Max Traylor®

Top Consultant Pressures

WHITEPAPER

Psychology is trendy in market research these days.

There are a number of AI-powered psychology-based platforms on the market. The question is, which one of those platforms are substantive vs just for show? Or, which one will provide you with insights that give you a real business advantage? In the following paper, we will review the psychometrics behind the Problem Profiles Scale, highlighting its validity and reliability. Our hope is that by the end of this paper, you will be able to see:

How much scientific rigor was put into the construction of Clarity

How trustworthy the results are

How to compare the rigor of our platform with other platforms on the market

CLARITY

BY MAX TRAYLOR

Consultants today are pushed and pulled by so many pressures, such as market, personal, and situational pressures. Since the advent of the pandemic, business practices have changed and have presented consultant's with unique challenges. These pressures push consultants to develop limiting beliefs and problematic behaviors, which impact their daily performance and profitability. Max Traylor helps consultants to find clarity—developing manageable practices that help them optimize their efficiency, creativity, energy, and profitability.

Clarity is a new platform that uses patent-pending systems for AI-powered projective tests. This platform helps consultants identify common problems consultants face as they try to run and grow their business. Clarity offers solutions tailored to each consultant to increase their creative energy and satisfaction with their job and life.

How do we do it? Clarity uses a unique combination of psychological science, data science, and machine learning algorithms to produce intelligent AI-powered projective test technology built on the following underlying conceptual principles:

- There exists a common set of problem profiles that impact a consultant's ability to grow and run their business.
- These psychological states may impact a consultant's burnout, life satisfaction, job satisfaction, and creativity. A consultant may or may not be aware how these psychological states affect them.
- Clarity uses what consultants' say and do in our proprietary projective tests to identify their psychological profile and suggest solutions.

What insights can you learn? By using Clarity, a consultant can get access to insights about their:

1. Problem Profile scores
2. What the scores mean
3. Recommendation for "dos and don'ts"
4. Recommendations on how to maximize creative energy

Key Benefits of Clarity

Why use our platform? While there are a number of ways the Clarity is beneficial, there are four primary benefits of Clarity:

Automated Predictive AI. Clarity automatically creates predictive algorithms unique to each consultant. The predictive AI not only provides consultants with their problem profiles and recommends solutions, but it also (1) predicts what would happen if the consultant doesn't implement changes with great accuracy, what's the potential impact (e.g., burnout in

30 days), and (2) if they do implement the changes, what's the potential impact (eg., more creative energy). The predictive AI grows as the consultant grows, getting better each time it is used.

Multiple Functions, Expandable Market. Clarity is designed for continuous use, where consultants are able to use the platform to:

- Log and update manageable practices they implement (e.g., I'm planning twice a week now)
- Track the impact of their "Problem Profiles" on work over time (e.g., I'm red-lining significantly less now)
- Receive recommendations for "dos and don'ts"
- Breakdown time and energy costs of work tasks

These key features can be marketed to the average consumer. Our research indicates that in the post-pandemic industry professionals, beyond consultants, also face many issues adapting to new work culture. Increase in burnout and minimal job satisfaction have led to what many refer to as the "Great Resignation" in the past two years. A platform, such as Clarity, can be of great use to the average consumer struggling with limiting beliefs and problematic behaviors in their daily lives.

Less Questions, More Insights. With Clarity, consultants have to answer less questions and get more insights. For years market researchers have been saying surveys in our industry are too long, leading to poor data quality from burnt out survey respondents. So at Inkblot Analytics, we wanted to create a solution that allows researchers to get the same amount of data by asking less questions.

Data Extravaganza. Clarity is connected to a number of web apps, consumer databases, and other martech solutions. This means that consumers who are interacting with these apps are sending data back to our database to "refresh" the data. This way, there is always a continuous flow of new data into the database every day. This keeps our models and benchmarks as updated as possible. For other platforms, updating data may happen once or twice a year. However, when something dramatic like the Covid-19 pandemic comes along, and consumer behavior shifts suddenly, you want a flexible and adaptable solution like Clarity that is constantly refreshing the thoughts, feelings, and behaviors of consumers.

How Does the Clarity Platform Work?

