

## Chapter 6:

# EXPLORING THE WORLD OF ASSESSMENTS

Differentiating the Simple from the Sophisticated

## GUESS OR ASSESS!

Assessments are a key component in evaluating and understanding people's behavior and play an integral part in assessing and developing leaders. The right assessments can reveal how people cope when faced with challenges, demonstrate their level of emotional resolve, self-awareness, and level of emotional intelligence. Assessments can also be used to show the depth of a person's critical thinking skills and problem-solving ability. This insight about an individual's personality, behavior, emotional intelligence, and critical thinking skills is essential not only to support the leadership development and succession planning process, but also to serve as an integral part of the selection process. You can't afford to make a bad hire and you cannot pick out problematic behaviors from an interview. In the selection process an external candidate is relatively unknown because there is no baseline as you would have from an employee within your organization. Understanding and using the right assessment tool can give you more insight into key behaviors as well as those problematic behaviors you are unable to uncover in other ways. Likewise, understanding and using the right assessment tool is an important aspect of a highly successful leadership development program as well as a worldclass selection program.

Let me emphasize that this chapter is more technical than others in the book. This is necessary because of the complexity of this subject. The material here will help you understand the differences and nuances among the majority of assessments available on the market. You will realize that the vast majority of assessments cannot delve deeply enough into personality and behavior to provide the depth of information necessary to support the talent management process and the organization's leadership development initiatives. In chapter 5, I discussed the value of using assessments as part of the feedback process for executive development. The information here will help provide more insight into that process and show you how to distinguish the differences between the simple and the sophisticated assessments.

It is unfortunate, but there is no rating service for assessments. The auto industry has its J. D. Power and Associates, but there is no such resource that compares and contrasts assessments. The main body of knowledge rests with the Buros Institute's *Mental Measurements Yearbook*, the independent review source for finding any information about a particular assessment. Any independent review of various assessments is valuable only to the degree to which the test publisher is willing to provide its validation and reliability studies. As you will learn, some test organizations aren't willing to release this data. While there are resources that critique various assessments, there is no resource that compares one assessment to another. My goal here is to review the key components of assessments so that you understand which one is best for your purposes. Many people are unaware that the vast majority of assessments cannot fully assess a person's personality to be effective in supporting the talent management and leadership development initiatives, let alone use them as an element of a dynamic selection model. Most people are predisposed to continue using a tool with which they're already familiar. Chances are they're using an assessment that isn't measuring what they need it to do and they're not even aware of this shortcoming. Any assessment, and I mean any, will hit the mark and be spot on in what it will reveal about yourself and anyone else. When people read their assessment results from the Myers-Briggs/MBTI or any other, nine out of ten people agree that it is reflective of who they are. The DiSC, Birkman, Predictive Index/PI, Profiles XT, Caliper, MBTI, and all the others will all be spot on with the data. Why? These tests measure what they say they will measure and, because they are all validated, meaning they will measure the traits they claim to measure, there is more than a 90 percent possibility that you will identify with the report you read.

This is a really important point to understand about assessments: it is not what an assessment tells you, it is what it *doesn't* tell you. What I mean by this is that almost every assessment report on the market produces what I call "soft language." This means that it mainly presents all the pertinent behavioral data in a positive light and is expressed in such a way that it is difficult to tell if there is any problematic behavior present.

The next thing to understand even beyond that point is what an assessment *can't* tell you. What this means is that some assessments do not measure certain traits that others do. Therefore, you will discover why I say it is what an assessment *can't* tell you that is an important distinction. This will become more understandable as I explain this in more detail later on. This information is critical because of its significant impact on using assessments for leadership development.

One of the interesting things I have observed is that some organizations are so entrenched with the assessments they use currently that even when they find out their current tool is insufficient, they don't want the hassle and inconvenience of introducing a new assessment or changing midstream. I've seen this reaction several times. If you are going to use an assessment for leadership development or even selection, you must recognize how you're using it and its strengths and limitations. If it's not providing the information you need, you must change to another system. If assessments can provide additional feedback to an executive for development and can help an organization gain an understanding of an executive's capabilities and potential, then using them should be a logical thing to do. That means that the assessments and the process you use should be highly capable of predicting an individual's behavior, and their critical thinking skills as well. The problem is that very few people understand the difference between a simple versus a sophisticated instrument and what the differences are between them as to what is measured. That is why you must differentiate between the various assessments that are available. Doing so requires someone who knows what questions to ask about a particular assessment test.

I have studied and researched assessments for nearly thirty years. I have been using them in my practice since 1991, and I have conducted nearly 13,000 assessments in my career. Over the years, in my work helping organizations build a selection and leadership development model, I have realized that using assessments is one of the most misinformed, misunderstood, and underutilized tools in the talent management and organizational development process. I also have evidence that they are just as misunderstood in the selection process, probably even more so. It is this firsthand experience that has led me to know how to understand the differences between assessments. That is what I hope to teach you so you know which assessments you should be using and why. I will also help you see their connection to leadership development, talent management, and selection. More often than not, is not enough to evaluate and assess the talents and abilities of an individual with just one assessment.

