Brand Blots

We know psychology is trendy in market research these days. This means 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 Brand Blots, highlighting the validity and reliability of key aspects of the platform. 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 this platform

How trustworthy the results are

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

Who are we? Brand Blots is a new patent-pending platform for AI-powered projective tests. This platform allows market researchers to administer, score, and interpret projective tests in a fast and efficient manner. With its proprietary AI modules, Brand Blots can predict a respondent's scores on over 100 attitudinal and behavioral aspects with great accuracy.

How do we do it? Brand Blots uses a unique combination of psychological science and data science to produce the first truly intelligent AI-powered projective test technology. This technology is built on the following underlying principles:

- Consumers have a variety of different psychological profiles.
- These psychological profiles are made up of 100's of different psychological traits.
- These psychological traits are detectable in the things you say and do.
- Brand Blots uses the things consumers' say and do in our proprietary projective tests to identify their psychological profile.
- These psychological traits are detectable in the things you say and do.

What insights can you learn? Many of our clients have told us stories of how they commissioned research from other research vendors in which the report was delivered with no real marketing strategy or tactics as implications of the research. That's why we built marketing implications directly into our platform. For example, not only can you learn about the psychographics of an individual, you can also profile, what we call, their consumption outcomes. Consumption outcomes are individual difference measures that apply directly to the way the shop and buy. For example, we have measures of the frequency at which they visit the grocery store, average spend per store visit, percent of basket spent on healthy items, etc. Our platform includes a number profiling items that are category specific and, in some instances, brand specific. That way, you don't just get to learn that your average shopper is Extroverted, you also get to know that because they're extroverted, they spend more on social occasions, in which they'll buy branded products over generic products, even if the price difference is 33%.

Key Benefits of Brand Blots

Why use our platform? What's the core benefit for market researchers? While there are a number of ways the Brand Blots platform is beneficial, there are three we hear about most often from our clients:

Less Questions, More Insights. With our platform, market researchers can ask fewer questions while gaining 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. And our platform now does that.

Constant Flow of New Data. Our platform 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 platform, updating data may happen once or twice a year. However, when something dramatic like the Covid-19 pandemic comes along, and consumer behaviors shifts suddenly, you want a flexible and adaptable solution like Brand Blots that is constantly refreshing the thoughts, feelings, and behaviors of consumers.

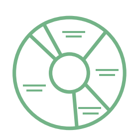
Better Open Ended Responses. If you've ever included an open ended response in your survey, you know consumers are always inclined to answer with one or two words--responses that are unusable for concluding any insights. With Brand Blots, consumers are more intrigued and interested in the task. They're curious and--to a degree--become more invested in providing a more detailed response. It is not gamification, but the engaging nature of how you interact with the task makes it more interesting to the consumer. This means more responses, with better quality responses, leading to more insights for the client.

How Does the Digital Inkblot Test Work?

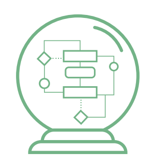
Brand Blots has over 36 types of projective tests. This paper is focusing specifically on the digital inkblot test. This inkblot test is a four step process:

Taking the digital inkblot test questions

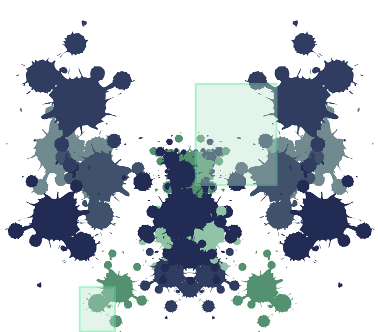
Coding the inkblot test responses



Measuring the traits you want to include in your participant's profile



Predicting the participant's scores on the referenced traits



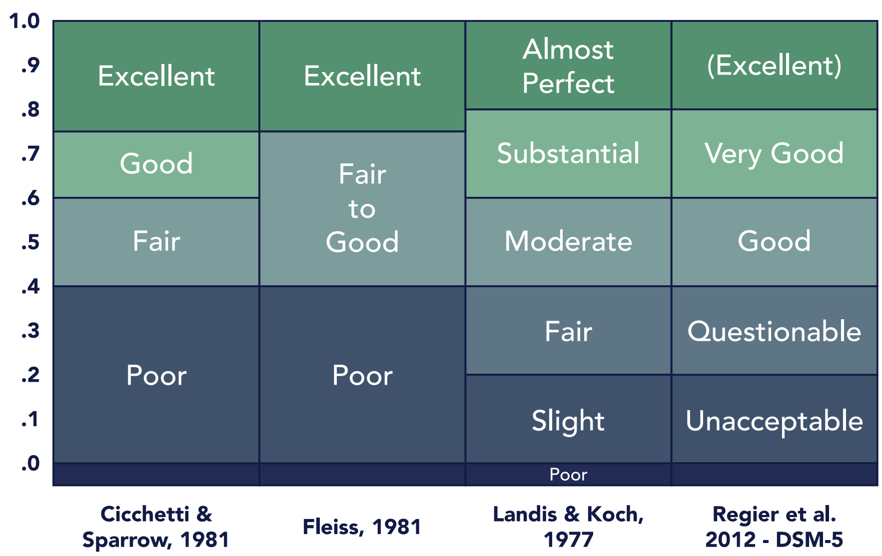
Each one of these steps has a scientific process built into them. For taking the inkblot test, we use an algorithm that measures the extent to which a respondent is intentionally trying to deceive the test, not take it seriously, or enter in bad quality data. For the coding step we use measures of inter-rater 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 model's error (the difference from the predicted score and actual score) to know how accurate/precise the models predictions are.

The Coding Step: Inter-rater Reliability

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 "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 are coding to agree on a minimum of 80% of the scoring they do.



The Profiling Step: Psychometrics of Measured Traits

Once the inkblot test data is scored, we're able to feed the scored results to our proprietary algorithms to help build a psychological profile. Often times, there are specific psychographics that a client wants to be a part of that profile, so we have the capability to select what traits to include or exclude from the profile. After all, with over 100 different psychological traits, including them all might not be as useful for the specific business challenge or marketing problem our clients have. However, regardless of what traits are included or excluded, we have to make sure we have solid psychometric properties of those constructs. Otherwise, we could say we're measuring extroversion, but we could really be measuring "likelihood to talk to strangers." In this section, we've picked a class of constructs called "Perceptual Style" and will take you through the science of how we validate the measures for each perceptual style measure.

A Set of Predicted Traits: Perceptual Style

This paper deals specifically with constructs that fall within the "Perceptual Style" or "Perceptual Attitude" construct type. Perceptual style is an individual differences variable that measures how individuals perceive their environment and organize information within it (Messick, 1984). Perceptual style is critical to understanding everyday behaviors such as interpersonal relations. For example, because perceptual style partially determines the way we process information, two people with conflicting perceptual styles may have an increased likelihood of miscommunication. Similarly, perceptual style also has concrete applications for marketers studying consumption outcomes. For example, how prospects search for (and process) information along the customer journey could greatly increase or decrease the likelihood of purchasing.

While there are a number of different perceptual styles in the literature, for the purposes of this paper we will only focus on ten. The ten perceptual styles used in this paper include:

Cognitive Scanning

A high score indicates the preference towards a fast review of some parts to make assumptions about the whole. For example, only reading the news headlines rather than the whole article itself. A low score indicates a preference towards a slow in-depth review of something in its entirety before processing the information.

Cognitive Tolerance

A high score indicates a comfortableness and even preference for working in unstructured situations or working with ambiguous items or on uncertain tasks. For example, someone who scores high on cognitive tolerance might be more likely to understand an advertisement using entendre, metaphors, or other devices that use the vagueness to engage in a kind of "doublespeak." A low score indicates a preference for structure, clarity, and literalness, making these individuals less likely to understand straightforward promo ads.

Cognitive Dependence

A high score indicates a preference for considering the historical, dependent, wholistic and relational nature of things when processing information. For example, when deciding between two brands to buy, someone who scores high on cognitive dependence would be more likely to disqualify a brand and not purchase them because of something they did two years ago. A low score indicates a preference for considering specific parts of the "current" information (i.e., from here-and-now). Someone who scores low might be more likely to be a "naive loyalist"—they like their preferred brand, regardless if competitors are cheaper, or have better functional benefits, etc. They like their preferred brand, for their solo reason and nothing else matters in the decision.

Cognitive Complexity

A high score on this construct indicates a preference for thinking about and processing information that is abstract, theoretical, and complex. This is juxtaposed with those who score low, who prefer to think and process information that is specific, concrete, and applied.

