I-Factor[®] Technical Paper

Brand Irresistibility Scale

WHITEPAPER

Psychology is trendy in market research these days.

There are a number of Alpowered 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 I-Factor, 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 the platform I-Factor



How trustworthy the results are



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

FACTOR®

In today's world, the marketplace and the way consumers receive and digest information have changed dramatically. Consumers are hypersaturated by lackluster brands competing for consumer attention. At the same time, consumers are increasingly demanding better access to and involvement with the growth of a brand. While love for a brand is important to achieve long-lasting consumer relationships, brand love can only go so far. Thus, brands have to craft relationships with consumers that go above and beyond brand love. Sawtooth helps detect brand equities that propel the brand from "lovable" to "irresistible", so that brands achieve long-lasting consumer relationships.

I-Factor is a new platform that uses **patent-pending systems** for Alpowered projective tests. This platform helps brands identify their brand irresistibility score - the intangible connection between a brand and a consumer. I-Factor offers a true barometer of a consumer's relationship with the brand relative to its competitors.

How do we do it? I-Factor 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:

- Traits of a brand affect all areas of the consumer-brand relationship
- Tapping into and enhancing traits that promote a brand's Irresistibility can help brands achieve long-lasting consumer relationships that increase their profitability.
- Brand irresistibility that can also be broken down into the 3 C's of Irresistibility: Comprehend, Crave, and Craze to create the I-Factor scale
- The I-Factor scale surveys consumers on their perception of the brand's irresistibility. Specifically, how much consumers know about a brand and its product, how deeply do consumers connect with the brand, and how willing are consumers to engage for a brand?
- The I-Factor platform offers brands solutions and detailed suggestions to enhance their emotional relationship with consumers

What can you learn? By using I-Factor, a brand can get access to insights about their:

- 1. Brand Irresistibility Scores or the I-Factor Score
- 2. Identify trends and context of I-Factor Scores relative to competitor brands
- 3. Get an in depth understanding of "Why" behind their I-Factor Score



- 4. Compare scores by competitors
- 5. Custom recommendations to improve their I-Factor score
- 6. Track changes to monitor brand progress

Why use our platform? While there are a number of ways I-Factor is beneficial, there are three primary benefits of I-Factor:

Automated Predictive AI. I-Factor automatically creates predictive algorithms unique to each brand. The predictive AI not only provides brand with their I-Factor Scores, but it also (1) predicts consumer satisfaction, recommendation, purchasing likelihood and other metrics to show the performance of the brand by I-Factor scores, and (2) automatically give solutions and recommendations for the brand strategy when they got different scores. The predictive AI grows as the brand grows, getting better each time it is used.

Less Questions, More Insights. With I-Factor, brands can get more insights by asking less questions. 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.

Both Quantitative and Qualitative. I-Factor studies both quantifiable data and emotional insight using visual metaphors, and there are more than 30 specialized questions to tap into a consumer's subconscious. More importantly, I-Factor can uncover consumers' deep-seated thoughts and feelings that they wouldn't normally give in a typical survey or interview.

How Does I-Factor Work?

This paper is focusing specifically on the I-Factor scale. I-Factor is a four step process:







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 brands

Each one of these steps has a scientific process built into them. For the **testing step** (i.e., when the participant takes the I-Factor 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

Due to the high velocity of data we sometimes receive, we have multiple coders who apply a specific scoring scheme to the image interpretation 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 image interpretation 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 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 accompanying 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.

This section is an add-on service.

1.0				
.9	Excellent	Excellent	Almost Perfect	(Excellent)
.8				
.7	Good	Fair	Substantial	Very Good
.6		to		
.5	Fair	Good	Moderate	Good
.4				
.3			Fair	Questionable
.2	Poor	Poor		
.1			Slight	Unacceptable
.0				
			Poor	
	Cicchetti & Sparrow, 1981	Fleiss, 1981	Landis & Koch, 1977	Regier et al. 2012 - DSM-5

The Profiling Step: Psychometrics of Brand Irresistibility

Once the test data is collected, we are able to use our proprietary algorithms to help build brand irresistibility scores. First, however, we have to make sure that our prosperity scales accurately and consistently measures each aspect or construct of brand irresistibility. 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 adoration for a brand, but in reality we might be measuring the "general likability of the brand."