A properly used assessment process can open up a dialogue with an individual about his behavior and tendencies which may be negatively impacting his work performance. This information would not be evident in a conversation because people with extreme egos, and especially those with narcissistic tendencies, do not like to talk about their weaknesses. Some people don't think they have any weaknesses while others aren't willing to admit to having them. Then there are those individuals who are just simply in denial. Therefore, using assessments in the process of leadership development and especially executive coaching is essential.

## Understanding Critical Thinking Skills

Beyond measuring behavior and emotional intelligence, it is also necessary to evaluate and measure the depth of a person's critical thinking skills and problem-solving ability. While this chapter is about utilizing assessments to evaluate personality, behavior, and emotional intelligence, it is important to discuss the use of assessments to measure critical thinking skills. As an individual advances in an organization, the more essential critical thinking skills, problem solving, and intelligence become. Often, the lack of effective problem-solving skills is a derailing factor for an executive and a contributing factor to the Peter Principle. We can use our leadership examples from Coca-Cola, Mattel, and Enron as classic examples of the Peter Principle.

To help you get a baseline to measure critical thinking skills, there are a number of different assessments that measure various types of intelligence. Just like personality assessments, not all problem-solving assessments measure the same things. Perceptual reasoning is one form of intelligence that can be measured, but is rarely done. Perceptual reasoning is an individual's natural ability and it is often referred to as IQ or intelligence quotient. The reason it is rarely done is because to accurately measure this you must proctor this particular assessment because there are four timed tests as part of the test itself. Today people want an online process that is quick and efficient and this one is time-consuming and cannot be done online. The mean score for IQ is 100 for the general population. A person's perceptual reasoning is independent of education, gender, or ethnic background. In one of our leadership studies with eighty high potential managers, the mean score for the group for IQ was 110. This means that the average high potential manager was 10 points above the mean.

There are only a handful of behavior and personality assessments on the market that include some sort of critical thinking or problem-solving skills component within their assessment. There is one test that has an IQ component that measures perceptual reasoning and that is the Culture Fair. A true IQ test is a separate test in and of itself and uses spatial recognition, not words or numerical reasoning. Neither the MBTI, DiSC, nor Predictive Index/PI, to name a few types of assessments, has any problem-solving component embedded within them, let alone a perceptual reasoning or IQ component. Later in this chapter I will discuss the 16PF<sup>®</sup> Questionnaire, one of the more sophisticated assessments on the market. It is the primary instrument of choice in our assessment model. It has a problem-solving component within the assessment, which is as reliable as the behavioral portion. It is not a perceptual reasoning/IQ test; it is a verbal and numerical reasoning test. When we administer our executive battery, we include a complete perceptual reasoning test. While it is valid and reliable in predicting a person's problem-solving ability it does have its limitations. It measures a person's ability to solve math problems and analogies where all the needed information is included and it does provide some insight into the person's problem-solving ability. However, the 16PF Questionnaire does not measure other kinds of problem solving that other cognitive assessments do. That is why we use several different cognitive assessments.

When we evaluate more senior-level hires or conduct internal evaluations, where critical thinking is an essential part of the role, we use three additional assessments that measure different forms of intelligence and problem solving as part of our managerial and executive assessment batteries. These added tests are always used in our leadership development and selection model for higher-level positions for selection, development, coaching, or promotion. These three additional tests include one that measures IQ (perceptual reasoning), another measuring inductive and deductive reasoning (indicates how well they apply logic to solving problems), and the other measuring rate of speed of processing and learning agility. These three additional assessments help us gain a full awareness of an executive's critical thinking skills, problem-solving ability, perceptual reasoning, and intelligence, regardless of the position.

## The Significance of Behavior and Performance

Behavior is the single biggest predictor of performance more than job skills, experience, education, and even critical thinking skills. While the other factors are relevant and have a bearing on job success—especially in higher-level roles, where critical thinking skills and intelligence come into play—behavior weighs in the most. Recall the failed executives mentioned in Chapter 1 and you can see that some underlying behaviors had as much

to do with their downfall as their lack of competencies and judgment. In addition to behavior, people's emotional intelligence is a critical factor. It is important to understand how they manage their emotions, cope with stressful situations, interact with people, and how much self-insight they have about their behavior, both good and bad. There are only a handful of assessments that can measure and uncover these traits in depth.

### Insight – What Can Be Revealed?

Here is an example of how an assessment can provide further insight about someone. In a conversation with an HR manager, we were discussing an individual's need for some executive coaching. The discussion revolved around the person being full of himself and also becoming stressed and emotional. If this is the case it would not be a surprise to find that a 360 assessment and peer interviews would produce the same information. If the person's assessment results reveal that he needs to be the center of attention, is overly dominant, and has a low tolerance for stress, you can begin to create the feedback loop to help him see what he can't see for himself and may need to work on.