A Set of Predicted Traits: Perceptual Style (cont.)

Brand Blots

Cognitive Tempo	A high score on this construct indicates a preference for thinking slowly and deliberately about information. A low score on this constructs includes a preference for thinking and making decisions quickly.
Cognitive Integration	A high score on this construct indicates a preference for gathering and aggregating information to process (i.e., "integrate") it. A low score indicates a preference for taking the information one has and "pulling it apart" (i.e., disaggregating) it to process it.
Locus of Control	A high score on this construct indicates a preference for internal locus of control (i.e., attributing things to the self). A low score on this construct indicates a preference for external locus of control (i.e., attributing things to the external environment).
Sharpening / Leveling	A high score on this construct indicates a preference for details and differences between things. A low score on this construct indicates a preference for more general similarities among things.
Cognitive Consistency	A high score on this construct indicates a preference for consistency, uniformity, and regularity whereas a low score on this construct indicates a preference for inconsistency.
Cognitive Style	A high score on this construct indicates a preference for gathering and organizing information to process, such as by organizing it visually in a drawing, on a list, or otherwise.

As you can see, the way in which we interpret perceptual stimuli has a huge impact on how we see the world. This especially applies to how we process information before making a purchase. So being able to profile consumers on each of these traits is critical information for a brand manager to have. That's why we created our own proprietary scales.

Our Predicted Traits Part 1: Scale Validity

For Brand Blots to work, we had to train and test how responses to the digital inkblot test were related to scores on each of the perceptual style constructs. To get scores for each construct, we had to write and test scales with acceptable psychometric properties.The first psychometric property we looked at was construct validity.

Construct Validity.

Measuring whether or not something is "valid" means assessing the extent to which the measure corresponds to reality. Construct validity is an assessment as to whether or not the measure is measuring what we want it to measure. For example, is our measure of Cognitive Tolerance really assessing a person's preference for ambiguity? 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 theoretical 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?

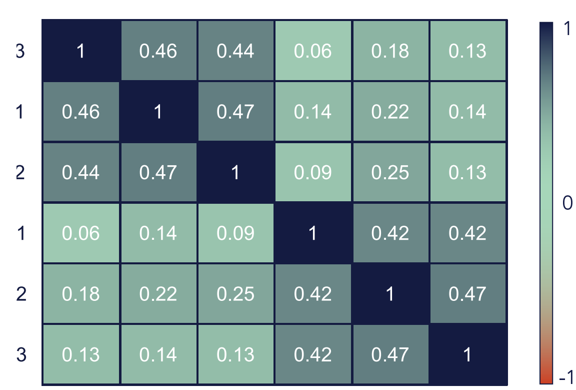
Structural Validity.

To assess structural validity, we look at models for both exploratory factor analysis and confirmatory factor analysis.

Exploratory Factor Analysis.

Exploratory Factor Analysis is a method by which a researcher can extract latent constructs that affect a person's behavior when responding to individual items. For example, the construct of Cognitive Tolerance is likely to affect the way a person responds to questions about their "comfort" with activities that vhave no clear goals or relationships that have no real label (e.g., undefined / it's complicated). The following steps were taken when creating the perceptual style measures:

- Step 1: Correlation Check
    - When determining what items to include or exclude in the factor analysis, we first looked for any items with small bivariate correlations ($r < .30$). Any correlations below that threshold were removed from the analysis. All other items were included.



- In addition to the traditional bivariate correlations, partial correlations were also run. Partial correlations are the correlation between two variables after controlling for the effects of all other variables. In effect, it's the correlation after the common variance is extracted. To support a factor analysis, researchers are assuming that there is a high degree of common variance between all the items. Therefore, we look for small partial correlations. Any items that have a partial correlation < .70 remain in the analysis, while any items that exceed that value are removed.

	Int_3	Int_1	Int_2	Tol_3	Tol_1	Tol_2
Int_3	1	0.16	0.17	-0.07	-0.11	-0.02
Int_1	0.16	1	0.21	-0.01	0.03	0.08
Int_2	0.17	0.21	1	-0.07	-0.03	0.05
Tol_3	-0.07	-0.01	-0.07	1	0.11	0.05
Tol_1	-0.11	0.03	-0.03	0.11	1	0.12
Tol_2	-0.02	0.08	0.05	0.05	0.12	1

Scale: 1, 0, -1

- 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.