Brand Blots has over 36 types of projective tests. This paper is focusing specifically on the digital inkblot test. This inkblot test is a four step process:

The Testing Step
Taking the Problem Profiles Scale

The Scoring Step
Scoring reponses to the image interpretation test

The Profiling Step
Identifying which profile is predominant for the individual

The Predicting Step
Predicting outcomes and solutions for consultants

Each one of these steps has a scientific process built into them. For the **testing step** (i.e., when the participant takes the Problem Profiles Scale), we want to make sure the data is good quality. So we use an algorithm that measures the extent to which a respondent is intentionally trying to deceive the test, not take it seriously, or enter in bad quality data. For the **scoring step** we use measures of inter-rater reliability. For the **profiling step**, we use classic psychometric measures of validity and reliability to know the traits we're measuring are trustworthy. For the **predicting step** we use the model's error (the difference from the predicted score and actual score) to know how accurate/precise the model's predictions are. Over the course of the rest of this paper, we'll go in depth on each of these aspects so that you can see just how science-based this tool is.

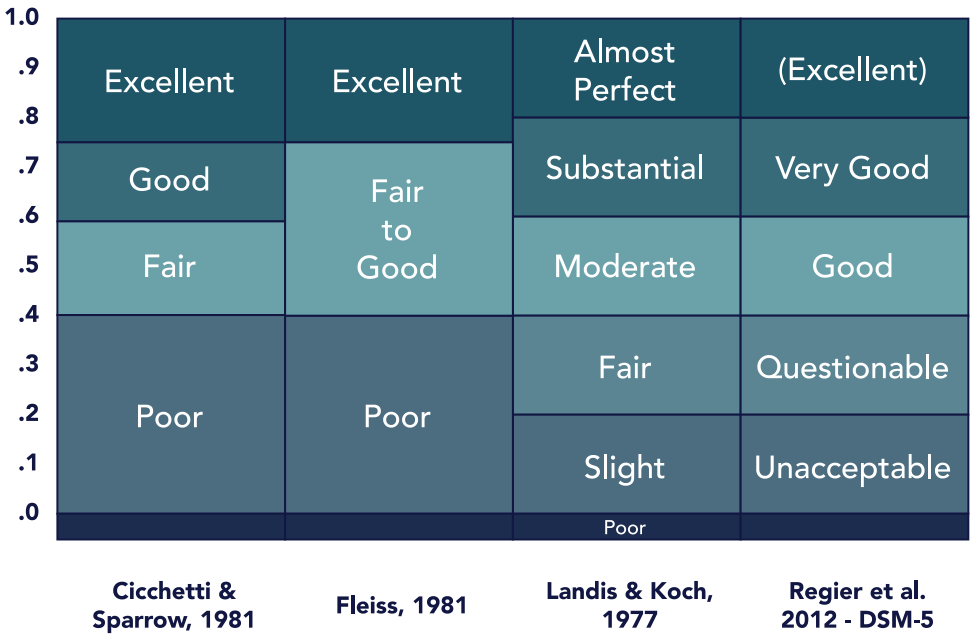
The Scoring Step: Measuring Inter-rater Reliability

CLARITY

Due to the high velocity of data we sometimes receive, we have multiple coders who apply a specific scoring scheme to the inkblot test responses. However, as you may suspect, everyone has a slightly different way of interpreting ambiguous data. As a result, all coders are put through a training program for how to score the inkblot test responses. Once the coders have sufficiently passed a scoring test, they are allowed to work on scoring project data. For any given project, we have 2 coders score the responses separately. No coder is able to see how any other coder has scored the responses, keeping all parties independent of possible scoring influences. However, to continually check that all coders are scoring the responses similarly, we calculate inter-rater reliability on all projects, and overall, on an ongoing basis.

Inter-rater reliability (IRR) is a statistic that measures the consistency of our coding methods. Basically, it's a check to see if our trained coders are applying the same codes to the same inkblot test responses.

Historically, there are a few different approaches as to what is considered a "good" versus "bad" reliability score. You can see these approaches, and their references, in the below chart. At Inkblot Analytics, we traditionally follow the inter-rater reliability approach outlined by Regier et al (2012), shooting for .80 reliability or above. This means that we always expect our coders to agree on a minimum of 80% of the scoring they do.



The Profiling Step: Psychometrics of Measured Traits

Once the test data is collected, we are able to use our proprietary algorithms to help build psychological profiles. First, however, we have to make sure that our scales are accurately and consistently measuring each problem profile. In other words, we have to make sure that our scales have strong psychometric properties. Without assessing the psychometric properties of constructs, we can't be certain if we are "tapping into" the construct we are interested in. For example, we may think we are "tapping into" the construct of extraversion, but in reality we might be measuring the "likelihood to talk to strangers."

To measure the extent to which a person falls into a given "Problem Profile" we created the Problem Profiles Scale. In this section, we walk you through the scientific process of how we evaluated the psychometric properties of the Problem Profile Scale, using the Workaholic profile as an example.