The proper assessment and, more importantly, a battery of quality assessments can help assess the difference between someone who is full of himself because he is smarter than other people and needs to see how innocent celebrations might make others see him as egotistical, versus someone who is truly egotistical and needs to be reminded of his imperfections. There are only a handful of assessments that can make the distinction between the two. The more sophisticated assessments can provide this kind of valuable information and measure any extreme behaviors that may be problematic and potentially derail an executive. In chapter 12, "Case Studies," I provide examples that show how you can combine various assessments to gain this powerful information. The case studies will reveal how utilizing a battery of assessments (several assessments that complement one another) can point to areas of concern that otherwise might go undetected.

Another advantage to using a more sophisticated assessment is that you have the added value of predicting future job success that will support the succession planning process. This is not something that can be achieved using any of the four-dimensional assessments that will be discussed later in this chapter. In one of the case studies I detail in chapter 12, we cautioned a client about promoting an individual. The assessments showed that while this employee was gifted, her talents would likely be overshadowed by some very difficult challenges present within her behavior. The client did, in fact, promote this individual only to terminate her within six months. In the detailed case study, you will see

how our evaluation was able to predict the outcome by utilizing the results of the assessment.

Assessments can also be used to identify high-potential individuals as you build your leadership pipeline and determine succession plans. To get at the heart of professional development, using a battery of assessments can help executives target their own personal gaps to aid in their professional development. In order to do this effectively, it is important to understand why using a battery of assessments is highly essential and extremely beneficial to assess an individual's true talent. Using a highly validated and reliable instrument is essential, but more importantly, utilizing one that truly measures the depth of a person's personality is just as critical. As I have already mentioned, the challenge is that not very many assessments have the ability to reach deeply into relevant aspects of behavior that can derail a person's career and also carry a high degree of validity and reliability.

## Understanding Validity and Reliability

It is essential to understand the significance between validity and reliability and why they are both important to understand when determining whether you are using a simple assessment or a sophisticated assessment. What is interesting to note is that the only topic people ask about is validity. It is the most common question that surfaces when the subject of assessments comes up. This is because most people don't understand enough about assessments to ask meaningful questions. All assessments really are validated. I have not found one yet that wasn't. So the response to the question is yes. But the more important query is, "to what degree and by whom?" While I'm going to address the aspect of validity, it is just as important to address reliability to help you understand why this is as critical to understand as validity. While validity and reliability are both important, there are many more aspects of a quality assessment that you will come to learn which go beyond these two terms. The two go hand in hand, but they are not one in the same when it comes to differentiating the simple from the more sophisticated assessments.

Technically speaking, it is rare that an assessment on the market has not been validated. In other words, does it measure what it says it measures? There are different degrees of validity that you will come to learn as well. The bigger question is: "To what degree have they been validated and how deeply have they been tested?" There are some assessments that I call "homegrown," meaning they were developed in-house by a company for its own use. There are also some that were developed by someone, usually a psychologist, who wanted to develop his or her own brand. Here is where you

really need to be careful because the odds are that the psychologist does not have a team of experts to support and maintain the ongoing studies necessary to keep it current. If you are using a particular assessment and Buros Institute's *Mental Measurements Yearbook* has not critiqued it, you need to be very vigilant regarding its validity and reliability. If you ask your assessment provider if its instrument is validated, the odds are high that you will get a yes. If you ask if it's reliable and they say yes, they should be able to tell you the test-to-retest coefficient. If they don't know what this means and don't know what it is, the response should tell you that the provider doesn't understand the difference. These terms are not one and the same, and it is important to understand the distinction when it comes to selecting not only an assessment, but a vendor. What most people do not realize is that many, if not all, assessment instruments are validated. The problem arises when assessments are validated only by the people who own and developed them and not by an independent review board. For an assessment to be validated, the review really needs to come from an outside third party, not from the company which produced the assessment.

Imagine if a car manufacturer conducted its own crash tests and published its results instead of the Insurance Institute for Highway Safety, or letting BP determine if its drilling operations are safe. Think of it as a company that audits its own financials rather than using an outside, independent accounting firm. If you were an investor who uses financials to purchase companies, or even stocks, you would want, perhaps even demand, an independent audit of the company's books before you buy or invest. You wouldn't take the company's word that its financials were appropriately audited. The same holds true for companies, and especially for independent professionals who develop their own assessments and validate their own research data and then claim their instrument is validated. You must bear this distinction in mind when you compare the different assessments on the market and the various test providers.

This chapter is about helping you understand what to ask and what to look for when selecting an assessment, as well as an assessment provider. I point this out because in my research there are, in fact, several test developers who have conducted their own validation studies in-house but have not had an independent review to critique and substantiate the results. As I have mentioned, one such third-party evaluator of assessments is the Buros Institute, publishers of the *Mental Measurements Yearbook*. An independent review board critiques and evaluates assessments in the marketplace and publishes its findings. If you cannot find the assessment you use on the Buros's website, there is a strong possibility the assessment has not been validated by this outside agency.



## Reliability – Its Significance and Importance

While all assessments are validated to some degree, another measure of their utility is their reliability. This is often referred to as “test-retest reliability” and is just as important as validity. To be clear, *validated* means it accurately measures what it claims to measure and accurately predicts what it claims to predict. “Reliable” means it gives the same results today as it will next week or next month. A valid bathroom scale indicates how much you weigh; a reliable bathroom scale shows you weigh the same twice in a row. When we say that a scale or test cannot possibly be more valid than it is reliable, we mean that you can’t say the scale got your weight right if thirty seconds later it showed a different weight.