	Tol_2	Int_2	Tol_3	Int_1	Int_3	Tol_1
Tol_2	1	-0.05	-0.05	-0.08	0.02	-0.12
Int_2	-0.05	1	0.07	-0.21	-0.17	0.03
Tol_3	-0.05	0.07	1	0.01	0.07	-0.11
Int_1	-0.08	-0.21	0.01	1	-0.16	-0.03
Int_3	0.02	-0.17	0.07	-0.16	1	0.11
Tol_1	-0.12	0.03	-0.11	-0.03	0.11	1

Scale: 1, 0, -1

- 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 the correlation matrix and identity matrices are similar then this is indicative of the possibility of a poor factor solution. However, if the two matrices are not similar, and the correlation matrix diverges significantly from the identity matrix, then there is support to continue to the next step of the analysis.

Correlation Matrix

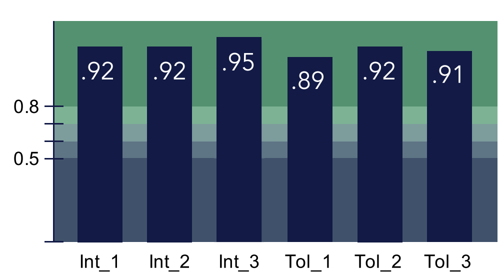
	Int_1	Int_2	Int_3	Tol_1	Tol_2	Tol_3
Int_1	1.0000000	0.46539508	0.45534639	0.14061258	0.2222927	0.1402906
Int_2	0.4653951	1.0000000	0.44277328	0.08653084	0.251622	0.1304822
Int_3	0.4553464	0.44277328	1.0000000	0.06386793	0.1819106	0.1342197
Tol_1	0.1406126	0.08653084	0.06386793	1.0000000	0.4244572	0.4225749
Tol_2	0.2222927	0.25162204	0.18191058	0.42445719	1.0000000	0.4724250
Tol_3	0.1402906	0.1413048224	0.13421975	0.42257457	0.472425	1.0000000

Identity Matrix

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1	0	0	0	0	0
[2,]	0	1	0	0	0	0
[3,]	0	0	1	0	0	0
[4,]	0	0	0	1	0	0
[5,]	0	0	0	0	1	0
[6,]	0	0	0	0	0	1

- 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



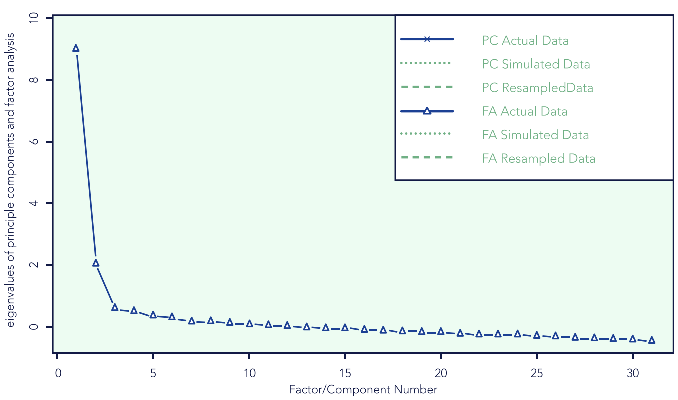
Our Predicted Traits Part 1: Scale Validity (cont.)

- **Step 2: Factor Check.** Once the correlations check out, 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 eigenvalues > 1

Eigen Values

9.6964339	2.8167664	1.3221164	1.1902187	1.072968	1.0061014	0.8869513	0.8741731
0.8559972	0.7870692	0.7528272	0.6955619	0.6796559	0.6707243	0.6497840	0.6220614
0.5939534	0.5566932	0.5498964	0.5111148	0.4667852	0.4495140	0.4391825	0.4286129
0.4052742	0.3896525	0.3720145	0.3558822	0.3159930	0.3052990	0.2807217	

