A Comparison Across Personality Assessments

2020 CONSUMER SYNOPSIS
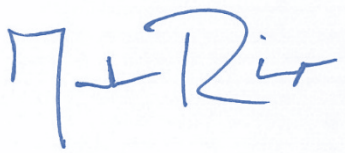
SOLVE

People Problems Solved.

Do a Google search for "best hiring assessment" or "most accurate leadership assessment" and you get a plethora of anecdotal recommendations on assessments. The most popular assessments are consistent: Myers-Briggs, DiSC, StrengthsFinder. These are also the assessments most business professionals have completed if they have been around long enough. I get a range of experiences regarding accuracy of these assessments – from love to hate. The one thing that is most consistent with the input I receive is its lack of consistency. Psychometricians refer to that as reliability (more on that in a moment).

I have spent much of my 20-year career advising organizations on hiring decisions. To date, I have researched over 100 assessments. Yet, every year, I am introduced to something new. Each assessment promises amazing accuracy – heralding itself as the best. However, there is rarely evidence to support the claim, and the evidence that does exist is often suspect.

What is a leader or human resource professional to do without unbiased information? Most typically, they go with the known brand. As a former colleague once told me, you rarely get fired when you choose a known brand. I am not inclined to argue the merits of that position. Rather, I hope to provide a guide that will help leaders and HR professionals identify the best assessments without needing to obtain a doctorate in psychometrics. Many of the great assessments are unfamiliar to those outside the world of organizational psychology. I hope to expose you to the rich opportunities that exist, which can enhance your hiring practices.

Why me? First, assessing talent is my superpower. I am not perfect, but my accuracy is north of 85% for leadership roles. In fact, I have guaranteed my accuracy to convince new clients to experience our approach because I have confidence it works given the hundreds of leaders we have assessed. Second, I am assessment agnostic. Our company, SOLVE, does not receive compensation for recommending assessments. Additionally, SOLVE does not sell its own assessments widely (I will not tell you mine is the best). Over the past three years, a handful of organizational psychologists from our firm reviewed technical manuals, combed the internet for facts on each assessment, and debated the merits and rigor of the claims the assessment vendors made. What follows is guidance on how to critique and, ultimately, choose the best "personality" assessment.

Enjoy!

Merle Riepe, PhD
President
SOLVE
www.solve.hr

Nine Things to Know Before Using a Personality Assessment

Before you begin assessing candidates, it is best to conduct due diligence on your assessment. First and foremost, you want to ensure your assessment is accurate – why pay for something that does not work?! Second, you need to avoid unfair hiring practices and reduce liability. Some assessments may inadvertently support biases that lead to adverse impact (a discriminatory effect on a protected group). As you attempt to diversify your employee base, you do not want an assessment that supports the biases you are attempting to quash. In addition to these, there are seven other questions to consider before incorporating a personality assessment into your hiring, promotion, and/or succession planning process.

Before addressing each aspect, it is important to present a few caveats. First, I am only considering "personality" assessments in this review. There are assessments for IQ, EQ, communication-style, values, you name it. In an effort to compare apples to apples, we restricted our review to the most popular personality assessments (or the ones most people asked about during conference presentations when I would share my expertise on the subject – "Hey Merle – is assessment XYZ good?"). Of course, good is relative. In so far as personality assessments are concerned, they are generally not as predictive as general mental ability or a strong structured interview. Having said that, the assessments we recommend are more accurate in predicting a candidate's performance than age, work samples, job knowledge tests, or years of experience. Second, a perfect prediction of a candidate is unattainable. If I had a nickel for every time I heard, "we tried that assessment and it was way off," I would have a lot of nickels. We are talking about intensely dynamic human beings and attempting to categorize the mosaic of traits that encompass their being. No assessment is perfect. No interviewer is perfect. Better than most (or at least better than chance) is our goal when utilizing hiring tools. Additionally, there are other factors at play in an employee's performance that an assessment cannot account for: an ineffective boss, disengaging work environment, poor onboarding, available candidates in the job market, life events, etc. In the end, we are not hiring the "best" or "right" employee, we are hiring the best candidate among the alternatives available.

