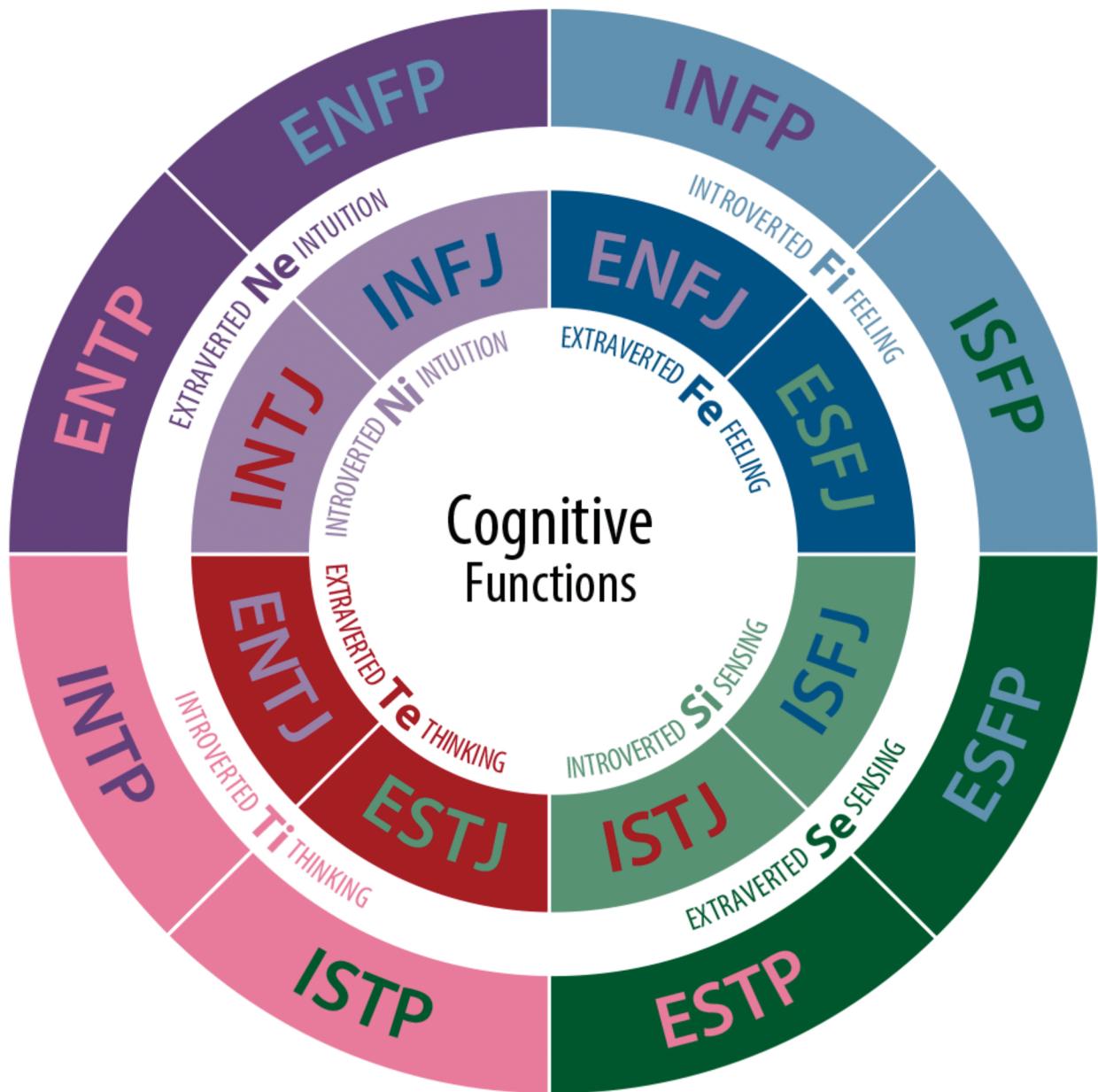


Should You Trust the Myers-Briggs Personality Test?



