1. Title of test:

16PF: Sixteen Personality Factor Questionnaire (5th edition)

2. Author(s)

16PF Authors: Raymond B. Cattell, A. Karen Cattell, & Heather E. P. Cattell

Manual Authors: Mary Russell, Darcie Karol

3. Publisher:

Institute for Personality and Ability Testing, Inc. (IPAT)

4. Date of publication


5. Grade or age levels covered

The 16PF 5th edition is intended for use with adults 16 yrs old and older.

6. Purpose of test or measure. Describe what the test is intended to measure and give your critical analysis of the tests ability to measure the intended construct(s).

The 16PF questionnaire is a broad assessment of normal adult personality. The purpose of the measure is to provide a comprehensive snapshot of an individual’s personality profile, by reporting on 16 primary personality factor scales as well as five broader global scales. This measure is supported by over 50 years of research and use in professional practice (Cattell, Cattell, Cattell, Russell, & Karol, 2002).
The 16PF can be used in a variety of settings. It can be used as a starting point in therapy, to give a counselor an overall view of a client’s personality profile and serve as a guide to an individualize counseling plan. The 16PF provides a measure of behavioral trends including anxiety and adjustment, which can facilitate diagnostic and treatment strategies. The 16PF may also be used as a way to evaluate management potential, facilitate leadership development and executive coaching, supplement individual and couple’s counseling, and inform career and vocational counseling. These varied uses are supported by a variety of scoring reports available for use with a 16PF assessment (Pearson Education, 2011).

As long as the caution of using the 16PF fifth edition as a measure of normal adult personality is maintained, I believe that the test measures the intended constructs. It is inappropriate for use in assessing abnormal personality. As the authors discuss, the 16PF is useful in predicting specific behavioral areas such as social skills, but does not address motivation behind behaviors (Rotto & McLellan, 2011).

7. a) How are the scores reported? b) What kind of scores does the instrument yield?

The 16PF records responses to 185 single items, which are responded to either true, ?, or false. This question mark is intended as a middle response for when neither of the other responses accurately represents the examinee but could be interpreted as a simple “I don’t know” (Rotto & McLellan, 2011). The Reasoning Factor appears at the end of the assessment and with separate instructions, as they are the only questions with correct and incorrect answers. The raw scores from these answers are converted to sten scores, a standardized measure which ranges from 1-10, with a mean of 5.5 and a standard deviation of 2. Sten scores between 4 and 7 are considered to be average, with outliers on either side. The extreme scores in a profile, those that fall outside of the average range, usually indicate the most distinctive traits of the examinee.
Numerous extreme scores indicate more distinctive expression of personality (Cattell et al., 2002).

The 16PF reports on sixteen basic personality scales. These scales are bipolar in nature, meaning that there is meaning ascribed to each end of the scale. The sixteen scales are:

- **Warmth**: Warm versus Reserved
- **Reasoning**: Abstract versus Concrete
- **Emotional Stability**: Emotionally Stable versus Reactive
- **Dominance**: Dominant versus Deferential
- **Liveliness**: Lively versus Serious
- **Rule-consciousness**: Rule-Conscious versus Expedient
- **Social Boldness**: Socially Bold versus Shy
- **Sensitivity**: Sensitive versus Utilitarian
- **Vigilance**: Vigilant versus Trusting
- **Abstractedness**: Abstracted versus Grounded
- **Privateness**: Private versus Forthright
- **Apprehension**: Apprehensive versus Self-Assured
- **Openness to change**: Open to Change versus Traditional
- **Self-reliance**: Self-Reliant versus Group-Oriented
- **Perfectionism**: Perfectionistic versus Tolerates Disorder
- **Tension**: Tense versus Relaxed

In the initial development of the 16PF the sixteen primary scales were derived using factor analysis from an extensive list of personality descriptors. From the sixteen primary scales, Cattell conducted a factor analysis to derive global factors on which the primary scales cluster.
together. These five global factors, also bipolar in nature, allow a simpler, broader view of personality that looking at the basic individual factor scales alone. These five global factors are Extraversion, Anxiety, Tough-Mindedness, Independence, and Self-Control (Cattell et al., 2002).