As you consider your current (or a future) personality assessment, here are the questions you should be asking:

1. *What is the purpose?*
2. *What is the desired outcome?*
3. *How accurate is the assessment?*
4. *Are the scores comparable?*
5. *What level of specificity will I receive?*
6. *What is the candidate experience?*
7. *How much does it cost?*
8. *Are there any legal considerations?*
9. *What level of value am I receiving?*

1. ***What is the purpose?***

    Will you be using the information to hire and/or promote the individual? Or, is it strictly for development purposes? If the latter, the need for scrutiny is less important; however, I implore you to review *all* assessments with rigor. Please do not settle for an inferior tool simply because it is "only for development." Your internal leaders deserve accurate information as they begin to explore their career opportunities – it is not the time to go cheap. For development-based processes, assessments should be (1) aligned with company values and leadership needs, (2) accurate, with strong reliability and validity, but not as high of a threshold as is required for hiring/promotion, and (3) simple to interpret, comprehend, and activates behavioral change (I can improve *this* or be more aware of *that*).

    If you are using an assessment for hiring, promotion, and/or succession planning decisions, you owe it to your company's stakeholders and employees, the candidates, and your professional reputation to use sound, proven tools. Otherwise, your simply adding noise to an already "loud" process, and further distracting decision makers from the signal.

2. ***What is the desired outcome?***

    Are you looking to screen out or screen in candidates? Personality assessments are more effective at separating great candidates from good candidates (i.e., screening in). If your goal is to screen out candidates early in the process who are unlikely to be effective in the role or generally disruptive, we recommend structured screening interviews and/or biodata assessments.

3. ***How accurate is the assessment?***

    Let us cut to the chase – why use an assessment that has the same level of accuracy (or marginally better) than a coin flip? A penny and a conditioned thumb are a lot less expensive and contain fewer distraction points (i.e., you know the coin is simply chance; with an assessment, we assume accuracy because we perceive it to be and is backed as "scientific" and/or "valid" via sales and marketing).

    We must understand an assessment's reliability and validity. Reliability is the consistency of the assessment – the dependability by which the hammer strikes the same spot. Validity (predictive validity to be specific) is the accuracy of the assessment in predicting your desired outcome – ensuring the hammer is striking the nail head and not your thumb. This information *must* be known before you use an assessment. Many vendors are slow to share this information. They highlight the strengths and hide the weaknesses, and/or claim "intellectual property" to justify withholding information. Given so much uncertainty, SOLVE has done the research for you. Quickly pass on those who do not provide this information.

4. *Are the scores comparable?*

A score without a relative comparison is useless. This is the predominant flaw in most strengths-based assessments – they cannot answer the question, *a strength compared to what or whom?* For example, I competed in several sports in high school and college. My strength was basketball. I was better than many, and I can name thousands upon thousands of others who are much better. What is the opportunity cost I would have incurred by focusing on basketball versus something that was a more unique strength with higher demand and prospect? The best assessments will provide both normative and scaling data. Normative data allows you to compare candidates to a group of leaders, accountants, call center representatives, etc. who have completed the assessment previously. Scales display the level of intensity, which may or may not include normative data, and allow you to make more accurate comparisons across candidates.

5. *What level of specificity will I receive?*

Granularity is very important for hiring and promotion decisions. Without specific recommendations, we are left to our own thinking to make judgments about how a person will behave. For humans, our thinking is prone to inertia and inaccuracies caused by cognitive biases. Information provided by assessments allow us to examine our biases and produce a healthy skepticism, which improves the accuracy of our thinking and the decisions we make. Does the report and feedback you receive from the assessment, challenge your thinking? Or does it most frequently confirm your intuition and biases? Many assessments provide reports that are equivalent to corporate horoscopes – written in a way they are easily misinterpreted and/or read as we hope the candidate to be. Finding an assessment that provides feedback specific to work behaviors and can forecast job outcomes is critical. Most organizations do not use these advanced assessments (yet) and incorporating them into your hiring process provides you an upper hand on your competition. If you are not experiencing counter-intuitive information and/or feelings from candidate assessment results, you deserve better.