As an example, the Myers-Briggs/MBTI, a very common, widely used assessment in measuring personality traits, comes up short on the test-retest reliability scale, placing only about half of all examinees in the same category (type) they were placed in the week before. Also, the Myers-Briggs is not validated as a selection tool because of its reliability, not its validity. This means that an instrument is valid in what it measures but its reliability may not give you the same results if you take it more than once. I have personally taken the MBTI. My results show me to be an INTP and then on another occasion show me to be an ENTP. Unless you are knowledgeable about the MBTI, these symbols are meaning- less to the average reader. What this means is that it shows me to be two different people, one more extroverted and one more introverted. So which am I? The answer is that I am both, but the MBTI pigeon-holes me into who I was at the moment I took it. What is important to understand is that this is one of the reasons why assessments can be so confusing. The Myers-Briggs does not measure personality as deeply as necessary to justify its use as a selection or development tool.

The DiSC is another validated instrument, but it is not reliable enough to be used as a stand-alone selection tool, either. It has the same problem as the MBTI in the test-retest area. It’s the reliability and the depth of what is measured, not its validity that is limiting its reach. This is a highly important distinction. The four-dimensional instruments like the MBTI, DiSC, Birkman, PI, and other popularized assessments may all have challenges in this area. I will discuss the four-dimensional model in more detail later in this chapter and its limitations in selection and development. As a point of reference, a well-designed assessment will yield a reliability factor of no less than .70 in its test-retest coefficient.

## Beyond Validity and Reliability – What’s Essential to Know?

Distortion, test construct, and test content are additional factors to consider. These features indicate what traits are and are not measured, as well as how deeply they are measured. Assessments are sold much like insurance products through distributors like DiSC and Profiles International. Some are sold from franchise operators like Predictive Index, and some are sold direct like Caliper. Most practitioners or providers will only use or provide one assessment. You can’t buy a Progressive Insurance policy from an Allstate or State Farm agent. Dealing with an independent insurance agency, you may be able to get multiple carriers. It is not likely someone would offer a Profiles XT and a Caliper together. I chose to remain independent and am in constant search of the best assessments on the market. I am always evaluating various assessments in an effort to keep abreast of what is on the market, what they measure, and how valid and reliable they are.

All assessments are validated to some degree and, while the question about validity and reliability are significant, there are other considerations that are also important. To help disseminate the simple from the sophisticated assessments, here are the questions you should ask to determine whether you’re using a well-constructed instrument:

- How is the test designed and structured?
- How is distortion measured—whether a person is faking good or faking bad?
- How many distortion scales, if any, are used and what do they tell you?
- What is the test-retest coefficient reliability and consistency of the assessment over time?
- Is it forced choice or open choice?
- How long does it take to complete the assessment?
- How many and what kinds of behaviors and factors are measured?
- How deeply are the dimensions measured?

These are just some of the considerations and your provider should be able to help you evaluate whether you are using a quality assessment. If they can’t, the odds are they don’t know what they are providing you. The more sophisticated assessments usually meet these criteria while the simpler ones do not. Understanding what differentiates the various

assessments and which ones will yield the greatest benefit is essential, and the payoff is well worth your time to evaluate what assessments to use.

## What they Measure and the Depth of What they Measure

The process of using an assessment is, in fact, a way to measure someone, something, or some trait. When looking at an assessment, you need to ask yourself the following questions:

- What is it that I want to measure?
- How deeply do I want to measure what I'm looking to evaluate?
- What is the instrument of choice I will use? Ruler? Yard Stick? Tape Measure? Micrometer? Laser?

These questions need to be asked because not every assessment measures the same traits and not every assessment measures every trait. This is what I referred to earlier when I said that it's what an assessment can't tell you that is important to understand. As an example, assessments like the PI, MBTI, DiSC, and Birkman are four-dimensional assessments; they measure four traits. The DiSC and the PI measure the same four traits. The Profiles XT measures nine behaviors and the 16PF Questionnaire measures sixteen. Some traits are measured with more depth than others. As an example, the Myers-Briggs, DiSC, and PI measure some but not all of the traits that are measured in the 16PF Questionnaire, Hogan, or OPQ. The 16PF Questionnaire, Hogan, OPQ, and a few others are more sophisticated assessments and measure more traits than the MBTI, DiSC, and PI.

What's even more significant is that they measure them at a deeper level than all the four-dimensional assessments. What I mean by measuring a trait deeper is that to weigh in on a certain personality trait, some assessments may have only three questions related to measure the trait and some assessments may have ten questions. Which do you think weighs in on the trait better? DiSC and PI don't have questions; they only use single words to describe behavior. This is why test construction is so important to understand when selecting an assessment. The Myers-Briggs assessment asks questions but only allows you to choose either/or answers, which pigeonholes the responses.

