

MASCULINITY-FEMININITY TEST

Authored by
mohammad looti

October 31, 2025

RECOMMENDED CITATION

mohammad looti (2025). *MASCULINITY-FEMININITY TEST*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=63634>

MASCULINITY-FEMININITY TEST

Primary Disciplinary Field(s): Psychology, Psychometrics, Gender Studies, Sociology

1. Core Definition

The **Masculinity-Femininity Test (MFT)** is a class of psychometric instruments designed to measure the degree to which an individual exhibits psychological traits, attitudes, interests, or behaviors traditionally associated with either masculinity or femininity within a specific cultural context. These tests operate on the premise that gender-related attributes can be quantified and assessed, positioning the individual along a conceptual dimension or continuum. Historically, the primary goal was to distinguish between individuals based on their conformity to conventional gender roles, often relying on self-report inventories that catalogue vocational preferences, emotional responses, activity choices, and societal attitudes.

In its earliest forms, the MFT operated under the assumption of a strict **bipolar continuum**, implying that high scores on masculinity necessarily meant low scores on femininity, and vice versa. This theoretical framework dominated psychometric attempts to measure gender differences throughout the mid-twentieth century. Test construction relied heavily on empirical keying--items were selected based on whether male and female normative groups answered them differently. Consequently, the resulting score did not necessarily measure an innate quality but rather the extent to which an individual's response pattern mirrored that of the culturally defined typical man or woman of that era.

The methodology employed by MFTs typically involves presenting a large battery of questions, statements, or paired comparison choices to the respondent. Scoring algorithms then aggregate the responses into a single index or, in later iterations, two separate indices. The conceptualization and scoring of MFTs have been significantly criticized and revised over time, reflecting evolving understanding within gender psychology, particularly the distinction between **biological sex**, **gender identity**, and **gender role expression**.

2. Etymology and Historical Development

The genesis of formal psychometric measurement of masculinity and femininity is traced back to the landmark work of Lewis M. Terman and Catherine Cox Miles. In 1938, they formulated the **Terman-Miles Attitude Interest Analysis Test (AIA)**, which is widely recognized as the first comprehensive attempt to quantify M-F traits. The AIA consisted of seven sections covering word association, inkblot interpretation, emotional responses, opinions, and interests, yielding a single score reflecting the degree of deviation from the average same-sex profile. Terman and Miles utilized this instrument to explore the psychological dimensions they believed differentiated men

and women, setting a precedent for subsequent research in the field.

Following the AIA, the concept was integrated into major personality inventories. Most notably, the Mf (Masculinity-Femininity) scale was developed for the Minnesota Multiphasic Personality Inventory (MMPI). The MMPI Mf scale, primarily developed by Starke Hathaway, was originally conceived to identify male subjects whose interests and behaviors deviated significantly from the conventional male norm, serving initially as a clinical indicator rather than a general personality measure. Its construction relied on empirical item selection, primarily comparing the responses of male homosexuals with heterosexual men--a controversial origin that rooted the scale in outdated clinical assumptions regarding sexual orientation and gender conformity.

Further instruments emerged, such as the M-F scales within the California Psychological Inventory (CPI), developed by Harrison Gough. These inventories continued to reflect the bipolar model prevalent until the 1970s. However, the theoretical utility of a single M-F dimension began to wane as researchers recognized that many individuals displayed high levels of traits traditionally assigned to both genders. This recognition paved the way for the radical conceptual shift toward androgyny, fundamentally altering the trajectory of M-F measurement.

3. Key Concepts and Components

The measurement of masculinity and femininity shifted dramatically from unitary concepts to multidimensional constructs in the 1970s. This conceptual evolution necessitated new instruments capable of assessing these dimensions independently. The most influential tools reflecting this change include the **Bem Sex-Role Inventory (BSRI)** and the **Personal Attributes Questionnaire (PAQ)**.

Bipolar Measurement (Pre-1970s): This early framework assumed masculinity and femininity were opposite ends of the same psychological continuum. An individual could be scored as more masculine or more feminine, but not both. Instruments like the Terman-Miles AIA exemplify this approach, focusing on interests and attitudes that differentiated the sexes.

Androgyny and Orthogonal Measurement (Post-1970s): Instruments like the BSRI, developed by Sandra Bem, measure masculinity (instrumentality/agency) and femininity (expressiveness/communion) as **separate, orthogonal dimensions**. This means an individual's score on one dimension is independent of their score on the other. This innovation allowed for the classification of four distinct gender-role orientations: **Masculine** (high M, low F), **Feminine** (low M, high F), **Androgynous** (high M, high F), and **Undifferentiated** (low M, low F).

Trait Focus: The BSRI and PAQ categorize items based on specific traits: Masculine items typically relate to assertiveness, dominance, and independence (instrumentality); Feminine items relate to tenderness, compassion, and sensitivity (expressiveness). The final score reflects the degree to which an individual self-describes using these culturally valued attributes.

Self-Perception and Social Desirability: Most MFTs rely on self-report. A significant component of the testing process is the respondent's awareness of, and willingness to conform to, culturally appropriate responses, making the scores susceptible to **social desirability bias**. Items are often scored based on how desirable they are considered for one sex versus the other.