6. *What is the candidate experience?*

This is the most overlooked criteria when considering an assessment. The assessment does not need to be "fun" or "entertaining" to be effective. It does not have to be innovative or unique. It needs to be job relevant and duration appropriate to the role. You want the assessment to measure relevant job information where candidates intuitively make the connection as to why it is important (i.e., face validity). Avoid broader personality assessments and adjective checklists because they are not as clear in their connection to the job. Work style assessments that ask candidates to respond to statements like, "*I enjoy presenting to groups*" appear more relevant than circling your top five adjectives or responding to abstract statements like, "*I know how to captivate people.*" Regarding duration, the time should be relative to the risk of a poor hire – more risk requires more data, which takes more time. Executives are accustomed to spending up to two hours completing online assessments, individual contributors around one hour, and no more than 20 minutes for entry-level positions.

7. *How much does it cost?*

This should not be a deciding factor, yet it continues to be an important factor for most. The most expensive personality assessments are less than $500, and that is rare. They tend to range between $25 and $100. In most cases, you get what you pay for. If an assessment is free to download and use, it is worth nothing. In fact, there is likely more risk than benefit. Many studies show the total cost of losing an employee can range from tens of thousands of dollars to 1.5-2X annual salary. Selecting an assessment based on price is equivalent to stepping over dollars to pick up a dime.

8. *Are there any legal considerations?*

Fortunately, most personality assessments have not demonstrated adverse impact, which will keep you off the OFCCP's naughty list. There are some assessments on the market that test for cognitive ability, IQ, decision-making acumen, and the like. Those sections of the assessment are likely to demonstrate adverse impact. The prevalence of adverse impact alone does not make an assessment "illegal." However, if you choose to use such an assessment you must have documentation demonstrating its job relevance. Local validity evidence is preferred, and do not rely on the vendor's technical manual to be sufficient. The EEOC is looking for *professional judgment based on accepted corpus of knowledge*. If you are unsure, call an expert. Some employment attorneys are knowledgeable, many organizational psychologists are experts, and either (or both) would be able to advise you.

9. *What level of value am I receiving?*

Knowing your purpose for using an assessment is critical to answering this question. The data from the assessments is not as valuable as the interpretation of the data from the assessment. Most assessments have too many data points. Hiring managers become overwhelmed and resort to searching for (or more easily recognizing) the data that confirms their biases (via subjective validation or availability heuristic).

Assessments fall into one of three value categories: off-the-shelf, semi-custom, and fully custom. *Off-the-shelf assessments* have an upper limit of accuracy, which is represented by the validity coefficient. *Fully custom assessments* can be optimized to the role and/or population you are measuring, and their accuracy is almost impossible to beat. Correspondingly, expect to invest six figures to develop the assessment. For high turnover positions with large volume (i.e., call centers, caregivers, entry-level roles), the value is worth it given the return in reduced turnover, absenteeism, theft, etc. *Semi-custom assessments* provide a middle ground that allows for additional accuracy beyond an off-the-shelf assessment but without the investment of a fully custom assessment. The additional accuracy comes from extra specificity. In these situations, consultancies will match the job with the role by way of focus groups and/or reviewing job descriptions. This added layer allows for custom

reporting of the assessment data, which can remove some of the biases inherent in the process, thereby, increasing the accuracy. But, remember, you must start with an assessment with strong validity evidence before customizing.

After considering these questions and selecting an assessment, it is important to track the effectiveness of the assessment over time. Many vendors do not provide this service, and those that do, typically charge tens of thousands of dollars to demonstrate their value to you – their customer. Again, organizational psychologists are a good resource to help with your data tracking, analyses, and interpretation. With data in hand, you can be confident you get the greatest value from your personality assessment, or know when it is time to make a switch.

A Comparison Across Personality Assessments

These personality assessments are evaluated and rated regarding their use in hiring and promotion decisions, not for employee development.