To measure brand irresistibility, we created the I-Factor scale. There are three primary factors that the I-Factor scale measures or the three C's: Comprehend, Crave, and Craze. Each of these factors can be further broken down into two facets each: Experience and Knowledge are facets for Comprehend, Adoration and Addiction are the facets for Crave, and Badge and Buzz are the facets for Craze.

In this section, we walk you through the scientific process of how we evaluated the psychometric properties of the I-Factor Scale, using the Comprehend factor as an example.

FACTOR

I-Factor – Brand Irresistibility



We determined what makes a brand irresistible by measuring three C's: Comprehend, Crave, and Craze.

For the Comprehend construct:



Comprehend Experience

A high score indicates that a consumer receives a great experience when using or buying from a brand. Such consumers feel like their life is enhanced by the experience products or services from the brand provide.



A high score indicates that a consumer knows about and seeks out knowledge about a brand. Such consumers know so much about the brand that they feel like they could be a salesperson for the brand.

We determine the extent to which individuals Comprehend a brand by adding up scores on Comprehend Experience and Comprehend Knowledge. We repeat this process for the remaining brand irresistibility constructs. Together, the three C's form the I-Factor score. Brands can use this information to target specific constructs within the three C's to improve brand irresistibility and ultimately connect and retain their consumers more effectively.

I-Factor Part 1: Scale Validity

For the I-Factor Scale to work, we had to train and test how responses to the scale were related to scores on each of the constructs 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 the construct Comprehend truly assessing the extent to which a person considers a brand as a representation of themselves? 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 construct?



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



Convergent Validity Does the construct, Comprehend, relate to other constructs it should be theoretically related to?



Nomological Validity Does the network of constructs around the construct, Comprehend, show relationships that are expected?

Construct Validity: Structural Validity.

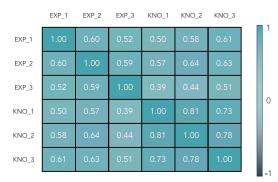
For the Comprehend construct, we want to make sure that the items for Comprehend Experience and items for Comprehend Knowledge are measuring their different facets of brand Craze. To do so, we assess structural validity by using both exploratory factor analysis and confirmatory factor analysis.

-FACTOR



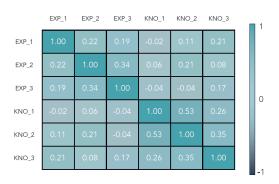
Exploratory Factor Analysis.

- 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 Comprehend Experience all have correlations, on average, above .60 with each other. Similarly, all three items Comprehend Knowledge have correlations above .75 with each other. Together, the items have correlations above .57 with each other. Therefore, all items for the Comprehend construct 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 vthis threshold. As you can see in the example below, the three items included in Comprehend Experience all have partial correlations below .7 with each other. Similarly, all three items in Comprehend Knowledge have partial correlations below</p>

.7 with each other. Together, all items have partial correlations below .7 with each other. Therefore, all items for the Comprehend construct were retained.



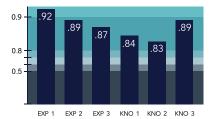
We also look at the anti-image correlation matrix, which contains the negatives of the partial correlation coefficients. Consequently, these values are the magnitude of the variable that can't be regressed on, or predicted by, the other variables. If variables can't be regressed on, or predicted by, 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. For the I-Factor Scale, correlations in the ant-image correlation matrix are close to 0. This means all items on both constructs are retained.