4. Theoretical Shift: From Bipolarity to Androgyny

The traditional MFTs were built upon a rigid sex-typing paradigm, which implicitly suggested that psychological health was correlated with conformity to one's traditional gender role. This model faced significant theoretical challenges throughout the 1960s and 1970s as clinical data suggested that restrictive adherence to gender stereotypes could be psychologically limiting. The inability of the bipolar model to account for highly flexible and psychologically well-adjusted individuals who possessed both "masculine" and "feminine" traits spurred a conceptual revolution.

The emergence of the **androgyny model**, championed by researchers like Sandra Bem, fundamentally redefined the goal of M-F measurement. Androgyny proposed that possessing a balance of instrumental (masculine) and expressive (feminine) traits was not only possible but often conferred greater behavioral flexibility and psychological adaptability, particularly in navigating diverse social situations. This adaptability was posited as a sign of greater mental health than rigid sex-typing.

This theoretical shift moved the focus of research away from simply diagnosing gender conformity toward studying gender-role flexibility and its relationship to psychological well-being. The new psychometric instruments, like the BSRI, provided the empirical means to test this hypothesis, demonstrating that androgynous individuals often scored higher on measures of self-esteem, cognitive complexity, and situational effectiveness compared to their traditionally sex-typed counterparts. This marked a profound turning point, rendering the older bipolar MFTs largely obsolete for research purposes.

5. Significance and Impact

Despite their limitations, Masculinity-Femininity Tests played a crucial role in the history of psychology, primarily by opening the door to the systematic empirical study of gender differences beyond simple biological categorization. The early MFTs provided the first quantitative data that allowed researchers to challenge traditional assumptions about the fixity of sex differences in personality and interests. They served as critical starting points for vocational and personality research for several decades.

The development of the Mf scales within major clinical inventories, although often misused, highlighted the relationship between gender nonconformity and certain psychological diagnoses, prompting further examination of social pressures and identity formation. Furthermore, the debates

surrounding the validity and utility of these tests directly contributed to the creation of modern gender psychology. Critics and reformers, using the weaknesses of the MFTs as evidence, successfully advocated for separating the constructs of biological sex and social gender, thereby legitimizing **Gender Studies** as a distinct academic field.

In applied settings, M-F measures were often used in **vocational guidance**. By assessing whether a person's interests and attitudes aligned more with typically male or typically female patterns, counselors attempted to predict job satisfaction in highly sex-segregated fields. For example, a high femininity score might suggest a better fit for roles in teaching or nursing, while a high masculinity score might point toward engineering or leadership roles. While such applications are now viewed as potentially reinforcing stereotypes, they represent a key historical impact of these psychometric tools.

6. Debates and Criticisms

The MFTs have faced extensive criticism since their inception, contributing to their eventual decline in prominence in contemporary psychology. These criticisms fall into several key categories:

Cultural Specificity and Temporal Instability: MFTs are inherently reflections of the cultural stereotypes and gender roles prevalent at the time and place of their construction. For instance, the interests deemed "masculine" in 1938 are drastically different from those in the 21st century. This means the tests quickly lose **validity** over time and cannot be reliably applied across different cultures or even different generations within the same culture, rendering them poor measures of stable, enduring personality traits.

Measurement of Stereotypes, Not Identity: A primary critique is that MFTs measure the degree to which an individual endorses or displays behaviors associated with **social stereotypes**, rather than measuring intrinsic gender identity or deeply held psychological dispositions. Critics argue that these instruments effectively assess conformity pressure rather than genuine psychological traits, failing to explain the underlying mechanisms of gender formation.

The Problem of Conflation: Early MFTs conflated several distinct variables: biological sex, sexual orientation, gender role, and personality. The MMPI Mf scale is a prime example, originally used to pathologize non-heterosexual men based on interests that deviated from the male norm. Modern psychology insists on rigorously separating these constructs to ensure ethical and accurate research.

Bias Against Androgyny and Undifferentiated Individuals: Even the later BSRI and PAQ models, while groundbreaking, faced criticism for assuming that high scores on both M and F (androgyny) were inherently superior to low scores on both (undifferentiated). Subsequent research suggested that the optimal profile often depended heavily on the specific context and cultural expectations, challenging the universal superiority of androgyny.

7. Contemporary Status

In current psychological practice and academic research, the use of unitary or bipolar Masculinity-Femininity Tests is largely abandoned. Modern psychological instruments focusing on gender tend to adopt a multidimensional and non-binary approach, recognizing gender as a complex spectrum influenced by social construction, individual identity, and biological factors.

The theoretical legacy of the MFT, however, persists in the recognition that certain psychological traits cluster according to social roles. While the specific instruments like the Terman-Miles AIA are historical artifacts, the conceptual framework they established--the difference between **agency (M)** and **communion (F)**--continues to inform personality research. However, these concepts are now typically decoupled from the rigid labels of "masculine" and "feminine" and are instead studied as components of general personality structures, acknowledging that these traits are distributed fluidly across all individuals, regardless of their gender identity or biological sex.

Further Reading

[Bem Sex-Role Inventory \(BSRI\)](#)

[Minnesota Multiphasic Personality Inventory \(MMPI\)](#)

[Terman, Lewis M.](#)

[Gender Role](#)