach, scale scores were treated as measures of unidentified constructs and empirical investigations were used to identify the construct by examining potential relations between specific scales’ scores with other measures of important social, emotional, and behavioral phenomena. Proponents of this approach also advocated for the use of “code types,” which were combinations of elevated scale scores reflecting purportedly even more distinct constructs. These approaches enabled a test-taker to be described using what research suggested other individuals who scored similarly were like. The use of the empirical correlates approach was furthered by movement in the larger field of personality assessment regarding the need for actuarial interpretation of test data. This shift toward wanting an empirically based approach to test interpretation was led by Paul Meehl who called for “a good cookbook” for interpreting test data in which he used the MMPI as an example (1956). This call led to numerous attempts to develop MMPI interpretation systems that would allow test scores to be interpreted using actuarial methods. This momentum was maintained over time and when the MMPI was revised in the late 1980s, it had become the most widely researched psychological test in the United States (Butcher and Rouse 1996). The rich tradition of empirically supported interpretation has continued with subsequent revisions to the MMPI, including the MMPI-2 and the MMPI-2-RF.

Cross-References
- Minnesota Multiphasic Personality Inventory – 2
- Minnesota Multiphasic Personality Inventory – 2 – Restructured From
- Personality Assessment

References