



**Breckenridge Institute**  
HARNESSING THE POWER OF CULTURE™

## **Reliability and Validity Report**

### **Sikora Strategy Preference Indicator™ SikoraSPI™**

*Prepared For:*

**Breckenridge Institute**

PO Box 5050  
Breckenridge, CO 80424  
1-800-303-2554

[info@breckenridgeinstitute.com](mailto:info@breckenridgeinstitute.com)  
[www.breckenridgeinstitute.com](http://www.breckenridgeinstitute.com)

*By:*

***Mark S. Majors Ph.D.***

**February 2007**

# TABLE OF CONTENTS

Overview.....	3
Development of the SikoraSPI™.....	4
Data Collection Procedure.....	5
Psychometric Properties of the SikoraSPI™.....	6
Scale Reliability (Internal Consistency).....	6
Consistency Over Time (Test-Retest Reliability).....	7
Validity of the SikoraSPI™.....	9
Means Analysis for the Nine Enneagram Types.....	9
Means Analysis for the Three Instinct Types.....	10
Conclusion.....	12
Mark Major's Bio.....	13

## Overview

The Sikora Strategy Preference Indicator™ (SikoraSPI™) is a 140-item Likert format instrument based on a model of personality style called the Enneagram. The process of developing the SikoraSPI™ was initiated by Breckenridge Institute™ in September 2006. The instrument was developed by Mario Sikora, Mark Bodnarczuk, and Elin Larson. The original question set contained 195 items, and currently the assessment contains 150 items (140 plus 10 research items). The results of the instrument provide scores on the nine Enneagram Strategy scales as well as three Instinct scales.

*The Strategies* are nine distinct patterns of interacting with our world. They are the way we go about interacting with others and satisfying the needs and desires of daily life. The strategies are at the heart of what is generally thought of as “personality type.” *The Instincts* are three distinct areas of fundamental biological drives. They determine what is important to us and what activities we tend to direct our attention to during the course of the day.

The remainder of this report will provide a brief discussion of the developmental process and the psychometric characteristics of the 12 scales that compose the SikoraSPI™.

## Development of the SikoraSPI™

The SikoraSPI™ project began in the fall of 2006 with the development of the 195-items and decision to use a six-response Likert-scale format. The developmental process focused upon data collection to refine and reduce the number of items, develop an accurate scoring method, and establish the reliability and validity of the measure. The final SikoraSPI™ research version contained 199 items for measuring the nine types and three instinct dimensions and four questions for scale validation. The four validity questions that were added to the end of the measure are presented below.

- If you know your Enneagram type, please select the appropriate choice
- How certain are you of your Enneagram type?
- If you know your dominant sub-type (instinctual variant) please select the appropriate choice
- How certain are you of your sub-type?

Currently the assessment contains 150 items (140 plus 10 research items). The results of the instrument provide scores on the nine Enneagram Strategy scales as well as three Instinct scales.

## Data Collection Procedure

Data collection for the SikoraSPI™ was done through email solicitation and web-based administration. This method provided consistency (all contact was uniform) throughout the collection process. Individuals were provided information about participating in the research of a new measure of personality type. By clicking on an Internet link the individual was taken to a web site to read instructions and complete the measure. All contacted participants were over 18 years of age and there were no known risks to participants.

In January of 2007 the initial data collection was completed (N=388). The sample was composed of 186 females and 202 males from a wide variety of business and college settings. To establish test retest reliability a second sample was collected after a 30-day interval (N = 205). The same solicitation and collection process was used as with the initial sample, and all of the 388 individuals who completed the SPI in the initial sample were contacted. The retest sample contained 105 females and 100 males.

The first phase of the data analysis focused upon the development of the 12 scales. An iterative process of statistical evaluation and scale validity resulted in reducing the number of items to the final 140-item format. The scale scoring is based upon individual item weighting and a normative process established by balanced sampling (balanced by type). The psychometric properties of the final nine type scales and three instinct scales are presented below.