We determined how applicable a problem profile was to a consultant by measuring the triggers that prompt an individual to fall into a particular profile (Spark) and how the psychological qualities of a given problem profile affect an individual's thoughts, feelings, and behaviors (Symptom). Each problem profile can be divided into Spark and Symptom subsets. Together, scores on each subset are added to formulate each individual's profile.

For the Workaholic problem profile:

Workaholic Spark

A high score indicates that an individual is experiencing multiple "triggers" associated with the Workaholic problem profile. For example, someone who scores high may feel that a busy work schedule proves their value, thereby prompting them to overwork themself to the brink of exhaustion and at the expense of their personal life. Low scores indicate that Workaholic triggers are scarce for the individual and so the Workaholic problem profile may not be a major source of the consultant's issues.

Workaholic Symptom

A high score indicates that an individual experiences multiple symptoms that correspond to the Workaholic profile. For example, someone who scores high may feel unfulfilled when they stop working. Low scores indicate that the thoughts and feelings of an individual minimally correspond to the Workaholic problem profile and so the Workaholic problem profile may not be a major source of the consultant's issues.

We can determine the extent to which the Workaholic profile represents a person by adding up their scores on Workaholic Spark and Workaholic Symptom. We repeat this process with the remaining problem profiles. Since every person has a mix of the problem profiles, we can use these scores to detect the one profile that is the primary source of a consultant's issues.

A profile is problematic when it starts to interfere with a consultant's ability to produce, deliver, and function. Our approach is not to get rid of the traits associated with a consultant's primary problem profile, but rather to recommend solutions or "fixes." Our approach is focused on achieving balance to mitigate negative outcomes like burnout.

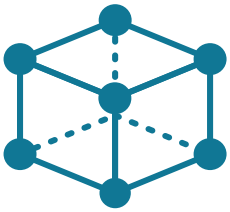
Problem Profiles Part 1: Scale Validity

CLARITY

For the Problem Profiles Scale to work, we had to train and test how responses to the scale were related to scores on each of the problem profiles and if the scale had acceptable psychometric properties. The first psychometric property we looked at was construct validity.

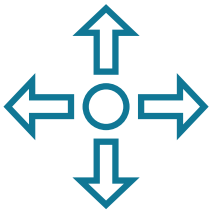
Construct Validity

Validity corresponds to the extent to which the scale accurately measures reality. Construct validity is an assessment as to whether or not the measure we created is measuring what we want it to measure. For example, is our measure of Workaholic really assessing a consultant's sparks and symptoms that contribute to their tendency to overwork themself? Or is it measuring something else? To test construct validity, we look at four areas:

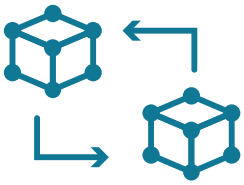


Structural Validity
Does the factor structure support that items are all measuring the same thing?

Convergent Validity
Is the construct related to other constructs it should theoretically be related to?



Divergent Validity
Is the construct unrelated to constructs it shouldn't be related to?



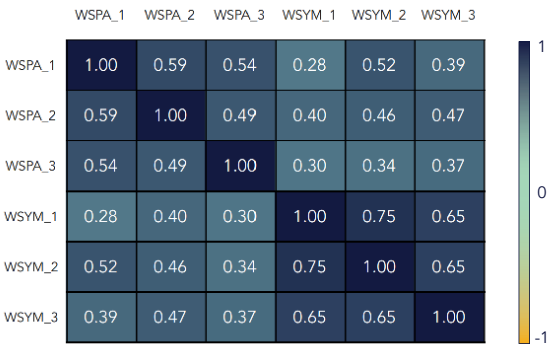
Nomological Validity
Does a network of constructs show relationships that are expected?

Construct Validity: Structural Validity.

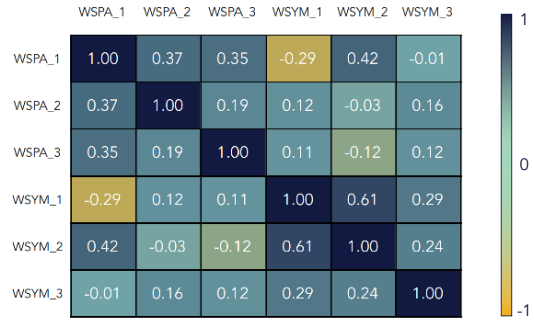
For the Workaholic profile, we want to make sure that items for Workaholic Spark are measuring spark, items for Workaholic Symptom are measuring symptoms, and together, the items are all measuring the Workaholic profile. To do so, we assess structural validity by using both exploratory factor analysis and confirmatory factor analysis.

