

MASCULINITY-FEMININITY TESTS

Authored by
mohammad looti

October 10, 2025

RECOMMENDED CITATION

mohammad looti (2025). *MASCULINITY-FEMININITY TESTS*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=41069>

MASCULINITY-FEMININITY TESTS

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

1. Core Definition

Masculinity-Femininity (M-F) tests are psychological instruments developed primarily in the mid-20th century to measure the degree to which an individual exhibits characteristics, interests, and attitudes deemed typical of men or women within a specific cultural context. These tests operate on the principle of comparing responses across large samples of men and women, retaining items that demonstrate the most significant sex differences in areas such as vocational preference, emotional expression, and general temperament. The underlying assumption of these scales is that masculinity and femininity are measurable constructs that exist along a continuum, allowing researchers to assign quantitative scores reflecting adherence to culturally established gender roles and traits.

2. Etymology and Historical Development

The history of formalized M-F testing begins with the landmark work of Lewis Terman and Catherine Cox Miles. Their seminal effort, the **Attitude-Interest Analysis Test**, commonly known as the **Masculinity-Femininity Test** (1938), was the first comprehensive instrument of its kind. Terman and Miles embarked on particularly extensive research to develop their scale, aiming to identify test items that successfully discriminated between the sexes across various demographics.

In seeking effective test items, Terman and Miles first searched the psychological literature for types of test material that yielded the most marked sex differences. They then prepared preliminary sets of items which were administered to diverse groups, including elementary, high school, and college students, unselected adults, professionals, and special groups such as athletes and adult homosexuals. This rigorous process led to the retention of items that produced significant statistical differences, resulting in a complex, seven-part test structure. This scale proved highly successful in differentiating between male and female groups across all age levels, from teenagers to octogenarians, providing considerable insight into the prominent sex differences observable in American culture at the time.

3. Key Instruments and Components

While the Terman-Miles test remains the most comprehensive, several other standardized tests integrated M-F scales based on fewer or more limited item sets. These subsequent instruments, though having value within their specific testing contexts, often exhibited low correlation with one another due to their specialized focus. These tests consistently utilized the common psychometric

principle of comparing responses that proved most characteristic of each sex in American culture.

Key instruments incorporating M-F scales include:

The Attitude-Interest Analysis Test (Terman and Miles, 1938): This seven-part test utilized a broad range of material to discriminate between the sexes, including word association, inkblot association, information tests, emotional and ethical attitudes, interests, opinions, and introverted response patterns. It sought a holistic measure of culturally defined gender traits.

The Strong Vocational Interest Blank (SVIB): This test included a masculinity-femininity scale focused specifically on interest patterns found to be highly characteristic of men versus women in vocational settings.

The Minnesota Multiphasic Personality Inventory (MMPI): The MMPI includes a scale designed to measure gender role adherence, often focusing on psychiatric and personality correlates of masculinity and femininity.

The Guilford-Zimmerman Temperament Survey: This inventory presents a yes-no questionnaire covering ten personality traits, one of which is the masculinity-femininity dimension. Characteristics indicative of **masculinity** included a disposition toward masculine activities, being hard-boiled, not easily disgusted, inhibiting emotional expression, and showing little interest in clothes and style. **Femininity** was characterized by interest in feminine activities and vocations, being easily disgusted, fearful, romantic, and emotionally expressive.

The Gough Femininity Scale (California Psychological Inventory, 1957): While based on a similar social-item approach as Terman-Miles, this later scale attempted to refine the measure by eliminating some of the more extreme or culturally specific social items, such as attendance at pool halls or beauty parlors.

4. Terman and Miles' Empirical Findings

Terman and Miles' extensive standardization research provided a detailed summary of observed sex differences within their standardization groups. They concluded that, from whatever angle they examined the data, the males evidenced a distinctive interest in exploit and adventure, outdoor and physically strenuous occupations, machinery and tools, science, physical phenomena, and inventions. Conversely, the females in their groups demonstrated a distinctive interest in domestic affairs and esthetic objects and occupations, preferring more sedentary, indoor, and direct ministrative occupations, particularly those focused on the young, helpless, or distressed.

These objective findings were supplemented by pronounced subjective and emotional differences. Males were consistently characterized as demonstrating greater **self-assertion** and **aggressiveness**, expressing more hardihood, fearlessness, and roughness in manners, language, and sentiment. Females expressed themselves as more compassionate and sympathetic, timid, fastidious, and esthetically sensitive. They were also judged as more emotional in general (or at

least more expressive of the four emotions considered) and often severer moralists, though they admitted to weaknesses in emotional control more readily than males. These findings collectively illuminated the more prominent gender differences codified within American society during the early 20th century.

5. Methodological Alternatives (Projective Tests)

The reliance of traditional M-F tests on explicit, consciously expressed attitudes and interests led some investigators to pursue entirely different psychometric approaches. A notable alternative was the **projective test** devised by Franck and Rosen (1949). This method operated on the assumption that gender differences manifest in subtle ways through fantasy and imaginative productions, aspects less susceptible to conscious social desirability bias than typical questionnaires.

In the Franck and Rosen test, the subject is presented with a series of simple lines and geometrical forms and instructed to draw a picture incorporating these elements. The resulting drawings are then analyzed and compared to productions found to be typical of representative members of the subject's sex. It is interesting that the scores derived from this projective method often failed to correlate with scores from the standard question-and-answer tests, suggesting they may be tapping into latent or unconscious aspects of masculinity and femininity that the social-type tests fail to reveal (Miller and Swanson, 1960).

6. Debates and Criticisms

The primary criticisms leveled against traditional M-F tests, such as the Terman-Miles scale, center on their extreme **cultural specificity** and their tendency to measure conformity to social customs rather than deep psychological or biological differences. Critics argue that the scores heavily reflect the particular culture and time period in which they were constructed, emphasizing social differences of the most extreme sort. This limitation is evidenced by the finding that only a few of the Terman-Miles items were found to differentiate between the sexes when the test was applied in other cultures, such as Holland.

Furthermore, these scales are criticized for giving undue consideration to conventional customs and cultural expectations while giving too little weight to underlying sexual and physical differences. This led to paradoxical results within their own standardization groups. For example, dressmakers, domestic employees, and women over sixty tended to rate highest in femininity due to their adherence to traditional interests. Conversely, a twenty-year-old woman, who might be at the height of her sexual attractiveness, would rate only moderately feminine if she happened to embrace common "masculine" interests such as sports. Similarly, a physically strong young man would rate only moderately masculine if he showed special interest in music or religion. In

essence, these tests were deemed to be overmeasuring social customs and underestimating the complexity of sexual and physical correlates of gender, leading to a focus on gender role stereotyping rather than underlying personality structure.

Further Reading

[Lewis Terman](#) (Wikipedia)

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

[Harrison Gough](#) (Wikipedia)

ARABPSYCHOLOGY.COM