With the vast amount of assessments on the market, it is difficult to understand which ones to use as well as how to evaluate the results. Most of the time the assessment used by a company is the assessment most readily available, not the one most researched. Other reasons a particular assessment is used is because someone recommended it or it's easy to use and understand. The DiSC and the PI fall into this category, as they are very simple and easy to use but also very limiting. Unfortunately, people usually select the least expensive and/or the simplest. However, selecting an assessment using this criteria is reducing the choice of an assessment to a commodity and means you're thinking of price, not substance or relevance, let alone validity and reliability. If a company does not use assessments as an integral part of its selection or development process then there will be less emphasis on the assessment process itself, which will lead to a shortage of valuable information about employees.

The assessments on the market are as different as the people they measure. There are simple assessments which include the four-dimensional assessments I have been referring to, and then there are those that are more sophisticated and are psychological in nature. The most important thing to understand is that the more sophisticated assessments measure behavior at a deeper and broader level and are more reliable. You get far more information when your assessment has the ability to evaluate the depth of an individual's behavior. The problem with the four-dimensional products on the market is that they fall short in their reach and ability to do so, and their reliability is limited.

## Four-Dimensional Assessments

The simple assessments I have been referring to are "four-dimensional," meaning they typically measure four behaviors: Dominance, Influencing, Steadiness, and Conscientiousness. The DiSC is indicative of this four-dimensional model. The PI/Predictive Index is also indicative of this model and uses the same four behaviors as the DiSC, but labels them differently. Predictive Index/ PI uses ABCD to define the same behaviors as the DiSC and they are interchangeable. Both instruments provide a three-graph model for visual representation to the degree to which a person exemplifies those traits. Both structure their assessments differently but they both ask you to select individual words that you feel define you and both assessments take about fifteen minutes to complete. These are examples of the simple instruments that measure general behavior traits. They are not sophisticated enough instruments to measure the depth of behavior to any significant level. I explain this in more detail in chapter 11 where I explore how behavior impacts leadership.

More sophisticated assessments are psychological assessments and require more training and expertise to interpret their meaning. The more sophisticated ones on the market are the 16PF Questionnaire and Hogan (both of which I use), and the OPQ, CPI, NEO, and a few others. These instruments require a higher level of understanding because they are more complex, and they usually take upward of forty minutes to complete because their questions (and results) are more extensive. Interpretation is more complex and more meaningful because you are using a stronger instrument that yields more information. Most HR professionals have used some type of four-dimensional assessment. Unfortunately, the consistency of the results is a concern. I personally have ended up with different results from the tests. In the PI, for example, I am High A, High B, Low C, and cut back or mid-level D. The pattern from my PI is identical to my DiSC pattern. This should come as no surprise when you study the similarity of the two instruments. Next is Profiles International and their product, the Profiles XT, the Birkman (using a four-color quadrant to demonstrate the four behaviors), the Caliper and a host of others.

The four-dimensional assessments I have listed are inexpensive, simple to administer, simplistic by design, and easy to take (ten to fifteen minutes), and they give you a computer-generated report of the findings, often captured in what is referred to as a pattern. In the Myers-Briggs (MBTI) you might be an INTP or an ENTP or an ENTJ, which is how people are labeled. The MBTI is a good coaching tool but it has another drawback as there are certain traits it does not measure which limit its reach and effectiveness. Besides these four-dimensional assessments, there are numerous others on the market as well, but typically they are what I call knockoffs of the versions I have mentioned. That is how Extended DiSC emerged.

HR professionals are generally knowledgeable about most of the four-dimensional assessments and some are even certified in some of these instruments. All of these four-dimensional assessments have their value in certain applications but not in selection or development. They are better suited for team building, understanding the differences between employees, and helping to better understand the communication style in others.

In the vast majority of the reports generated from the four-dimensional assessments, you will find “soft” language. Essentially, this means it is difficult to tell if there are any problems or areas of concern because the results are always presented in such a favorable light. Given the results and the language in most reports, it is hard to discern whether a person has any underlying problematic or extreme behaviors. This is more problematic on the selection side than the development side. That’s because on the selection side you will typically not have any prior data to draw from about an individual’s behavior or performance. If the report you receive does not give you any red flags or

areas of concern, how do you know if you are making a bad hire or a wrong promotion? If the assessment you are using does not provide the kind of results that provide insights into the kinds of behaviors you need to understand, it's likely you will make poor decisions on selection and promotion. To help you understand this concept, chapter 11 will get more into evaluating how behavior impacts performance and leadership.

If you take the DiSC assessment, your personality will be defined in similar terms as the MBTI in the form of a pattern. Taking an MBTI you will be labeled by a pattern, such as an INTP or an ENTP as I mentioned earlier. In the DiSC results you will be labeled by a pattern as well. Your pattern might identify you as a *Counselor* or *Perfectionist* or one of fourteen other patterns that the DiSC identifies. The PI does the same thing and uses sixteen patterns as well to describe people. The two tests just have different names for their patterns, but if you read the pattern descriptions, you will find they read almost identical. Another problem with this is that people become labeled by their pattern.