## Psychometric Properties of the SikoraSPI™

The psychometric properties of the SikoraSPI™ that will be presented are the results of reliability and validity analysis. The reliability that is reported focuses on individual scale internal consistency reliability and stability over time (test-retest reliability). The validity of a multi-scale measure like the SikoraSPI™ is indicated by the validity of the individual scales that it contains. The ability of each of the nine type scales to differentiate its intended group (type) from other groups is established, as well as the scale's accuracy in placing individuals who know their type on the correct type scale. The internal consistency scores for the SikoraSPI™ are presented below in the form of Chronbach's Alpha Coefficient.

### Scale Reliability (Internal Consistency)

Test reliability refers to the extent to which the SikoraSPI™ scores can be expected to reliably perform each time the measure is used. There are two generally recognized measures of reliability. The first is internal consistency which is an estimate of how closely the responses to items in the SikoraSPI™ on a given scale relate to each other. It is generally agreed that an alpha coefficient of .70 or higher indicates adequate internal consistency.

The SikoraSPI™ instrument contains 12 scales with a total of 140 items. Some items appear on more than one scale, but the weighting of those items is unique for each scale. To get a sense of what is the common range of internal consistency reliability (Cronbach's Alpha Coefficient) we can examine the reported alphas from other common measures such as the MBTI® Form Q and M. The scales on the MBTI® Form Q range from 5 to 9 items on each scale. The average Chronbach's Alpha reliability for these scales is .77. The scales on MBTI® form M range from 21 to 28. The average Chronbach's Alpha reliability for the four scales is .92. These are common ranges of scale alphas from very well constructed measures. Below are the internal consistency Chronbach's Alphas for the SikoraSPI™ instrument.

SikoraSPI™ Type Scale	Alpha	# items
Type One: Perfect	.88	4
Type Two: Connected	.98	13
Type Three: Outstanding	.98	13
Type Four: Unique	.97	13
Type Five: Detached	.96	13
Type Six: Secure	.89	17
Type Seven: Excited	.87	17
Type Eight: Powerful	.88	12
Type Nine: Peaceful	.97	13
Overall Average: (Sample size N=338)	.93	

The range of number of items per scale on the Type scales of the SikoraSPI™ is 4 to 17 items per scale with the average internal consistency reliability Alpha = .93. This indicates a very high level of internal consistency. Below are the reliability alphas for the three Instinct scales.

Sikora SPI Instincts	Alpha	# items
Preserving	.94	8
Navigating	.91	10
Transmitting	.94	9
Overall Average:	.93	
(Sample size N=338)		

The SikoraSPI™ Instinct scales have an average internal consistency of Alpha = .93, which indicates a very strong connection statistically within the items on each of the three scales. The results of the internal consistency analysis for the SikoraSPI™ show that the measure has been constructed with items that are of similar meaning and utility.

### **Consistency over Time (Test-Retest Reliability)**

The second measure of reliability to be reviewed is the measure of reliability over time or test-retest reliability. The statistic that is used to express this measure of reliability is the Pearson product moment correlation. To produce this analysis a group of individuals are sampled on the same measure two times over an interval of time (30 days). The results of the SikoraSPI™ test retest analysis for all 12 scales are presented in the tables below.

Sikora SPI Type Scale	Pearson <i>r</i>
Type One: Perfect	.87
Type Two: Connected	.90
Type Three: Outstanding	.85
Type Four: Unique	.87
Type Five: Detached	.89
Type Six: Secure	.89
Type Seven: Excited	.93
Type Eight: Powerful	.91
Type Nine: Peaceful	.83
Overall Average:	.88
(Sample size N=199)	

SikoraSPI Instincts™	Pearson <i>r</i>
Preserving	.71
Navigating	.84
Transmitting	.86
Overall Average:	.80
(Sample size N=199)	

The results of test retest analysis indicate that SikoraSPI™ is very stable over time. The average correlation for the nine Enneagram scales is .88, and for the three Instinct scales it is .80. These results are above average for measures with these scale lengths.

The information presented on the SikoraSPI™ reliability shows that the measure has both internal consistency and stability over time. This is important information for the assurance that this measure will produce consistent results across administrations and internal stability for reliable results. Of equal importance is the indication of an instrument's validity.