Scree Plot



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

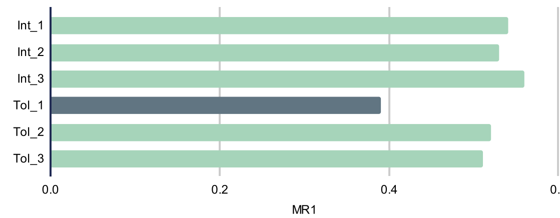
FACTOR VARIANCE FOR A PA SCALE HERE

  - » **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:

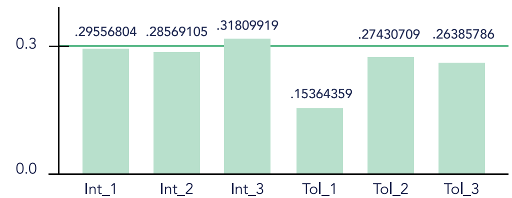


» For statistical significance of factor loadings, there are a few different approaches that researchers can take. However, factor loadings significance changes as a function of sample size. Consequently, we generally adhere to the following significance of factor loadings given the sample size.

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

- **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).



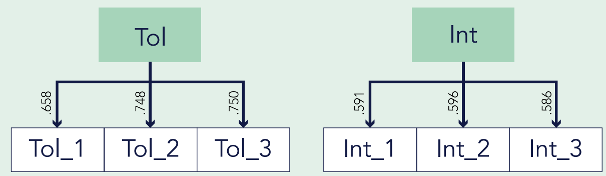
Our Predicted Traits Part 1: Scale Validity (cont.)

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. Thereare 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.

> .70	Ideal
.50 - .69	Minimally Viable
< .50	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.

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

Example residuals for a PA scale

- 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 significant improvements can be made to the model and therefore represent a poor factor structure.

Example modification indices for a PA scale

Our Predicted Traits Part 1: Scale Validity (cont.)

Brand Blots

- 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 can a number 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

- AVE > .5. 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.

Construct Validity: Convergent Validity.

Another form of construct validity is known as convergent validity (Campbell & Fiske, 1959). Convergent validity refers to the relationship between variables that should be theoretically related. We assess this in two core ways:

- With regular bivariate correlations. When looking to support convergent validity, the use of bivariate correlations can show us just how related different measures are. At Inkblot Analytics, we use the accompanying rules of thumb.

> .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

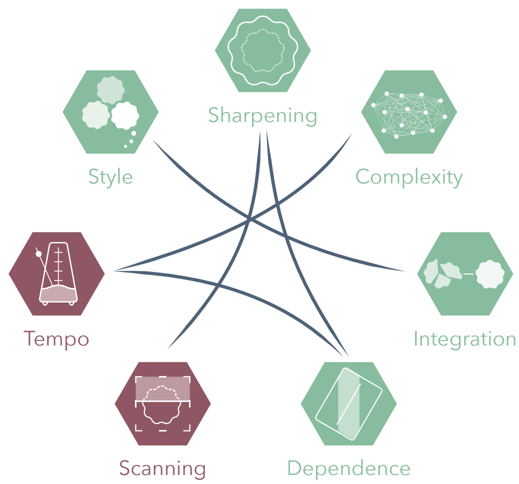
- With Confirmatory Factor Analysis (CFA).
  - » AVE > Correlation. Convergent validity is supported by finding two constructs are related, but are NOT the same construct. For this to be shown, the variance extracted by a factor should be GREATER than, the variance explained by the related construct. So when doing a CFA, we're looking for the AVE for two factors to be > 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, you 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.
  - » Test a two factor model and see if fit gets worse. I mean… how much more can you really say?

Construct Validity: Discriminant Validity.

- CFA. Compare the AVE with the square of the correlation estimates. AVE > squared correlation estimates.

Construct Validity: Nomological Validity.

Typically, at Inkblot Analytics, we use other construct types for convergent and discriminant 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 you score on cognitive dependence, the higher you score on cognitive sharpening (looking for differences and details).
- The higher you score on cognitive dependence the lower you score on cognitive tempo (slow, deliberate thinking).
- The lower you score on cognitive tempo (slow deliberate thinking) the higher you score on cognitive complexity.
- The lower you score on cognitive scanning (a deep reading instead of a scan), the higher you score on cognitive complexity.
- The lower you score on cognitive scanning (a deep reading instead of a scan), the higher you score on sharpening (looking for detail and difference).
- The higher you score on cognitive style, the higher you score on cognitive integration.