If you listen carefully to MBTI believers, they will tell you who they are by telling you what pattern they are. "I am an ENTJ," they say, as if everybody knows what an ENTJ is. You will find the same thing with those who use the PI as you may be an *Altruistic* profile or the DiSC where you are a *Counselor* profile, as these kinds of assessments tend to label people. I use these examples to help state the case that these types of assessments are good for team building or understanding one another, as well as better understanding of why others do what they do. They are not designed to assess and evaluate leadership development or selection, nor are they strong enough to carry an assessment process that can support a world class leadership development program.

If you look at the PI and the DiSC training material, you would see they are extremely similar, if not identical in many ways. The DiSC material says you should not use the DiSC as a selection tool, but the PI touts theirs as a selection tool. I know the DiSC has the same challenges the MBTI does as it relates to reliability, and it is highly likely that every assessment on the market that uses a four-dimensional mode, including the PI, does as well. The construct of the DiSC and PI are different in that one is forced choice and the other is open choice. Both use a selection of individual words in their approach, but they construct them differently in the way you choose them. However, at the end of the day, they really aren't that different.

I have taken the Myers-Briggs, DiSC, PI, Profiles XT, and others just to see what they measure. I have been using the DiSC since 1991. While I am a certified distributor of the DiSC instrument, I know it is only acceptable to use in certain situations and I do not recommend it to my clients as a stand-alone selection tool, and especially not a

leadership development tool. The DiSC has its place and is a great team-building tool. When used in a battery of assessments it can provide additional supporting data in evaluating behaviors, something you will see in chapter 12 on case studies. Understanding the different nuances of assessments will help you gain an understanding of what you are looking for and what to ask when you're looking for the right assessment and a practitioner to support you. There are also specialty assessments like the Firo-B, the Murray Interests and the Holland Variables, and StrengthsFinder instruments that measure interest and motivation. These are unique instruments that can add to an individual's ability to gain more self-insight and can be part of the coaching equation, but they aren't revealing enough about a person's behavioral makeup. Another instrument somewhat different from many of the others is called the Profiles XT from Profiles International. It is a hybrid, meaning that it measures nine behaviors, incorporates the six Holland Variables, and has four cognitive scales. While it is a good instrument, it is lacking in a few areas. Most importantly, it lacks reach in the areas of workplace coping skills and emotional resilience; both significant factors that are identified by true psychological assessments and something I will explain more in-depth shortly. The Caliper, a very similar instrument to the Profiles XT, also falls short in the same way since it does not measure workplace coping skills.

### Sophisticated Instruments

From my research some of the top instruments are the 16PF Questionnaire, Hogan, CPI, OPQ, and NEO. These are considered psychological in nature and provide deeper personality insights related to what is referred to in the world of psychology today as the "Big 5 Theory," which I will address in a moment.

These assessments are more expensive and not readily available for general distribution because of the level of training needed to use and resell them. Of these five, I use the 16PF Questionnaire and the Hogan, depending on the nature and needs of the situation. In chapter 12, "Case Studies," I will explain in detail the best uses for the 16PF Questionnaire and the Hogan. Both these tools can uncover behaviors that most assessments can't. In addition, I'll show how the Hogan brings confirming information to the results when used in our executive battery.

## The 16PF “Big 5” Theory and Psychology Today

Raymond Cattell, who developed the 16PF Questionnaire, was one of the early pioneers of personality testing. The following is taken from an article called “Big Five Personality Traits” and describes the relevance of the 16PF Questionnaire and its historical background which makes it one of the most reliable assessments on the market. This is from the article:

Sir Francis Galton in 1884 made the first major inquiry into a hypothesis that, by sampling language, it is possible to derive a comprehensive taxonomy of human personality traits. In 1936 Gordon Allport and S. Odbert put Sir Francis Galton’s hypothesis into practice by extracting 4,504 adjectives which they believed were descriptive of observable and relatively permanent traits from the dictionaries at that time. In 1940 Raymond Cattell retained the adjectives, and eliminated synonyms to reduce the total to 171. He constructed a personality test for the clusters of personality traits he found from the adjectives, called Sixteen Personality Factor Questionnaire. Then, in 1961, Ernest Tupes and Raymond Christal found five recurring factors from this 16PF Questionnaire. The recurring five factors were: surgency, agreeableness, dependability, emotional stability, and culture. This work was replicated by Warren Norman, who also found that five major factors were sufficient to account for a large set of personality data. Norman named these factors surgency, agreeableness, conscientiousness, emotional stability, and culture; and these factors are through which the 5 Factor consensus has grown.

Wikipedia.Org, “Big Five Personality Traits,”  
November 30, 2015

The 16PF Questionnaire is now in its fifth edition and has recently updated its norms.