Exploratory Factor Analysis.

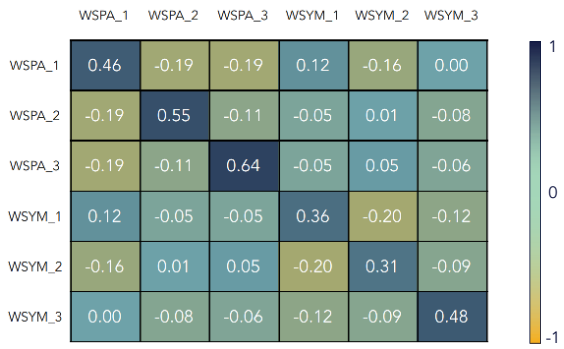
- Step 1: Correlation Check
  - » To determine which items to include or exclude in factor analysis, we first examined the **bivariate correlations** to identify any items with small bivariate correlations (r <.30). Items with correlations below this threshold were removed from the analysis and all others were retained. As you can see in the example below, the three items included in Workaholic Spark all have correlations above .3 with each other. Similarly, all three items in Workaholic Symptom have correlations above .3 with each other. Together, the items have correlations above .3 with each other. Therefore, all items for the Workaholic profile were retained.



  - » Traditional bivariate correlations only provide a part of the picture, so we also examined **partial correlations**. Partial correlations refer to the correlation between two items after controlling for the effect of all other items. In other words, partial correlations are the correlations that are left over after the common variance is extracted. As a rule of thumb, we include items with a partial correlation <.70 in the analysis and exclude items that exceed this threshold. As you can see in the example below, the three items included in Workaholic Spark all have partial correlations below .7 with each other. Similarly, all three items in Workaholic Symptom have partial correlations below .7 with each other. Together, all items have partial correlations below .7 with each other. Therefore, all items for the Workaholic profile were retained.

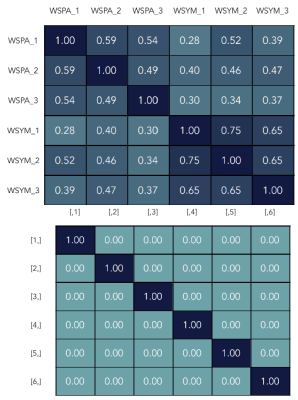


  - » We also look at the **anti-image correlation matrix**, which contains the negatives of the partial correlation coefficients. Consequently, these values are the magnitude of the variable that can't be regressed on, or predicted by, the other variables. If variables can't be regressed on, or predicted by, the other variables, then the variables are not likely related. If variables aren't related, then they will not likely load on the same factor. Consequently, large magnitudes indicate the possibility of a poor factor solution. However, as you can tell from the light colors in the corrogram heat map, all correlations in the ant-image correlation matrix are close to 0. This means all items on both constructs are retained.



CLARITY

» **Bartlet test of sphericity** compares the correlation matrix to the identity matrix, checking to see if there is any redundancy between the variables. High redundancy is indicative that the variables have common variance and therefore can be loaded on similar factors. If there is high redundancy, then the correlations in the correlation matrix should be higher in magnitude. Therefore, when it's compared to the identity matrix (where values are mainly 0), the two matrices will not be similar. If there is little redundancy, then the correlations in the correlation matrix should be close to zero. This means when it is compared to the identity matrix, the two matrices will be similar, indicating the possibility of a poor factor solution. In the case of the Workaholic profile, the correlation matrix was significantly different from the identity matrix.



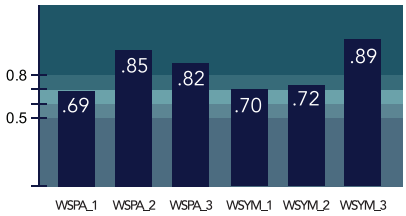
	WSPA_1	WSPA_2	WSPA_3	WSYM_1	WSYM_2	WSYM_3
WSPA_1	1.00	0.59	0.54	0.28	0.52	0.39
WSPA_2	0.59	1.00	0.49	0.40	0.46	0.47
WSPA_3	0.54	0.49	1.00	0.30	0.34	0.37
WSYM_1	0.28	0.40	0.30	1.00	0.75	0.65
WSYM_2	0.52	0.46	0.34	0.75	1.00	0.65
WSYM_3	0.39	0.47	0.37	0.65	0.65	1.00