ASSESSMENT	OVERALL	PRICE	TIME (MINUTES)	RELIABILITY CRONBACH'S α	TEST-RETEST	COMPARABILITY SCALES	NORMS	VALIDITY CONCURRENT	PREDICTIVE
WAVE Professional Styles	98	$$$	40	🟢	🟢	🟢	🟢	🟢	🟢
WAVE Work Strengths	82	$$	20	🟢	🟢	🟢	🟢	🟢	🟢
OPQ32i	82	$$$$	60	🟢	🟢	🟢	🟢	🟢	🟢
NEO PI-3	75	$$	45	🟢	🟢	🟢	🟠	🟡	🟢
ProfileXT	46	$$$$	60	🟢	🟢	🟢	🟢	🟡	🟡
Hogan Personality Inventory (HPI)	45	$$$	20	🟡	🟢	🟢	🟢	🟡	🟡
Birkman Signature	45	$$$$	30	🟡	🟢	🟢	🟢	🟡	🟡

ASSESSMENT	OVERALL	PRICE	TIME	RELIABILITY		COMPARABILITY		VALIDITY	
			MINUTES	CRONBACH'S α	TEST-RETEST	SCALES	NORMS	CONCURRENT	PREDICTIVE
16PF	43	$$	50	🟢 green	🟢 green	🟢 green	🟢 green	🟠 orange	🟡 yellow
Caliper	42	$$$$	75	🟡 yellow	🟡 yellow	🟢 green	🟢 green	🟠 orange	🟡 yellow
Predictive Index Behavioral Assessment	15	$$	20	🟢 green	🟠 orange	🟢 green	🟡 yellow	🟠 orange	🟡 yellow
FIRO-B	14	$$	15	🟢 green	🟢 green	🟢 green	🔴 red	🔴 red	🟡 yellow
EverythingDISC	14	$$	15	🟢 green	🟢 green	🟢 green	🟢 green	🟡 yellow	🔴 red
StrengthsFinder	11	$$	40	🟠 orange	🟢 green	🔴 red	🔴 red	🟠 orange	🟠 orange
CPI 434	10	$$	60	🟢 green	🟡 yellow	🟠 orange	🟢 green	🟡 yellow	🔴 red
RembrandtAdvantage Portrait	5	$$$$	40	🟡 yellow	🟡 yellow	🟠 orange	🟢 green	🔴 red	🔴 red
TTI Behaviors	4	$$$	30	🟢 green	🟢 green	🟠 orange	🟠 orange	🔴 red	🔴 red
Kolbe RightFit	4	$$$	10	🔴 red	🟡 yellow	🟠 orange	🟡 yellow	🔴 red	🔴 red
Myers-Briggs Type Indicator (MBTI)	3	$	30	🟡 yellow	🟡 yellow	🔴 red	🟡 yellow	🔴 red	🔴 red

KEY:

VERY POOR   POOR   MODERATE   STRONG   EXCELLENT
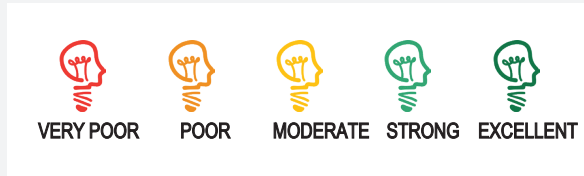
For the latest ratings
and more information visit
SOLVE.hr/hiring-the-right-people/

RELIABILITY: The consistency, or replicability, of the assessment.

CRONBACH'S α: A measure of reliability to test the consistency of the test items.

TEST-RETEST: A measure of reliability obtained by administering the same test twice. The scores from test 1 and test 2 can then be correlated in order to evaluate stability over time.

COMPARABILITY: Designed to assist score interpretation.

SCALES: Set to distinguish different score ranges.

NORMS: How a given score compares to others in a similar group.

VALIDITY: The accuracy of an assessment.

CONCURRENT: Validity evidence suggesting the results can accurately predict performance of job incumbents.

PREDICTIVE: Validity evidence suggesting the results can accurately predict performance of actual hires.

SOLVE

People Problems Solved.