## Validity of the SikoraSPI™

Test validity involves determining the extent to which the SikoraSPI™ actually measures the characteristics that it claims to measure. The SikoraSPI™ was designed to help in the process of individuals discovering their Enneagram Type and their Instinct Type. This requires that the instrument be able to accurately sort an individual into the correct type (one out of nine). There are no sorting instruments that can perform this task perfectly, as there are no universally accurate questions that function well for everyone. Validity is the information that reveals a scale or instrument's ability to do what it is intended to do. The accuracy of the SikoraSPI™ has been established by scale means analysis and overlap.

### Means Analysis for the Nine Enneagram Types

One way to indicate the ability of a scale to separate relevant groups or assign individuals into the correct group is through a simple means analysis. Below is a table representing the means and standard deviations of each of the nine Type scales across groups who have reported, with confidence, that they know their correct Enneagram Type.

		SCALE								
Known Type		Type1	Type2	Type3	Type4	Type5	Type6	Type7	Type8	Type9
1	Mean	<b>62.5</b>	47.2	52.2	47.1	54.0	55.4	46.0	52.1	49.4
	N	31	31	31	31	31	31	31	31	31
	Std. Dev.	5.7	9.2	8.0	9.6	8.6	7.2	8.6	6.4	8.8
2	Mean	50.5	<b>62.6</b>	48.3	48.7	39.4	48.0	53.0	48.2	49.1
	N	19	19	19	19	19	19	19	19	19
	Std. Dev.	8.2	5.9	5.9	8.7	7.4	11.3	7.2	7.7	8.9
3	Mean	44.7	50.1	<b>62.9</b>	49.6	47.6	48.0	54.0	53.8	45.5
	N	25	25	25	25	25	25	25	25	25
	Std. Dev.	10.0	9.4	7.6	6.3	9.1	11.1	6.6	6.5	4.6
4	Mean	47.5	52.7	49.1	<b>65.5</b>	47.5	44.0	46.4	46.2	47.1
	N	23	23	23	23	23	23	23	23	23
	Std. Dev.	8.9	9.4	9.4	7.5	10.1	9.6	8.4	11.0	8.5
5	Mean	49.6	40.5	46.4	49.2	<b>64.1</b>	51.0	42.5	48.1	49.1
	N	17	17	17	17	17	17	17	17	17
	Std. Dev.	10.6	9.1	7.6	8.6	3.5	6.9	7.7	8.4	6.8
6	Mean	48.4	49.2	41.5	44.7	51.5	<b>61.0</b>	45.7	45.6	46.7
	N	18	18	18	18	18	18	18	18	18
	Std. Dev.	8.2	8.0	8.3	7.8	7.2	6.2	10.8	8.2	7.5
7	Mean	44.7	51.5	48.3	53.1	45.1	45.9	<b>63.4</b>	51.3	49.8
	N	27	27	27	27	27	27	27	27	27
	Std. Dev.	7.9	9.2	6.3	7.8	9.2	10.8	5.6	7.0	8.9
8	Mean	51.3	44.4	55.8	45.9	50.3	49.2	51.4	<b>63.9</b>	43.2
	N	21	21	21	21	21	21	21	21	21
	Std. Dev.	8.3	10.8	10.2	10.8	8.6	8.6	9.5	6.4	6.6
9	Mean	48.4	51.0	44.3	46.5	51.5	49.1	46.1	42.7	<b>62.5</b>
	N	35	35	35	35	35	35	35	35	35
	Std. Dev.	8.8	6.8	8.8	6.9	8.1	8.0	7.9	10.8	10.3
<b>Total</b>	N	216	216	216	216	216	216	216	216	216
<b>d %</b>	<b>Overlap</b>	<b>13.4%</b>	<b>13.4%</b>	<b>11.3%</b>	<b>8.6%</b>	<b>11.3%</b>	<b>17.3%</b>	<b>11.3%</b>	<b>10.3%</b>	<b>9.4%</b>

The above table was produced from those individuals who reported that they know their Enneagram Type (question presented previously). The means for the scale on the corresponding type group is presented in bold. All of the scales are based upon norms developed from the known type group (sample N = 216) and are standard t-scores with a mean of 50 and standard deviation of 10. Examination of the information in the table reveals that all of the nine type scales separate individuals of their intended type groups extremely well. The Cohen's *d* effect size is used to describe how well a scale separates its intended group from others. The overlap percentages (*d* %; the percentage of overlap based on Cohen's *d*) presented at the bottom of the table represent the amount of scale score overlap of the Type group, on their Type scale, with the general population (the sample with the specific type group removed). These scales provide a clear delineation of the intended target type from other types on all scales.