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1.00	0.00	0.00	0.00	0.00	0.00
[2,]	0.00	1.00	0.00	0.00	0.00	0.00
[3,]	0.00	0.00	1.00	0.00	0.00	0.00
[4,]	0.00	0.00	0.00	1.00	0.00	0.00
[5,]	0.00	0.00	0.00	0.00	1.00	0.00
[6,]	0.00	0.00	0.00	0.00	0.00	1.00

» Lastly, the **Kaiser-Meyer-Olkin Measure of Sampling Adequacy** measures the extent to which the variance of the items might be caused by an underlying factor. The higher proportion of variance caused by underlying factors, the better your factor solution might be. Consequently, the following is what we use to determine whether or not to continue with the factor analysis.

> .80	Meritorious
> .70	Middling
> .60	Mediocre
> .50	Miserable
< .49	Unacceptable

There is cause for concern, if the KMO drops below .60. In the case of Workaholic Spark, two items have values above .80, indicating that they are meritorious and one item with a value before 0.69 indicating it is nearly middling. In the case of Workaholic Symptoms, two items have values above .70 indicating they are middling and one item has a value above .80 indicating that it is meritorious. Since KMO represents sampling adequacy, these results are unsurprising as a limited number of consultants were sampled for the study, but not concerning enough to discontinue the analysis.

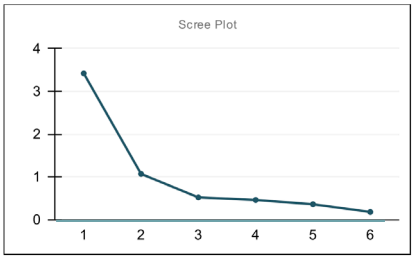


- Step 2: Factor Check. Once the correlations check out for each profile, and a final list of items are retained, we then run the factor analysis. The first thing to consider in this process is how many factors to retain in the solution. To determine this, we use two general principles:

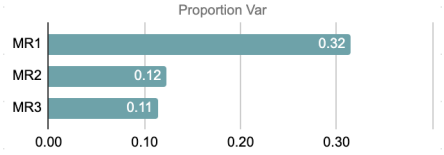
» Only retain factors with eigen values > 1.

Eigen values

3.42	1.07	0.52	0.46	0.36	0.18



» Only retain factors with variance > 5% OR factors whose variance sum to 60% or more.



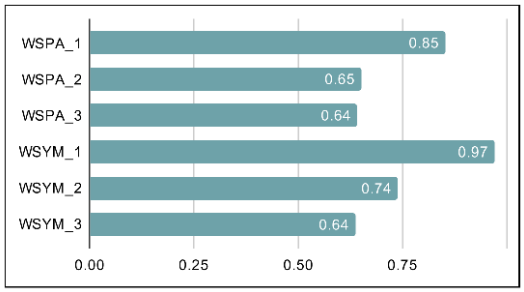
» In both cases, the data represent a solution for two factors: Workaholic Spark and Workaholic Solution indicating that items measuring the Workaholic profile can be represented with the Spark factor and Solution factor.

- **Step 3: Item Check.** Once we've decided on the number of factors that should be retained, the question becomes what items are associated with the factors (and which items are not).

» For practical significance of factor loadings, we follow the below approach:

> .70	Indicative of a well-defined structure
.50 - .69	Practically significant
.30 - .49	Minimally viable for a factor structure
< .30	Unrelated

You can see the following example:



» Items for Workaholic Spark and Workaholic Symptom have factor loadings above .6. This means that items for both factors are practically significant. Additionally, items did not cross-load across factors. Analytically, this means that items measuring Workaholic Spark didn't have factor loadings greater than .3 for Workaholic Symptoms and vice versa. Conceptually, items for Workaholic Spark measured spark and items for Workaholic Symptoms measured symptoms. Together, the correlation between the factors was .84, indicating that while items can be separated to spark and symptom, they still overlap and collectively measure the Workaholic Profile.

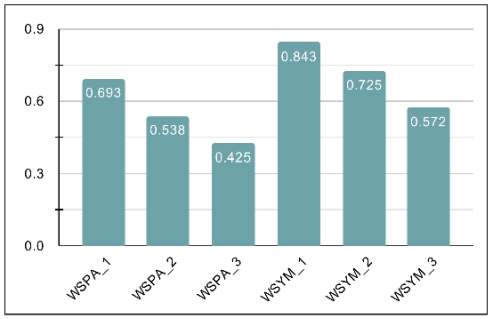
» For statistical significance of factor loadings, there are a few different approaches that researchers can take. However, factor loadings significance changes as a function of sample size. Consequently, we generally adhere to the following significance of factor loadings given the sample size.