	EXP_1	EXP_2	EXP_3	KNO_1	KNO_2	KNO_3	
EXP_1	1.00	-0.11	-0.11	0.01	-0.04	-0.08	1
EXP_2	-0.11	1.00	-0.17	-0.02	-0.07	-0.03	
EXP_3	-0.11	-0.17	1.00	0.02	0.02	-0.07	0
KNO_1	0.01	-0.02	0.02	1.00	-0.15	-0.08	
KNO_2	-0.04	-0.07	0.02	-0.15	1.00	-0.10	
KNO_3	-0.08	-0.03	-0.07	-0.08	-0.10	1.00	-1

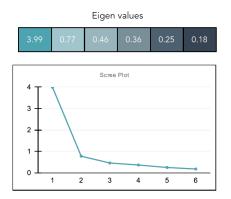
- Bartlet test of sphericity compares the » correlation matrix to the identity matrix, checking to see if there is any redundancy between the variables. High redundancy is indicative that the variables have common variance and therefore can be loaded on similar factors. If there is high redundancy, then 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 Comprehend construct, the correlation matrix was significantly different from the identity matrix.
 - P3
 P3<
- 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

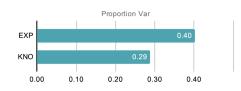
There is cause for concern, if the KMO drops below .60. All items for CComprehend Experience and Comprehend Knowledge have values above .80, indicating that they are meritorious fit.



- 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



» While the eigenvalues present evidence for a one-factor solution, in that all six items come together to represent the Comprehend construct, the proportion of variance for the two facets indicate that items measuring the Comprehend construct can also be further divided and represented with the Experience facet and Knowledge facet.



- 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
.5069	Practically significant
.3049	Minimally viable for a factor structure
< .30	Unrelated

You can see the following example:



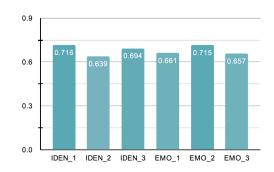
- Items for Comprehend Experience have factor loadings above .58, indicating a welldefined structure. Items for Comprehend Knowledge have factor loadings above 0.65, indicating practical significance. This means that items for both factors are at a minimum practically significant. Additionally, items did not cross-load across factors. Analytically, this means that items measuring Comprehend Experience didn't have factor loadings greater than .3 for Comprehend Knowledge and vice versa. Conceptually, items for Comprehend Experience and items for Comprehend Knowledge measured two key aspects of the Comprehend construct. Together, the correlation between the factors was .84, indicating that while items can be separated into facets, they still overlap and collectively measure the Comprehend construct.
- 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

- » With a sample size greater than 200, we can safely conclude that factor loadings for Comprehend Experience and Comprehend Knowledge 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).



» All items have communalities above .5 indicating that all items should be retained, as a considerable amount of variance is accounted for by the factor solution.

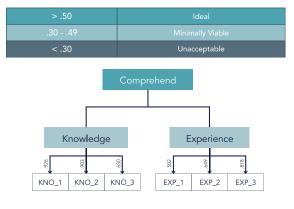
Confirmatory Factor Analysis.

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 Comprehend Experience and Comprehend Knowledge load highly on their respective factors.

- 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 Comprehend Experience and Comprehend Knowledge 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.
- 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, no modification individes were flagged indicating a well-defined structure. Thus, we retained our hypothesized model.



< .20	No problem
.2139	"Red flag"
> .40	Unacceptable

	KNO_1	KNO_2	KNO_3		EXP_1	EXP_2	EXP_3
KNO_1	0.000	0.036	-0.016	EXP_1	0.000	-0.033	0.019
KNO_2	0.036	0.000	-0.027	EXP_2	-0.033	0.000	0.025
KNO_3	-0.016	-0.027	0.000	EXP_3	0.019	0.025	0.000



Scale Validity > Construct Validity > Structural Validity



• 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
TU	> .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	0.980
TLI	0.963
RMSEA	0.096
SRMR	0.035

- » Notice that our model fits the data very well, with the exception of RMSEA. This is likely a function of the sample size. Since all other empirical evidence points to a good fit, we move forward.
- 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 Comprehend Experience and Comprehend Knowledge, the average variance extracted was greater than .50.