Laith Al-Shawaf
Mar 9, 2021 — 12 min read



The Myers-Briggs Type Indicator is the most popular personality test in the world. It's a favourite among Fortune 100 companies, government agencies and regular people. More than 1.5 million people take it every year. It is a thriving multimillion-dollar-a-year industry. And as any psychologist worth their salt will tell you, it's mostly bullshit.

Here's why.

Reason 1: It Is Based on Carl Jung's Ideas

The first reason is one of the least important, but it's good to start with history. The Myers-Briggs personality test is based on Jung's ideas, which are—to put it mildly—empirically unsubstantiated. Jung was a Swiss psychiatrist and psychoanalyst who worked in the first half of the twentieth century. He had an abiding interest in religion, mythology, alchemy and astrology. These interests reverberate through his ideas, some of which are closer to mystical or supernatural claims than scientific ones. Jung didn't place much emphasis on testing his ideas through rigorous empirical studies, a problem that was admittedly not unique to him during his time. After Jung died, two non-experts with limited training in psychometrics or test construction created the Myers-Briggs Type Indicator (MBTI), a personality test founded on Jung's empirically unsubstantiated ideas.

This is not enough on its own to reject the MBTI—a test could conceivably be created by non-experts and founded on unsupported ideas, and yet do an accurate job of measuring human personality.

But it is certainly reason to be cautious.

Reason 2: The Test Lacks Predictive Validity: It Is Not Good at Predicting Outcomes in the Real World

The point of personality tests is not just to tell you about yourself, but also to predict outcomes in the real world. But time and time again, studies show that the MBTI is not a good predictor of career outcomes, romantic relationships or anything else we might care about.

By contrast, there are scientifically validated personality tests that usefully predict everything from career performance to likelihood of divorce to how likely you are to

have a psychological disorder (for example, see [here](#)).

Predictively successful tests exist, but the Myers-Briggs is not one of them.

Reason 3: Human Personality Falls Along Continua, Not Into Discrete Categories

This has implications that matter.

Some variables are continuous: they fall on a continuum or spectrum. Other variables are categorical: they fall into neat, discrete categories or types. For example, your height is continuous: it falls somewhere on a spectrum from extremely short to extremely tall. You could be anywhere on that spectrum—say 5'4, 5'9 or 6'2 and three quarters. By contrast, religion is a categorical variable: you are either Jewish, Christian, Muslim, atheist, Jain (or whatever).

If I measure how far you can toss a coin, the answer is continuous, because distance falls on a spectrum, and the coin can land anywhere on that spectrum. If I measure how that coin lands (which side faces up), the answer is categorical: it can only be heads, tails or on its side. There are no other options, and the answer falls neatly into one of those categories.

Human personality is continuous, not categorical. It is like height, not religion. It is like how far you can toss a coin, not how the coin will land. The problem is that the Myers-Briggs framework boldly claims that personality falls into categories or types. And it further claims, without any real evidence, that there are exactly 16 types.

There are different models of human personality. One prominent model called the [Big 5](#) suggests that each person falls somewhere on a continuum of openness to experience ranging from low to high, somewhere on a continuum of low to high conscientiousness, somewhere on the extraversion continuum, somewhere on the agreeableness continuum, and somewhere on the neuroticism continuum (OCEAN for short). [Each person has a score on all five of these continua](#). If you give a person a test and get their score on all five dimensions, you've got a good (albeit [incomplete](#)) sketch of their overall personality—that is, who they are and how they differ from other people.

A different model of personality called [HEXACO](#) suggests that personality can be better statistically captured with six key dimensions, not five. These six traits are honesty-humility, emotionality, extraversion, agreeableness, conscientiousness and openness to experience.

The two models differ, but they agree on the core fact: personality falls along several continua, not into distinct types. For example, the best way to conceptualize introversion/extraversion is by locating it on a spectrum, not by imagining that people fall into one of two types called *introverted* and *extraverted*. Both these models have more evidence in their favour than the Myers-Briggs.