Factor Loading	Sample Size Needed for Significance*
.30	350
.35	250
.40	200
.45	150
.50	120
.55	100
.60	85
.65	70
.70	60
.75	50

*Significance is based on a .05 significance level (α), a power level of 80 percent, and standard errors assumed to be twice those of conventional correlation coefficiencies

» Even though our sample size was limited (~100), we can still conclude that factor loadings for Workaholic Spark and Workaholic Symptom are statistically significant.

- **Lastly, when determining what items to retain, we look at communalities.** Communalities are the proportion of each variable's variance that can be explained or accounted for by the factors. As a general rule of thumb, we shoot for Communalities > .5 (i.e., retaining items in which a half of the variance of each variable should be accounted for by the factor solution).



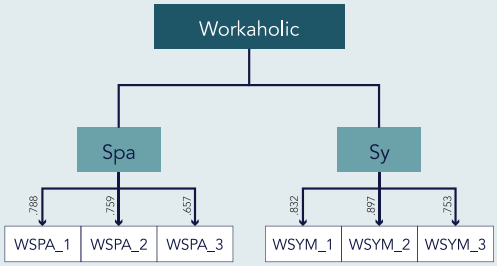
» Notice that two of the three items for Cognitive Integration are below the .30 threshold. Similarly, all three items for Cognitive Tolerance are below the .30 threshold. Results like this indicate that while the scale and items are still usable given all the other criteria they've passed, the items in each of these scales will need to go through a revising process to improve their psychometric properties.

Confirmatory Factor Analysis.

Exploratory Factor Analysis is only half of the equation. At Inkblot Analytics, we also use Confirmatory Factor Analysis to help with structural validity. While exploratory factor analysis was a data-driven approach, confirmatory factor analysis is a theory-based approach that helps us "confirm" if our theory matches the data. There are four things we look for in a confirmatory factor analysis that supports structural validity:

- **Standardized loading estimates should be high.** Standardized loading estimates are the same as standardized regression coefficients—they quantify the magnitude and direction of the relationship between the item and the factor. We want the relationship between the item and the factor to be high, therefore standardized loadings estimates must be high to be retained for adequate structural validity. More specifically, we use the accompanying rule. Notice, items for Workaholic Spark and Workaholic Symptom load highly on their respective factors.

> .50	Ideal
.30 - .49	Minimally Viable
< .30	Unacceptable



- **Standardized residuals should be small.** Standardized residuals are a calculation of the error in a model. Basically, it is a calculation of the magnitude of difference between observed and expected values. If our factor structure is not valid, then there is likely to be more error. Consequently, for items to be retained, we look for low values.
  - » Notice that the values for both Workaholic Spark and Workaholic Symptom are less than .2. This means that the expected values are a close match to the observed values. Very little error was produced when we estimated our theoretical model.

< .20	No problem
.21 - .39	"Red flag"
> .40	Unacceptable

	WSPA_1	WSPA_2	WSPA_3
WSPA_1	0.000		
WSPA_2	-0.019	0.000	
WSPA_3	0.040	-0.015	0.000

	WSYM_1	WSYM_2	WSYM_3
WSYM_1	0.000		
WSYM_2	0.015	0.000	
WSYM_3	0.049	-0.054	0.000

- **Model Indices should be small.** Modification indices represent the improvement a model would see (that is, improvement in units of chi-square values) if a particular relationship was added or deleted to the model. For a factor structure to be structurally valid, we want to minimize the number of modification indices and their values. Our current rule of thumb is that modification indices > 4 suggests improvements can be made to the model and therefore represent a poor factor structure.
  - » CFA is a theoretically guided analysis. So the researcher must be selective in what modification indices to use. The algorithm will give any/all modifications that can be made to your model, not just the ones that are theoretically relevant. In this case, four modification indices were flagged. The pathways recommended were to add correlation paths between items of Symptom and between items of Symptom and Spark. This is not surprising, as all items are related to each other and adding any of these paths would not change the interpretation of the model. The modification suggestions are theoretically trivial. To be through, we tested each modification suggestion and did not find a significant improvement in model fit for all. In other words, adding any of the four recommended paths had a minimal (non-significant) impact on our final conclusions, so we retained our hypothesized model.