## The 16PF and the Big 5 Theory

The 16PF categories for the Big 5 model are: EX–Extraversion, ER–Emotional Resilience, TM–Tough-Minded, IN– Independence, SC–Self-Control. These are referred to as global factors. The global factors are then broken down further into subfactors that are referred to as primary factors to give deeper meaning to the trait. This is where the 16PF (Primary Factors) comes from. As an example the “D” of the DiSC and the correlating factor “A” of the PI both identify the level of dominance a person displays. The 16PF Questionnaire uses IN for the word “Independence” to define



dominance as well. The difference is that the 16PF model breaks down the world of dominance into four subcomponents to help you see how the dominance is structured and what is underneath it. This is why this is a very detailed, in-depth assessment. More information on the use of these assessments will be found in chapter 11.

As I have discussed previously, I incorporate several different assessments in my practice, particularly when administering a battery of assessments for executive evaluation and development and, in so doing, the main instrument of choice is the 16PF Questionnaire. I have used it in nearly 13,000 evaluations over the last twenty-five years in the areas of selection and promotion, employee interventions, coaching, executive development, and succession planning. I have found it to be the single most reliable and accurate assessment in predicting behavior and emotional resilience. It is one of the most, if not the most, sophisticated instrument on the market. Its value and accuracy have been proven countless times in my validation studies in a variety of different roles to include the world of leadership assessment. It has also consistently held up well across international borders from Asia to Europe to South America, and is available in more than thirty languages. As I mentioned, another valuable tool I also incorporate into our assessment battery is the Hogan. I employ this when conducting our executive evaluations, executive coaching, and especially interventions. The Hogan has very similar characteristics to the 16PF Questionnaire. In fact, it complements the 16PF Questionnaire and provides a component called “derailers” that looks at how a person copes with challenges. However, there is one limitation with the Hogan. To get the same data I get from the 16PF Questionnaire in one report it requires two different reports from Hogan. Needing two reports from Hogan adds to the cost but, more importantly, the interpretation becomes more complex. Having to compare and contrast the two reports makes interpretation more challenging because you have to look at two different reports to interpret the meanings. Still, the Hogan is one of the better assessments on the market because it carries a high degree of validity and reliability and it is a great complement to our executive battery.

To show you the reach of the 16PF Questionnaire, we were chosen to conduct all of the leadership evaluations for the thirteen top executives in a \$4 billion company located in the United States. Its parent company was located in Europe and the total global value of the organization was \$11 billion. We were chosen to evaluate and create development and coaching plans for the US executives and Dr. Steve O’Shaughnessy, a UK-based consultant, was chosen to do the same for the European executives. During a conference call to coordinate efforts on the project, I discovered that Dr. O’Shaughnessy also used the 16PF Questionnaire as well as the Hogan. I asked him why he used the 16PF Questionnaire in his practice and he said, “I believe the 16PF Questionnaire is a gold

standard of psychometric testing.” Who would have thought that two firms, thousands of miles away, asked to evaluate leadership traits around the world for an international company, would use the same instruments to measure behavior and create leadership development plans? Perhaps it speaks to the integrity and worldwide reach and universal acceptance of the 16PF Questionnaire and Hogan assessments.

## Test Results – Integrity and Distortion

### Understanding Faking Good and Faking Bad

When using any instrument, you need to know its true ability to predict personality. That’s the validity and reliability piece I have been referencing. Beyond the validity and reliability of the assessment itself, there is the integrity of the assessment results of the individual who took the assessment. This is called *distortion*. The problem exists because not many assessments can discern distortion to any depth nor have a way to measure true distortion. Distortion is the degree to which an individual is faking good or faking bad. There is the validity and reliability of the assessment and then there is the validity and reliability of the results of someone’s responses. These two results are entirely separate. You have to be able to measure distortion to see if the person is presenting himself in a different fashion from who he really is and be able to determine if the results of the assessment can be counted on as a fair representation of the individual who is taking the assessment. In other words, did he put on a different face when he took the assessment? This is why using more sophisticated assessments and especially multiple assessments is so valuable. A person might fool one assessment, but it is not likely that individual can fool a battery of assessments. This is yet another reason the 16PF Questionnaire is my instrument of choice because it has three distortion scales built into it.

## Distortion Scale -- Impression Management (IM)

### 24 Weighted Questions

Impression Management, the first scale, is the tendency of the applicant to try to conform to societal expectations of making a good impression while taking the personality test. It is a person’s slant on himself. It is normal for job applicants to try to look their best, especially in sales, so they will usually score above average on the Impression Management scale. Low scores may indicate self-esteem issues.

## Distortion Scale – Infrequency (INF)

### 32 Response Indices

The second scale, Infrequency, is based on the number of times the individual selected the “b” response or the “?” alternative to one of the item stems (questions) for items on which most people readily make a decision. High scores on this scale correlate with indecisiveness, possible reading problems, concentration difficulties, or noncompliance, and may suggest a candidate is trying to hide something. In contrast, the MBTI, DiSC, PI and many others cannot measure Infrequency because of the way their assessments are constructed.

## Distortion Scale – Acquiescence (ACQ)

### 103 True/False Questions

The third scale, Acquiescence, measures the person’s tendency to agree, indiscriminately, with items on the test. High scores here may indicate disinterest in taking the assessment or may be related to comprehension or concentration problems.