Finally, the 16PF reports three validity (or reliability) scales which help to account for response bias. These scales are the Impression Management Scale (IM), a bipolar scale which assesses whether or not an examinee is willing to admit behaviors that aren’t socially desirable, the Acquiescence Scale (ACQ), which measures an examinee’s tendency to agree with whatever statement is presented by choosing true regardless of whether or not the statement was actually factual about the responder, and the Infrequency Scale (INF), which is used to measure the degree to which examinee responses are likely to be the result of random responding (Cattell et al., 2002).

Scores are reported on one of numerous available score reports, including the Basic Score Report, which covers the basic sixteen personality scores. Numerous resource books are available to aid interpretation. Responses can be hand scored or computer scored by the publisher through mail in forms, software or internet-based services. More commonly used the Basic Interpretive Report (BIR), which, in addition to the basic sixteen personality scores, contains a profile of each global factor and the related primary factors, arranged from most extreme to average. Brief interpretive statements are provided for each of the global factors on which the examinee scores outside the average range; for average scores, scores are reported, but no narrative information is given.

In addition to the interpretive statements about global factors, The BIR also includes several additional sections. One of these sections is about the three response style indices measured by the 16PF, which serve as measures of reliability and response bias. A section about
criterion scores contains information based on the relationship of the examinee’s 16 basic scores with measures such as self-esteem, adjustment, and aspects of social skill. The vocational interests section presents occupational choices along the six general occupational themes according to Dr. John Holland’s typology, presenting a predicted sten score for each of Holland’s themes. These six themes are realistic, investigative, artistic, social, enterprising, and conventional, each relating to a set of interpretive statements regarding occupational interests relating to the themes. This vocational section includes interpretive statements. The relationship between the 16PF and these themes were examined by administering the 16PF to a sample group along side Holland’s Self-Directed Search, Form R. Finally, the Item summary section contains an item-by-item listing of the examinee’s response choices as well as a summary of statistics regarding those choices. The item summary also presents the examinee’s raw scores (Cattell et al., 2002).

8. How long does it take to administer? How does the length of administration impact the tests’ utility in counseling?

The manual indicates that 16PF can be administered in 35-50 minutes when administered by hand, and 25 to 35 minutes when administered by computer. The 16PF can be administered individually or in groups.

This test can be administered in less than the length of a typical counseling session or class period. In addition, particularly with the computer administration, it would be easy to administer this test with a larger group of students and get back information quickly. If the test was much longer, I would worry bout students’ ability to focus on the assessment without becoming disinterested.
9. How much does it cost to purchase and administer? How does the cost impact the tests’ utility in counseling?

The manual itself costs $55 when purchased from the publisher. In addition to the manual, various starter kits are available, depending on what use you have in mind for the 16PF, including starter kits for those using it as a tool in career/vocational counseling, human resources, or couples counseling. These starter kits range in price from about $100 to $150. Testing booklets are available in packs of ten for $23. Answer sheets are available in packs of 25 for $20-$22, depending on the quantity ordered. Mail in score reports (BIRs) cost from $20-$50 each, depending on the quantity submitted and the type of score report desired. Using a basic starter kit, administration would be between $50 and $75 per administration. Using additional material, assuming the manual is already owned, and mail in score reports, the cost per administration for an additional 25 examinees would be between $24 and $54 per administration, depending on the type of score report ordered (Pearson Education, 2011). Administration is also available via computer software and internet based services. For the BIR report, the computer administration costs approximately $25 per administration.

Cost predictably has an impact on the use of any given assessment. If the cost is to severe in comparison to the information gained from the scores of the assessment, the assessment is not valuable enough to account for the cost of administration. In the case of the 16PF, the cost is not necessarily prohibitive, though it might be difficult to use with larger groups.

10. What evidence is provided for validity (e.g., how is the information presented, is it sufficient to make a decision whether or not to use, etc.)?