A continuum has an infinite number of points on it, and we are dealing with five or six continua here. So if we really wanted to count how many different personality types there are, the answer would be: infinity multiplied by itself five or six times. If that strikes you as a perversely unhelpful way of describing things, you're right. Because personality traits fall on a continuum, we shouldn't be talking about discrete personality types at all.

The Myers-Briggs takes the beautiful, richly graded complexity of human personality and squeezes it into 16 types. The framework is neat and easy to remember, and the taxonomy of personality types sounds cool. But the resultant picture of human personality is unfaithful to reality.

Reason 4: The Types Used by the MBTI Have Arbitrary Boundaries

Shoehorning personality into types creates the problem of arbitrary boundaries: once you've specified the types that supposedly exist, you must define their boundaries. But it is very hard to define the boundaries in a non-arbitrary way. For example, imagine I take the variable of height, which is continuous in the real world, and try to place it into categories. I could claim that there are three categories of height and label them *short*, *medium* and *tall*. But now I have a problem: what should the cut-offs of the categories be? Does a person who is 5'10 fall into the medium category or the tall one? What about someone who is 5'5—do they count as short or medium? And on what basis should I make this decision? There is no principled way for me to define the boundaries of the categories, and any method I pick will be arbitrary. There is no real *medium* or *tall* in nature, so I cannot look to

the external world to help me decide. I have to decide how to define my category boundaries using intuition, majority vote or some other arbitrary method.

The MBTI tries to deal with this problem by using a median split. This means that many people are tested, and their scores on a trait like extraversion are arrayed from the very lowest to the very highest. For example, if we test 50 people, we end up with a list of all 50 extraversion scores, running from the lowest to the highest in sequential order. The median split method says: let's split this list right down the middle. We'll consider all scores above the middle one (the median) to be extraverted, and we'll consider all scores below the median to be introverted. But the problem remains: why use the median? Why not the average? Why not some other method? The answer is that there is no principled, non-arbitrary way to decide. And no method will solve your key problem: you have taken a variable that is continuous in nature and tried to fit it into your procrustean bed. This creates many downstream problems, of which the issue of arbitrary cut-offs is only the first. The next has to do with the test's reliability.

Reason 5: The Myers-Briggs Has Poor Reliability

Psychologists measure whether or not a test is good in several ways. One important measure is called test-retest reliability. This means that the test should give you roughly the same results if you take it once and then take it again a few weeks later. A bit of fluctuation is normal, because everybody changes a little depending on context. But your personality doesn't fluctuate wildly from week to week, so a good test shouldn't offer wildly different results from one testing session to the next.

The MBTI violates this rule with bravado and fanfare. It has terrible test-retest reliability, often reclassifying participants into the opposite personality type from one testing session to the next. For example, some studies find that within a period of five weeks, it reclassifies a full 50% of participants into a different personality type. Reliabilities this low are considered unacceptable in psychology.

Like the problem of arbitrary boundaries, the problem of poor reliability is a consequence of taking a continuous variable and trying to make it categorical. You are forced to define the boundaries of your categories, so you decide to use the median. But this creates a new problem: many people score just above or just below the median. In fact, many more cluster around the median than fall at the

extremes. If these people take the test again three weeks later, and feel a tad more gregarious (or energetic or anxious), the test could easily reclassify them into the opposite personality type. This is exactly what the MBTI does on a regular basis.

Reason 6: The Myers-Briggs Misleadingly Implies That There Are Big Differences Between Types and Minimal Differences Within a Type

By placing people into discrete personality types, the Myers-Briggs implies two additional things.

First, it implies that those who fall just below the cut-off for a given type are fundamentally different from those who fall just above it—the test places them into different personality types. But in reality, borderline introverts and borderline extraverts are very similar to one another.