- Model Fit Indices should indicate a good fit. Lastly, there are a number of model fit values that provide an overall assessment of how well the model fits the data. We use many of these to assess model performance and overall structural validity. The table below will show you what values we use for our cutoff.

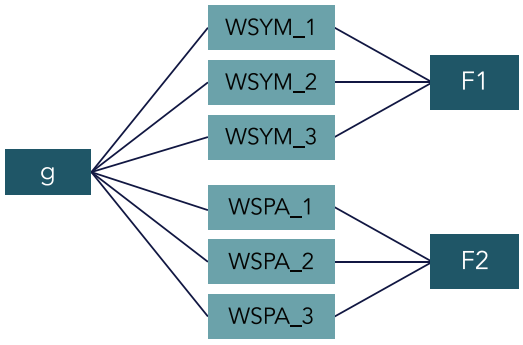
Factor Loading	Standard for Acceptable Fit
TLI	> .90 (marginal fit) ; > .95 (good fit)
CFI	> .90 (marginal fit) ; > .95 (good fit)
RMSEA	< .08
PClose	> .05 (i.e., not statistically significant)
SRMR	< .08
CD	The closer to 1, the better the fit
AIC	When comparing models, the lower the better
BIC	When comparing models, the lower the better

CFI	.951
TLI	.908
RMSEA	.130
SRMR	.046

- AVE > .5. With CFA, the average variance extracted is calculated by the average of the variance explained by the factor for each item that loads on it. Said differently, it's the sum of the squared standardized loadings of all items on a factor, divided by the number of items on that factor. If an AVE < .50, then it suggests that error explains more about the item's variance than is explained by the factor structure. For both Workaholic Spark and Workaholic Symptom, the average variance extracted was greater than .50.

Bifactor Modeling.

So far, exploratory and confirmatory factor analysis has shown us that the data represent Workaholic Spark and Workaholic Solution. The high correlation between the factors indicates that both factors measure a similar construct. Conceptually, we know it is the Workaholic profile, but we still need to explore this analytically. Bifactor Modeling is an advanced statistical method that we use at Inkblot to assess the structural validity of our scales. In this case, we use bifactor modeling to analytically explore if our scale measures Workaholic Spark, Workaholic Symptom, and in general the Workaholic profile simultaneously. The workaholic bifactor model looks like:



Notice that all items not only load well on the general Workaholic profile, but also they also load well on their respective factors. This tells us that our items can be split into Spark and Symptom AND together they represent the Workaholic profile.

Construct Validity: Convergent & Divergent Validity.

Another form of construct validity is known as convergent and divergent validity. Convergent validity refers to the relationship between variables that should be theoretically related. Divergent validity refers to the relationship between variables that should not be theoretically related.

With Regular Bivariate Correlations.

Another form of construct validity is known as convergent and divergent validity. Convergent validity refers to the relationship between variables that should be theoretically related. Divergent validity refers to the relationship between variables that should not be theoretically related.

> .90	Indicates the same construct
.70 - .89	Convergent validity for highly related constructs
.50 - .69	Convergent validity for somewhat related constructs
.40 - .49	No man's land
.20 - .39	Divergent validity for somewhat unrelated constructs
.10 - .19	Divergent validity for highly unrelated constructs
0 - .09	Indicates no relationship

With Confirmatory Factor Analysis.

Another form of construct validity is known as convergent and divergent validity. Convergent validity refers to the relationship between variables that should be theoretically related. Divergent validity refers to the relationship between variables that should not be theoretically related.

- AVE > Correlation. Convergent validity is supported by finding two constructs are related, but are NOT the same construct. For this to be shown, the variance extracted by a factor should be GREATER than the variance explained by the related construct. So when doing a CFA, we're looking for the AVE for two factors to be greater than the correlation between the two factors.
- A model with cross-loadings should be a poorer fitting model. When performing a CFA, if construct validity is to be theoretically supported, there should not be any cross-loaded items. If there were to be cross loaded items, removing them should make the model better. To test this out, we force some items to cross-load (that is, load on to the original construct and the related construct). By doing this, your model should get worse. If it gets better, then you know both constructs might be measuring the same thing.