In assessment, validity refers to a test’s ability to measure the concept it was designed to measure. The 16PF has been evaluated for validity in several ways. Most notably, the test is in its
5th edition, with over fifty years of use in practice and research to support its validity (Rotto & McLellan, 2011).

Construct validity demonstrates that the test measures sixteen distinct personality traits. When the test was being developed, factor analysis was used to “discover 16 primary factors that explain the larger domain of personality descriptors in the English language” (Cattell et al., 2002, p. 71). Cattell chose to use oblique factors, which are permitted to intercorrelate; the sixteen basic factors intercorrelate along global factors.

The 16PF uses several other assessments and measures to assess the validity of its factors. Scores on the sixteen primary factors can be correlated with similar personality constructs measured via other instruments. In its development, the 16PF fifth edition was compared to the Personality Research Form (PRF), the California Psychological Inventory (CPI), the NEO PI-R, and the Myers-Briggs Type Indicator (MBTI). While correlations existed between these tests and a variety of the basic factor scales found in the 16PF, the correlations were particularly strong with the 16PF’s global factors. For example, the 16PF extraversion global factor correlates positively with the MBTI Extraversion type and negatively with the Introversion type (r = .68 and r = -.61). The Reasoning Factor was validated separately, as this scale measured reasoning ability instead of personality. This factor was validated against the Information Inventory and Scale 2 of the Culture Fair Intelligence Test. The factor was found to be a “generally unbiased gauge of reasoning skill,” however, the author notes that it was not designed to be used as a measure of intelligence (Cattell et al., 2002, p. 90). The 16PF has additional validity in the form of a demonstrated ability to predict certain criterion scores, such as those in the domain of self-esteem and creative potential.
11. What evidence is provided for reliability? (e.g., how is the information presented, is it sufficient to make a decision whether or not to use, etc.)?

Reliability in testing and assessments relates to the consistency of a measure, including how well it accounts for statistical error. Most 16PF scales have a standard error of measurement that is close to 1 sten point. 68% of the time, the true score for an examinee will fall within 1 sten point of his or her obtained score. The true score will fall within two sten points of the obtained score 95% of the time.

The 16PF was evaluated for internal consistency using a test-retest model. Over a two week interval, a test-retest trail with 204 examinees had reliability for the sixteen primary factors ranging from .69 to .86, with a mean reliability score of .80. For the global factors tested with the same group of examinees, the two week test-retest reliability was higher, ranging from .84 to .91, with a mean reliability score of .87. The two month interval for test-retest was conducted with a sample of 159 examinees, resulting in a reliability score of .56 to .79 for the sixteen basic personality factors, with a mean of .70. On a two month interval, the reliability score for the global factors ranged from .78 to .82, with a mean of .78. As an additional measure of internal consistency for the 16PF is the use of Cronbach’s coefficient alpha, calculated on the norm sample. The coefficient alpha values ranged from .68 to .87, with a mean of .76 (Cattell et al., 2002).

The 16PF have three response style indices build into the construction of the assessment. The first of these scales is the Impression Management Scale (IM), a twelve item, bipolar scale that does not contribute to any of the primary personality scales. The IM scale evaluates social desirability, namely, whether a examinee is willing to admit less socially desirable behaviors. The Infrequency (INF) scale is composed of 32 items. A high score on the INF scale indicates
that the examinee responded to a “relatively large number of items” differently than the majority of people. Such a score might indicate random responding, inability to decide, or comprehension difficulty. The Acquiescence (ACQ) scale measures the tendency of an examinee to answer true to an item, regardless of the content of the item. A high score on any of these scales should be cause for an administrator to further review the possible causes of such scores (Cattell et al., 2002).

12. What kind of ‘norming’ information is provided? What is the make-up of the norming group?

The 16PF Fifth Edition was updated in 2002 with new norming information based on the 2000 U.S. Census. The initial sample was composed of 31,244 16PF fifth edition protocols taken from IPAT’s Test Services Division, received by the testing service between January 1, 1999, and May 15, 2001. After those protocols that did not report demographic data were eliminated, the sample size was 16,133. From this sample, simple stratification was completed using statistics from the 2000 U.S. Census to determine target numbers based on gender, race, age, and education level. This resulting sample size was 10,261 individuals.