Construct Validity: Convergent & Divergent Validity.

Other forms of construct validity are 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.

When looking to support convergent and divergent 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
.7089	Convergent validity for highly related constructs
.5069	Convergent validity for somewhat related constructs
.4049	No man's land
.2039	Divergent validity for somewhat unrelated constructs
.1019	Divergent validity for highly unrelated constructs
009	Indicates no relationship

With Confirmatory Factor Analysis.

- » 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.

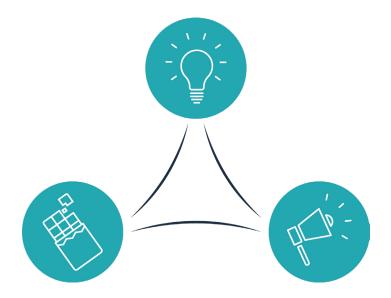
» Test a bifactor model and see if it gets worse. A bifactor model is usually used when you want to test the presence of a general factor that all items load onto. This approach helps identify the plausibility of a scale having multiple factors that are theoretically uncorrelated.

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:



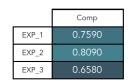
- The higher a consumer scores on the Craze, the higher they score on the Crave.
- The higher a consumer scores on the Craze, the higher they score on the Comprehension.
- The higher a consumer scores on the Comprehension, the higher they score on the Crave.

These relationships between constructs make sense. When a consumer feels a high degree of "Craze" or willingness to engage and share a brand, the more likely they are to be connected to the brand or "Crave" the brand. Similarly, when a consumer feels a high degree of "Craze" towards a brand, the more likely they are to know a lot about a brand and its product, or "Comprehend" the brand. Finally, when a consumer feels a high degree of "Crave" towards a brand. While district, all of these factors are interrelated to create how irresistible a brand is to a consumer.

I-Factor Part 2: Scale Reliability



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 Comprehend Experience, we get the following:



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

	Comp
KNO_1	0.8500
KNO_2	0.8980
KNO_3	0.8880

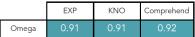
Again all items are above the .50 threshold.

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
.6069	Acceptable

As you can see below, not only do the facets Comprehend Experience and Comprehend Knowledge meet the .70 threshold for reliability, but also the general Comprehend construct meets the reliability threshold, suggesting good reliability.



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			
α ≥ 0.9			Excellent			
$0.9 \ge \alpha \ge 0.8$						
$0.8 \ge \alpha \ge 0.7$			Acceptable			
$0.7 \ge \alpha \ge 0.6$			Questionable			
$0.6 \ge \alpha \ge 0.5$			Poor			
0.5 > α			Unacceptable			
	EXP			KNO	Comprehend	
	Alpha	0.80		0.81	0.90	

When looking at Comprehend Experience and Comprehend Knowledge, scale reliability is .80, indicating good internal consistency. For the general Comprehend construct scale reliability is .92 indicating excellent internal consistency.

-FACTOR

Summary



At this point, I hope you can see just how much rigor goes into the platform, I-Factor, and the Brand Irresistibility 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--Comprehend Experience and Comprehend Knolwedge--have good to excellent psychometric properties.

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



Notice

The entire contents of this presentation are owned by or licensed to Inkblot Holdings, LLC and cannot be reproduced, communicated or distributed without the express permission of Inkblot Holdings, LLC. Further, the information disclosed herein may contain, or relate to, one or more inventions which are the subject of U.S. and/or foreign patent application(s), and/or, are the subject of trade secret protection. The contents of this presentation are intended only for use to evaluate a potential business relationship with Inkblot Holdings, LLC and cannot be used for any other purpose. Inkblot Holdings, LLC reserves the right to pursue any legal remedies for unauthorized use, communication or distribution of this presentation or its contents, in order to preserve or enforce its rights.

Copyright © 2023 Inkblot Holdings, LLC