These relationships between constructs within the same construct class make sense.

Our Predicted Traits Part 2: Scale Reliability

Brand Blots

Construct Validity: Discriminant Validity.

- Item-to-total correlations > .5
- CFA's Construct Reliability

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

Our Predictions: Model Fit

TEXT WILL GO HERE

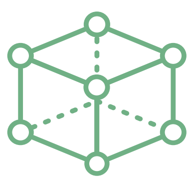
References

- TEXT

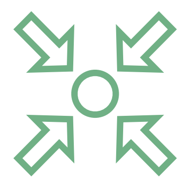
For Brand Blots to work, we had to train and test how responses to the digital inkblot test were related to scores on each of the perceptual style constructs. To get scores for each construct, we had to write and test scales with acceptable psychometric properties.The first psychometric property we looked at was construct validity.

Construct Validity.

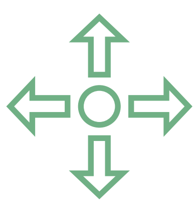
Measuring whether or not something is "valid" means assessing the extent to which the measure corresponds to reality. Construct validity is an assessment as to whether or not the measure is measuring what we want it to measure. For example, is our measure of Cognitive Tolerance really assessing a person's preference for ambiguity? Or is it measuring something else? To test construct validity, we look at five areas:



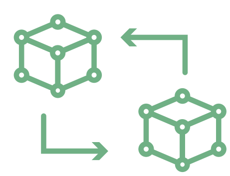
Structural Validity
Does the factor structure support that items are all measuring the same thing?



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

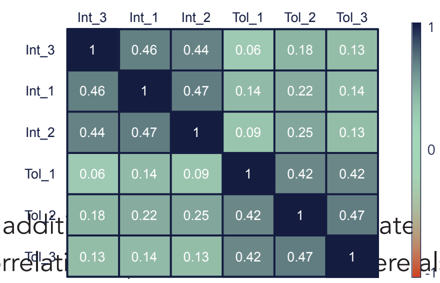
Structural Validity.

To assess structural validity, we look at models for both exploratory factor analysis and confirmatory factor analysis.

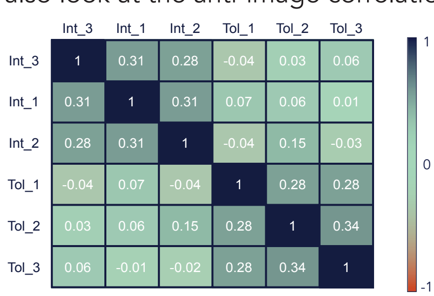
Exploratory Factor Analysis.

Exploratory Factor Analysis is a method by which a researcher can extract latent constructs that affect a person's behavior when responding to individual items. For example, the construct of Cognitive Tolerance is likely to affect the way a person responds to questions about their "comfort" with activities that have no clear goals or relationships that have no real label (e.g., undefined / it's complicated). The following steps were taken when creating the perceptual style measures:

- Step 1: Correlation Check
    - » When determining what items to include or exclude in the factor analysis, we first looked for any items with small bivariate correlations (r < .30). Any correlations below that threshold were removed from the analysis. All other items were included.



    - » In addit ate correlati ere also run. Partial correlations are the correlation between two variables after controlling for the effects of all other variables. In effect, it's the correlation after the common variance is extracted. To support a factor analysis, researchers are assuming that there is a high degree of common variance between all the items. Therefore, we look for small partial correlations. Any items that have a partial correlation < .70 remain in the analysis, while any items that exceed that value are removed.
    - » For statistical significance of factor loadings, We also look at the anti-image correlation

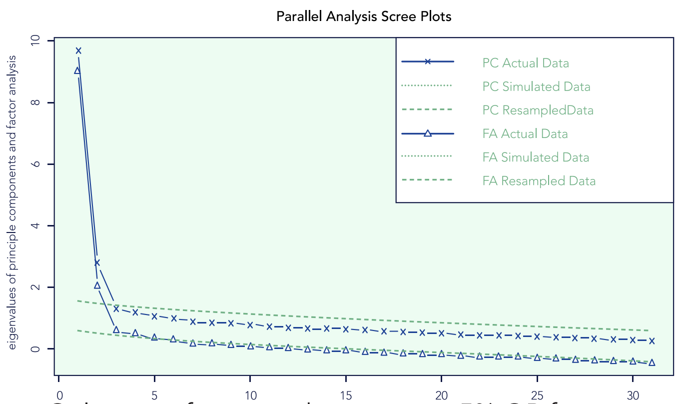


matrix, which contains the negatives of the partial correlation coefficients. Consequently, these values are the magnitude of the variable that can't be regressed on, or predicted by, the other variables. If variables can't be regressed on, or predicted by, the other variables, then the variables are not likely related. If variables aren't related, then they will not likely load on the ame factor. Consequently, large magnitudes indicate the possibility of a poor factor solution.