This norming sample is composed of 5,124 males (49.9%) and 5,137 females (50.1%). The age of those in the sample ranged from 16-82 years old, with a mean age of 32.7 years. The majority of individuals included in the sample (75.3%) had at least some college education. Racially, the sample was made up of individuals that identified as 77.9% Caucasian, 10.8% African American, 3.6% Asian American, 0.8% American Indian, 1.5% Multiracial and 5.4% as other races. 8.6% of the sample identified as having Hispanic origins (Cattell et al., 2002).

The reported percentages of the norming sample were similar to those of the 2000 U.S. census, particularly among racial and gender dimensions, where the difference between the
percentages in the U.S. population and the sample population were often within a single percentage point. The difference in percentage between the two groups is compared below:

<table>
<thead>
<tr>
<th>Population</th>
<th>Percent in Sample</th>
<th>Percent in U.S. 2000 Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>49.9</td>
<td>49.1</td>
</tr>
<tr>
<td>Male</td>
<td>50.1</td>
<td>50.9</td>
</tr>
<tr>
<td>White</td>
<td>77.9</td>
<td>77.4</td>
</tr>
<tr>
<td>Black/African American</td>
<td>10.8</td>
<td>11.4</td>
</tr>
<tr>
<td>American Indian</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>Multiracial</td>
<td>3.6</td>
<td>3.8</td>
</tr>
<tr>
<td>Other</td>
<td>5.4</td>
<td>4.7</td>
</tr>
<tr>
<td>Hispanic Origin</td>
<td>8.6</td>
<td>11.0</td>
</tr>
</tbody>
</table>

The demographics of the sample as compared to the census were less consistent when evaluated in terms of age and educational level. Overall, the sample population was significantly younger than the U.S. Census figures, with the elderly (ages 65 and older) significantly underrepresented. For instance, the percentage of examinees in the norming group from 15 to 24 years of age was more than double that reported by the 2000 U.S. Census (36.2% in the norming group versus 18% according to the census), while the percentage of individuals ages 65 years or older according to the census was more than twenty times larger than the percentage reported in the norming group (16.0% according to the census versus 0.7% in the norming group). The age group percentages are presented below:

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Percent in Sample</th>
<th>Percent in U.S. 2000 Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-24 yrs.</td>
<td>36.2</td>
<td>18.0</td>
</tr>
<tr>
<td>25-44 yrs.</td>
<td>41.7</td>
<td>38.0</td>
</tr>
<tr>
<td>45-54 yrs.</td>
<td>15.7</td>
<td>17.0</td>
</tr>
<tr>
<td>55-64 yrs.</td>
<td>5.6</td>
<td>11.0</td>
</tr>
<tr>
<td>65 yrs and older</td>
<td>0.7</td>
<td>16.0</td>
</tr>
</tbody>
</table>

The norming group also had a significantly higher level of education than the 2000 census indicated in the U.S. population. For instance, 47% of individuals in the norming group reported being college graduates, while only 30.5% of the respondents to the 2000 census
indicated the same. While 19.8% of census respondents reported some college education, 28.3% of the individuals in the norming group reported some college education, a difference of more than ten percent. The education level percentages are reported below:

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Percent in Sample</th>
<th>Percent in U.S. 2000 Census</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School Grad or Less</td>
<td>24.7</td>
<td>49.7</td>
</tr>
<tr>
<td>Some College</td>
<td>28.3</td>
<td>19.8</td>
</tr>
<tr>
<td>College Grad.</td>
<td>47</td>
<td>30.5</td>
</tr>
</tbody>
</table>

The disparity in percentages along the lines of education and age raise some flags of caution for administrators of the 16PF, particularly when administering the test with individuals that were significantly underrepresented in the norming group, such as those with a lower level of education or those over 65 years of age. It is possible that the assessment results will be less reliable in these populations.