Second, by placing people into types, the test implies that all people within a given type are reasonably similar—all extraverts share a fundamental core quality, as do all introverts. This implies that extreme scorers in a type are reasonably similar to borderline scorers in the same type. This is wrong—extreme introverts are actually very different from borderline introverts.

For example, imagine that your cut-off for extraversion is 50%. You would categorize someone who scores 10% or 45% as an introvert and someone who scores 55% as an extravert. But in reality, the borderline introvert (who scored 45%) is much closer to the borderline extravert (55%) than she is to the extreme introvert (10%). And yet the test would group the borderline and extreme introverts together, despite the vast gulf between them (a 35 point difference)—and would consider the borderline introvert wholly different from the borderline extravert despite their similarity (a 10 point difference).

Like the two problems that preceded it, this problem is a consequence of the same core issue: trying to force a continuous variable into categories. In type-based tests like the Myers-Briggs, those who barely make it into a category are considered full and equal members in the category, no less than extreme scorers in the same category. And those who barely make it into one category are considered wholly different from those who barely make it into the other category. Conceptually, this is exactly backward.

Reason 7: When You Turn a Continuous Variable Into a Categorical One, You Throw Away Information

There is another reason most researchers caution against dichotomizing a continuous variable: doing so involves discarding valuable information.

Let's say I measure the heights of all the students in my classes. I get a rich gamut of exact heights ranging from 4'11 to 6'2. Then I decide to sort people into the categories of short, medium and tall. To do this, I have to throw away information. I end up going from each person's exact height (precise) to the rough and ready descriptions of *short*, *medium* and *tall* (less precise). I have chosen to discard the rich complexity and shades of difference between my students in favour of the cruder, less precise categories. I've chosen, in other words, to throw away information.

There are some delimited circumstances in which it can be helpful to do this. But these circumstances are rare, and they do not apply to personality measurement. Most of the time, brutally hammering a continuous variable into a categorical one is not a good idea.

Some human variables *are* categorical (for example: did he kill that person or not?). In such cases, some of which fall into areas of psychopathology, type-based approaches can be helpful. But the evidence suggests that this does not apply to most personality traits.

Reason 8: The MBTI Doesn't Measure Neuroticism

Neuroticism is an important personality variable. It describes how much of a person's mental life is taken up with negative emotions, especially feelings of anxiety, sadness and vulnerability. People who score higher in [neuroticism](#) experience episodes of negative emotion that are more frequent, more intense and longer lasting than people who score lower on this trait. People high in neuroticism are more vigilant about danger in the world, and they see more threats in ambiguous stimuli, [which can be a good or bad thing](#), depending on whether they are in a safe or unsafe environment. Neuroticism is one of the most important predictors of romantic relationships (on average, it predicts [greater dissatisfaction and relationship dissolution](#)). It also predicts—on average—worse career outcomes,

worse health outcomes and a wide array of psychopathologies. It is cross-culturally universal. It is not unique to humans: it exists in a wide variety of species.

Everyone falls somewhere on the neuroticism continuum. The Big 5 model measures this trait and calls it *neuroticism*. The HEXACO model measures it and calls it *emotionality*. It is a crucial dimension of human variation. There are other personality models not discussed in this essay, and they also measure neuroticism. The sole exception is the Myers-Briggs: it completely ignores neuroticism, choosing not to measure the trait at all. If the goal is to capture human personality, this is a staggering failing.

Why would the Myers-Briggs ignore a personality variable that is linked to life outcomes ranging from career attainment to health problems to divorce? Perhaps because the test-makers are incentivized to make the results broadly appealing. If you don't measure neuroticism, you never have to give your participants bad news when the test is done, and you don't risk losing any customers.

That said, it's impossible to say with certainty why the Myers-Briggs ignores neuroticism. What we *do* know is that neuroticism is conspicuously missing from the test, despite predicting important life outcomes. Notice, too, that of the 16 personality types suggested by the MBTI, not one is negative. (You can take a look at the unbridled positivity of all 16 types [here](#)). The Myers-Briggs is in the comfortable—and financially judicious—position of never being the bearer of bad news.