    - » 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 the correlation matrix and identity matrices are similar then this is indicative of the possibility of a poor factor solution. However, if the two matrices are not similar, and the correlation matrix diverges significantly from the identity matrix, then there is support to continue to the next step of the analysis.
    - » 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.

> .8	Meritorious
> .7	Middling
> .6	Mediocre
> .5	Miserable
< .49	Unacceptable

- Step 2: Factor Check. Once the correlations check out, 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 eigenvalues > 1

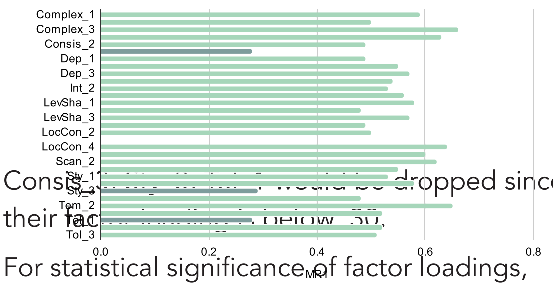


Only retain factors with variance > 5% OR factors whose variance sum to 60% or more

FACTOR VARIANCE FOR A PA SCALE HERE

- ... e 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



Consis_3 ... would be dropped since their fac... below .30.

- » 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 coefficients.

- Lastly, when d... to retain, we look at cor... alities are the propor... variance that can be ex... or by the factors. As a g... 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).

Example communality for a PA scale

Our Predicted Traits Part 1: Scale Validity (cont.)

Brand Blots

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. Thereare 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.

> .70	Ideal
.50 - .69	Minimally Viable
< .50	Unacceptable

Example loadings for a PA scale

- 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.

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

Example residuals for a PA scale

- 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 significant improvements can be made to the model and therefore represent a poor factor structure.

Example modification indices for a PA scale

Our Predicted Traits Part 1: Scale Validity

Brand Blots

For Brand Blots to work, we had to train and test how responses to the digital inkblot test were related to scores on each of the perceptual style constructs. To get scores for each construct, we had to write and test scales with acceptable psychometric properties.The first psychometric property we looked at was construct validity.

Construct Validity.

Measuring whether or not something is "valid" means assessing the extent to which the measure corresponds to reality. Construct validity is an assessment as to whether or not the measure is measuring what we want it to measure. For example, is our measure of Cognitive Tolerance really assessing a person's preference for ambiguity? 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 theoretical 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?

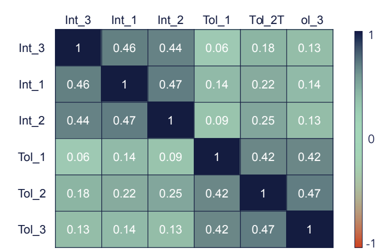
Structural Validity.

To assess structural validity, we look at models for both exploratory factor analysis and confirmatory factor analysis.

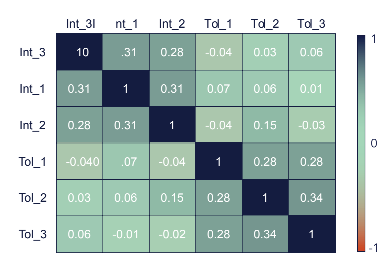
Exploratory Factor Analysis.