The updated fifth edition also reports sex specific norms for three of the sixteen basic personality factors. On the factors “Warmth,” “Sensitivity” and “Apprehension,” it was found that males tended to report differently than females, which was not the case with the other factors. The gender-specific and combined gender norms are available as scoring options for these three factors.

13. Is the test appropriate for multicultural populations? (expand your consideration beyond just race/ethnicity to other populations/aspects of identity as well). What evidence is provided (or not provided) upon which to make this decision?

The 16PF fifth edition makes an effort to include aspects of identity such as race and ethnicity in the norming group. The publisher also notes that the questions contained in the fifth edition were designed to be non-invasive, meeting the Equal Employment Opportunity Commission (EEOC)’s requirements, as related to the Civil Rights act (1990) and the Americans
with Disabilities Act. The aim of this design was to ensure that test result differences among different ethnic, gender, and age groups were minimized (IPAT).

Several elements of the 16PF consider multicultural issues. The Rule-Consciousness factor measures “the extent to which cultural standards of right and wrong are internalized and used to govern behavior” (Cattell et al., 2002, p. 46). Scoring high in this section does not necessarily account for the cultural values of the client’s culture, but the values endorsed by the majority of “western” culture, “having remnants of the Protestant Puritan ethic” (Cattell et al., p. 46). This measurement severely limits the interpretation of scores for individuals from minority cultures. The 16PF is designed to be easy to read and understand, with an overall readability at the fifth grade level. However, this makes the test less accessible to those for whom English is a second language. In addition, the manual notes that because of the verbal nature of the reasoning scale, individuals who speak little English could be expected to obtain a lower than accurate score. The 16PF manual also notes that the Vigilant factor does not account for life circumstances that may impact the adoption of a vigilant stance, often associated with members of oppressed minority groups (Cattell et al., 2002).

14. What special competencies does the administrator need to administer, score, and interpret this test, based on the author(s) recommendations?

The 16PF fifth edition is, in the words of the manual, “virtually self-administrable,” (Cattell et al., 2002, p. 7) indicating the ease with which a counselor can use the assessment. The 16PF can be used in a variety of settings, for a range of purposes, by a variety of professionals “who understand the limits of psychological testing and interpretation and possess a fundamental understanding of psychometrics” (Cattell et al., 2002, p.105). Those administering the 16PF should also be familiar with the Standards for Educational and Psychological Testing as
presented by the American Psychological Association. Specific training exists for those who would like to become experts at using the 16PF, but is not required by the publisher before use of the instrument. In order to purchase the instrument from IPAT, IPAT requires the competition of a user qualification form. In order to receive materials from IPAT, they ask that professionals provide the organization with “credentials regarding licensure, education, training and experience, and/or affiliation that supports the professional’s stated ability to use materials for their intended purposes” (IPAT user qualification form, n.d.).

The authors of the manual recommend that the administrator stress to examinees the importance of not languishing over answers, instead encouraging examinees to pick the response that initially seems most suited to them. The manual suggests it might be appropriate to remind examinees of this at least once during the administration process. The administrator should also be cautious in scoring and interpreting the score, paying particular attention to the reliability scales that could indicate reporter bias. In interpreting the score, the manual also notes that the BIR should not be given directly to the examinee without the aid of a professional to explain such results (Cattell et al., 2002).

15. Is this a test you plan to use in the future? Why or why not? If you will use it – discuss how. If not, describe the relevant cautions for counselors who do use it.

It is very possible that I will use the 16PF in the future as a professional counselor. After completing this evaluation, I believe that the assessment has been significantly validated, both in its current edition throughout its history of assessment. However, I will hesitate to use this assessment with some minority groups, as well as with groups from cultures that differ significantly from the norming group, that is, Americans in the 21st century. As I will be working as a school counselor, my use of the test will also depend on the cost of administration, though I
believe the BIR provides a wealth of information that would be valuable to a school counselor and an adolescent examinee considering post-high school options. I would be more likely, however, to use the 16PF Adolescent Personality Questionnaire, as the manual points to that particular assessment as most appropriate for ages 11 to 22.
References