But My Results Still Seem Pretty Accurate

Most people who hear about these problems agree that they are serious. A test that neglects neuroticism, shoehorns continua into categories, has low reliability and doesn't predict life outcomes is a bad combination. But even for some very bright people, an issue lingers: despite all this, their MBTI results seem pretty accurate.

Do you feel this way about your Myers-Briggs test results? If so, consider the following.

First, people tend to believe that their horoscope describes them well, even when they're reading the wrong horoscope. People seem to want to read themselves into

[things](#).

Like horoscopes, the Myers-Briggs test seems to exploit this fact by offering people personality descriptions that contain Barnum statements. A Barnum statement is a strategically vague description that says little of substance, but is written with just the right amount of ambiguity, so that people can read into it and conclude that it describes them beautifully. It is named for the famous showman P. T. Barnum, who supposedly said that there's a sucker born every minute (fittingly, he never said this: it is a [misattribution](#)).

If the Myers-Briggs appears to describe you well, it is because it is constructed to do so, because you want to read yourself into it, because it stays tactfully silent about neuroticism, and because if you already partially believe in it, confirmation bias will entrench your belief.

By contrast, consider the MBTI's key flaws: arbitrary boundaries, poor reliability, failure to predict life outcomes, bypassing of neuroticism, etc. To know about these failings, you would need to be a dataphile or have seriously researched the empirical studies on the issue.

Fine—but It's Better Than Nothing, Right?

If this sounds like a reasonable retort, consider the following.

First, [as many thinkers have pointed out](#), the illusion of knowledge is more insidious than a lack of knowledge, and harder to overcome. By painting an inaccurate and crudely pixelated picture of human personality, the MBTI is an obstacle to a more accurate understanding. In this sense, it may be worse than nothing.

Second, taken to heart, this misinformation could hurt people. It could steer people away from careers or mates that supposedly don't fit their personality type. If the basis for this advice is inaccurate, we risk harming people's lives.

Third, the choice isn't between the Myers-Briggs and nothing. The choice is between the Myers-Briggs and predictively successful continuum-based models like HEXACO and the Big 5. Both these models of human personality reliably [predict important life outcomes](#) and avoid the key mistake of shoving continuous variables into categories.

Why Is the Myers-Briggs So Popular Then?

I'm not sure why the Myers-Briggs is so popular despite its shortcomings. But candidate reasons include: (1) it has excellent advertising and money to back it; (2) the test is easy to take, easy to administer and easy to calculate; (3) the results are easy to interpret and understand; (4) the test tactfully avoids telling the reader anything negative; and (5) some evidence hints that we might be cognitively disposed to think in terms of [dichotomies](#) and [dualisms](#) rather than continua (*introverted vs. extraverted* is more intuitive and less cognitively taxing than a continuum with an infinite number of points on it), leading us to prefer the cruder and less accurate model.

Concluding Remarks

What should we take away from this?

Personality psychology is a thriving field that has led to important discoveries about how humans differ from one another and how they are the same. Unlike some subfields of psychology, it has [done exceptionally well in the replication crisis](#). Personality might seem wispy and intangible, but it can be measured rigorously if you have the right tools. And it is of great practical importance: it predicts (on average) [everything from salary to career performance to romantic and friendship relationships to running afoul of the law](#).

But the Myers-Briggs is not the tool for the job, as it fails spectacularly at these tasks. If companies and governments want to make evidence-based decisions, they should avoid the MBTI and opt for models of human personality and career propensity that have better validity and reliability.

As for regular people like you and me, we should not be lulled into the facile belief that the Myers-Briggs can tell us much of substance about ourselves, or that it can help us find the right career or mate. We'll have to rely on our own minds for that—plus the chancy contingencies of an indifferent universe.

The Myers-Briggs offers only the illusion of self-knowledge, not the real thing. A closer look reveals that it rests on a foundation of empirically unsubstantiated

Jungianisms and unsound psychometric practices.