Exploratory Factor Analysis is a method by which a researcher can extract latent constructs that affect a person's behavior when responding to individual items. For example, the construct of Cognitive Tolerance is likely to affect the way a person responds to questions about their "comfort" with activities that vhave no clear goals or relationships that have no real label (e.g., undefined / it's complicated). The following steps were taken when creating the perceptual style measures:

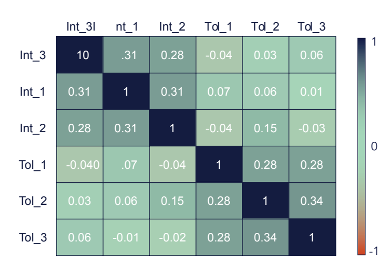
- Step 1: Correlation Check



  - » When determining what items to include or exclude in the factor analysis, we first looked for any items with small bivariate correlations ($r < .30$). Any correlations below that threshold were removed from the analysis. All other items were included.



  - » In addition to the traditional bivariate correlations, partial correlations were also run. Partial correlations are the correlation between two variables after controlling for the effects of all other variables. In effect, it's the correlation after the common variance is extracted. To support a factor analysis, researchers are assuming that there is a high degree of common variance between all the items. Therefore, we look for small partial correlations. Any items that have a partial correlation < .70 remain in the analysis, while any items that exceed that value are removed.



  - » 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.

Our Predicted Traits Part 1: Scale Validity

For Brand Blots to work, we had to train and test how responses to the digital inkblot test were related to scores on each of the perceptual style constructs. To get scores for each construct, we had to write and test scales with acceptable psychometric properties.The first psychometric property we looked at was construct validity.

Construct Validity.

Measuring whether or not something is "valid" means assessing the extent to which the measure corresponds to reality. Construct validity is an assessment as to whether or not the measure is measuring what we want it to measure. For example, is our measure of Cognitive Tolerance really assessing a person's preference for ambiguity? Or is it measuring something else? To test construct validity, we look at five 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 theoretical 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?

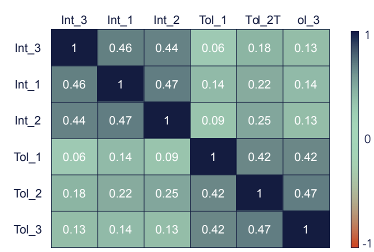
Structural Validity.

To assess structural validity, we look at models for both exploratory factor analysis and confirmatory factor analysis.

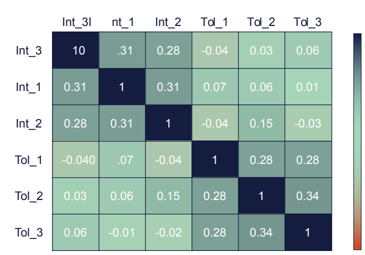
Exploratory Factor Analysis.

Exploratory Factor Analysis is a method by which a researcher can extract latent constructs that affect a person's behavior when responding to individual items. For example, the construct of Cognitive Tolerance is likely to affect the way a person responds to questions about their "comfort" with activities that vhave no clear goals or relationships that have no real label (e.g., undefined / it's complicated). The following steps were taken when creating the perceptual style measures:

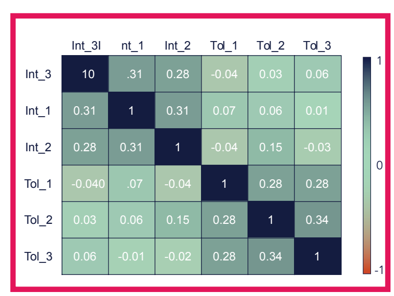
- Step 1: Correlation Check
    - » When determining what items to include or exclude in the factor analysis, we first looked for any items with small bivariate correlations (r < .30). Any correlations below that threshold were removed from the analysis. All other items were included.



    - » In addition to the traditional bivariate correlations, partial correlations were also run. Partial correlations are the correlation between two variables after controlling for the effects of all other variables. In effect, it's the correlation after the common variance is extracted. To support a factor analysis, researchers are assuming that there is a high degree of common variance between all the items. Therefore, we look for small partial correlations. Any items that have a partial correlation < .70 remain in the analysis, while any items that exceed that value are removed.



    - » 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.



Our Predicted Traits Part 1: Scale Validity (cont.)

Brand Blots

- 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 the correlation matrix and identity matrices are similar then this is indicative of the possibility of a poor factor solution. However, if the two matrices are not similar, and the correlation matrix diverges significantly from the identity matrix, then there is support to continue to the next step of the analysis.

- 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.

> .8	Meritorious
> .7	Middling
> .6	Mediocre
> .5	Miserable
< .49	Unacceptable