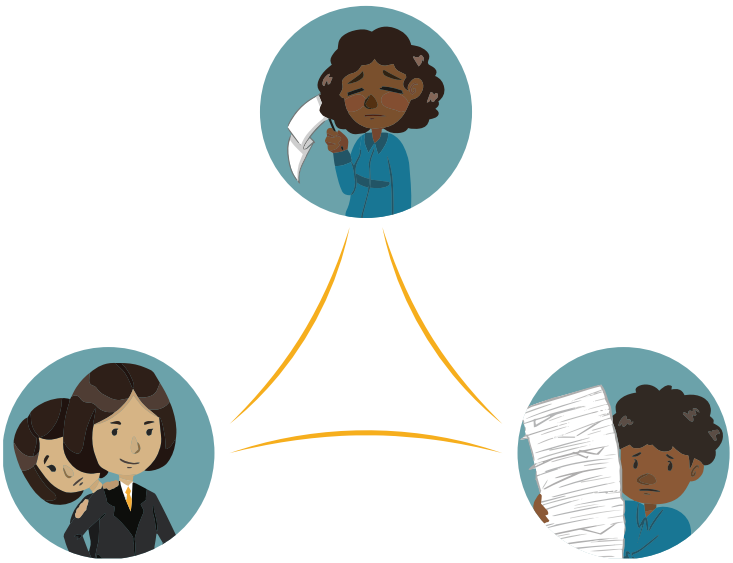
With Bifactor Modeling.

- Test a bifactor model and see if it gets worse. A bifactor model is usually used when you want to test the presence of a general factor that all items load onto. This approach helps identify the plausibility of a scale having multiple factors that are theoretically uncorrelated. factors to be greater than the correlation between the two factors.

Convergent and divergent validity analysis are add-on features.

Construct Validity: Nomological Validity.

Typically, at Inkblot Analytics, we use other construct types for convergent and divergent validity, while using variables from the same construct type for nomological variability. For nomological validity we look at a correlation matrix and identify the biggest correlations. In theory these relationships should correspond to how you would theoretically think variables within the same construct type would be related. For example we found the following correlations:



- The higher a consultant scores on the workaholic profile, the higher they score on the imposter profile.
- The higher a consultant scores on the workaholic profile, the higher they score on jack-of-all-trades.
- The higher a consultant scores on the jack-of-all-trades profile, the higher they score on the imposter profile.

These relationships between constructs make sense, as a consultant who identifies with one problem profile is likely to identify with another.

Item-to-total correlations > .5.

One of the first things we look at is to what extent each scale item correlates with a composite score of the scale (i.e., with all items for the scale scored properly). Generally speaking, we look for an item-to-total correlation of at least .50. When looking at the scores for Workaholic Spark, we get the following:

	raw.r
WSPA_1	0.6805
WSPA_2	0.6803
WSPA_3	0.5719

Notice all items are above the .50 threshold. Similarly, when looking at the scores for Workaholic Symptom, we get the following:

	raw.r
WSYM_1	0.7225
WSYM_2	0.8125
WSYM_3	0.7268

Again all items are above the .50 threshold.

CFA's Composite Reliability >.70.

We calculated the composite reliability of the CFA models. This includes both Alpha and Omega values of reliability. Generally speaking, we use the following criteria:

> .70	Suggests good reliability
.60 - .69	Acceptable

As you can see below, both Workaholic Spark and Symptom meet the .70 threshold for reliability. Additionally, the general Workaholic problem profile meets the reliability threshold.

	Workaholic	Symptom	Spark
Alpha	0.85	0.78	0.87
Omega	0.89	0.76	0.87

Chronbach's Alpha > .70.

One of the most prolific ways of checking scale reliability is by calculating Chronbach's alpha. When calculating scale reliability at Inkblot Analytics, we use the following standards:

Cronbach's alpha	Internal consistency
$\alpha \geq 0.9$	Excellent
$0.9 \geq \alpha \geq 0.8$	Good
$0.8 \geq \alpha \geq 0.7$	Acceptable
$0.7 \geq \alpha \geq 0.6$	Questionable
$0.6 \geq \alpha \geq 0.5$	Poor
$0.5 > \alpha$	Unacceptable

When looking at Workaholic Spark, scale reliability is .78, indicating acceptable internal consistency. It flags to our internal team to keep revising the measure. For Workaholic Symptom, scale reliability is .87. This falls into the good internal consistency range. Similarly, the general Workaholic construct has a reliability of .85.

Summary

CLARITY

At this point, I hope you can see just how much rigor goes into the platform, Clarity and the Problem Profiles Scale.

From the perspective of scale construction and use, scales must have adequate psychometric properties to be used. Both example scales reported on in this paper--Workaholic Spark and Workaholic Symptom--have good to excellent psychometric properties.

From the perspective of our machine learning, our algorithms get checked for their error on a regular basis, with overall error reaching +/-9 points from a person's true score.

No matter what part of the tool you're looking at, our results are backed by a rigorous vetting process.