What does the means analysis indicate? Because accurate information was collected from a substantial number of individuals (N = 216) who report that they clearly know their Enneagram Type, the means results in the above table are a powerful proof that the SikoraSPI™ is doing what it is intended to do with respect to the nine Type scales. Very few individuals will score with a false high on any of the nine scales. Multiple high results (above t-score ≥ 62) on an individual's result profile are an accurate indication of a much stronger than average utilization of a Type strategy. Further the low percentage of overlap from the Cohen's *d* analysis reveals the measure's strong Hit Rate (a measure's ability to accurately classify or group an individual). The average overlap across all nine type scales is 11.8% or the average Hit Rate is 88.2% for the SikoraSPI™. Very few measures (and no other Enneagram Type measure) will provide this level of scale clarity.

### Means Analysis for the Three Instinct Types

The table below provides the same mean and Cohen's *d* percentage information for the three instinct scales as was presented above for the nine type scales.

		Instinct Scales		
Known Instinct		Preserving	Navigating	Transmitting
<b>Preserving</b>	Mean	<b>54.8</b>	45.1	45.4
	N	34	34	34
	Std. Dev.	8.0	12.3	10.7
<b>Navigating</b>	Mean	40.0	<b>57.0</b>	46.4
	N	21	21	21
	Std. Dev.	13.9	6.0	14.8
<b>Transmitting</b>	Mean	44.3	47.0	<b>55.9</b>
	N	26	26	26
	Std. Dev.	10.8	13.3	5.6
<b>Total</b>	N	81	81	81
<b>d %</b>	<b>Overlap</b>	<b>18.9%</b>	<b>20.5%</b>	<b>22%</b>

The results of means analysis on the three instinct scales show that all three function very well at separating individuals into the reported groups. The average amount of overlap

between the instinct scale's intended groups with the other groups is 20.5%. In other words almost 80% of the scores for the intended groups on the instinct scales are above the non-intended groups. This can be considered an 80% hit rate for these scales. This is consistent with other high quality measures such as the MBTI® form M and others.

The validity results for the Sikora SPIT™ indicate that the instrument is measuring what it is purported to measure. Self-reported known types (Enneagram and instinct type), provided by individuals who indicate high levels of certainty for their types, make it possible to establish clear levels of validity. Sampling individuals from workshops or individuals who have little opportunity to become firm in their individual type belief causes measurement error and the need for the use of Kappa inter-rater reliabilities to attempt to establish validity. This compounds error on error, as the inter-rater agreement and the initial beliefs of newly experienced individuals are known to be frequently flawed.

## Conclusion

The results of the psychometric analysis performed on the SikoraSPI™ reveal a psychological instrument that is both reliable and valid. No measure is perfect and there are no perfect items, but this measure of Enneagram type is performing extremely well and can be considered a valuable tool in helping individuals to gain important information about their personality with a minimum of error and confusion.

## Dr. Mark S. Majors' Bio

Dr. Mark S. Majors is a counseling psychologist with extensive psychometric experience that includes data analysis on the 1994 Strong Interest Inventory and the MBTI® Form M. He was coauthor for the new MBTI® Form Q Manual, as well as the developer of the IRT scoring. Mark has been a Research Scientist at Consulting Psychologist Press (CPP) and Director of Research at the Center for the Application of Personality Type (CAPT). He is widely published in the areas of psychometrics, personality type, and counseling psychology and is the developer of the Majors Personality Type Inventory™ (MajorsPTI™), Occupational Environment Measure™ (MajorsOEM™), and codeveloper of the Interstrength® X-Styles Assessment. In addition to being a counselor and test developer, he is president of a small private college. The college specializes in training pastoral counselors with an emphasis on the use of personality assessment for conflict resolution through the acceptance of differences and personal growth. Mark received his BS and MS degrees in psychology from Iowa State University, and his PhD in counseling psychology from the University of Nebraska-Lincoln.