As an example, while the 16PF Questionnaire has three distortion scales, the Profiles XT only has one. On this one particular distortion scale, the 16PF Questionnaire uses 103 True/False questions and the Profiles XT uses thirty-five yes/no questions. The three different distortion scales used by the 16PF Questionnaire make it superior in its ability to spot inconsistencies in a person’s responses. While there are several thousand published articles on the 16PF Questionnaire, further evidence of its value is presented in Dr. Michael Karson’s book, *16PF in Clinical Practice*, Dr. Jim Schuerger and Dr. Heather Cattell’s book, *Essentials of 16PF Assessment* and Wendy Lord’s two books *Personality in Practice* and *Overcoming Obstacles to Interpretation*. These authors are among the foremost authorities on the 16PF Questionnaire.

While I continue to explore the assessment landscape to search for the best instruments, my trust remains with the 16PF Questionnaire as one of the premier instruments that can support a quality selection model and a true leadership development model.

## Workplace Coping Skills / Emotional Resilience

I stated earlier that an important distinction is what an assessment *can’t* tell you. What this means is that some assessments measure traits that others don’t. It’s another

important reason you need a more sophisticated instrument in measuring behavior and that is to capture a person's emotional resolve and coping skills. The 16PF Questionnaire is one of the more sophisticated instruments that can do this as it not only measures sixteen general dimensions of behavior, but also measures an individual's problem-solving ability as well. The four-dimensional assessments, and even the Profiles XT with its nine behaviors, does not reach deeply enough in measuring the emotional resolve of an individual. In our leadership work regarding high-potential executives, we know that there are eleven core traits in successful leaders that can be identified by the 16PF Questionnaire. In our case studies we can see when the 16PF Questionnaire has predicted the Peter Principle on numerous occasions.

### Why Use a Battery of Assessments?

I have attempted to explain some of the differences in the various kinds of assessments and which ones to use. I have also explained why I have chosen the ones I use to help you understand what assessments will be most useful for your organization. My recommendation is that you use the 16PF Questionnaire and the Hogan because they are among the best on the market. Furthermore, I have also created a battery of assessments that include three behavioral/ psychological assessments and three assessments related to measuring critical thinking and problem-solving skills. The combination of these six assessments is, in my view, the best all-around assessment model to help identify an executive's strengths and weaknesses. This assessment battery has an enormous predictive value and is useful for the selection, leadership development, and coaching models.

### Selecting and Evaluating Your Provider

By now you should understand there are simple assessments (four-dimensional assessments) and more sophisticated assessments (psychological in nature and The Big 5 Theory), and why it is critical for you to understand the differences. However, there's more to evaluate than the assessment itself. During the entire assessment process, you must use a licensed practitioner who knows how to interpret the assessment, not just someone who sells you a report or recites what the report already tells you. The practitioner must be able to extract meaning from the report. The wrong assessment in the hands of an inexperienced practitioner will likely have a higher probability of failure, whether it's used for selection or development. Even the right

assessment in the hands of an inexperienced practitioner is concerning. The following quote from Dr. Dave Watterson, another expert on the 16PF Questionnaire summed it up nicely. He explained just why it is so important to properly select the best assessment and the best provider or practitioner when he said:

“Merely buying a test doesn’t necessarily guarantee you desired results. Just like a scalpel in the hands of a skilled surgeon, it is the surgeon not the scalpel that makes the difference. Looking at a particular test should not be the only criterion for its use. Looking at the process and whether you are working with a skilled practitioner who understands how to use the instruments is paramount.”

## In Closing

Assessments are extremely complex leadership tools. Even the most experienced HR professionals can find them confusing because there is no uniform rating system covering all the assessments. There are some third-party evaluations but still choosing and using these tools remains challenging. Here are some closing thoughts.

Behavior is the biggest predictor of performance. In an effort to raise productivity and performance, a company can use assessments to improve their bottom line, reduce turnover, save on hiring costs, aid in succession planning, and help target leadership development. It is important to use the right assessment, particularly when you need to measure in-depth behavior and critical thinking skills. Currently, the four-dimensional assessments on the market are simply not deep enough to measure behavior and assess talent to the degree necessary.

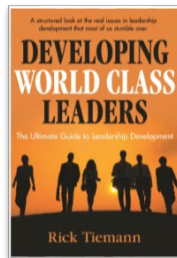
Assessing real talent is best achieved through a battery of assessments that complement each other. A complete battery of assessments measures a person’s behavior, emotional intelligence, how he manages himself when facing challenges as well as his critical thinking skills, problem-solving ability, and his intelligence. This type of application requires a more sophisticated set of assessments as well as an experienced practitioner who can gain the most insights from the information. By utilizing and incorporating more sophisticated assessments, along with the guidance of a skilled practitioner, an organization can optimize its ability to develop world class leaders and hire and retain the very best talent.

~~ Rick Tiemann

From the book

*Developing World Class Leaders:  
The Ultimate Guide to Leadership Development*

by Rick Tiemann



To obtain a copy of the book, please contact:



80 East US Hwy 6 | Valparaiso, IN 46383 | 219.477.6378  
[www.theeg.com](http://www.theeg.com)