Facilitator

Technical Report

Universal

Multi-Rater SOCIAL STYLE & Enhanced Versatility Profile

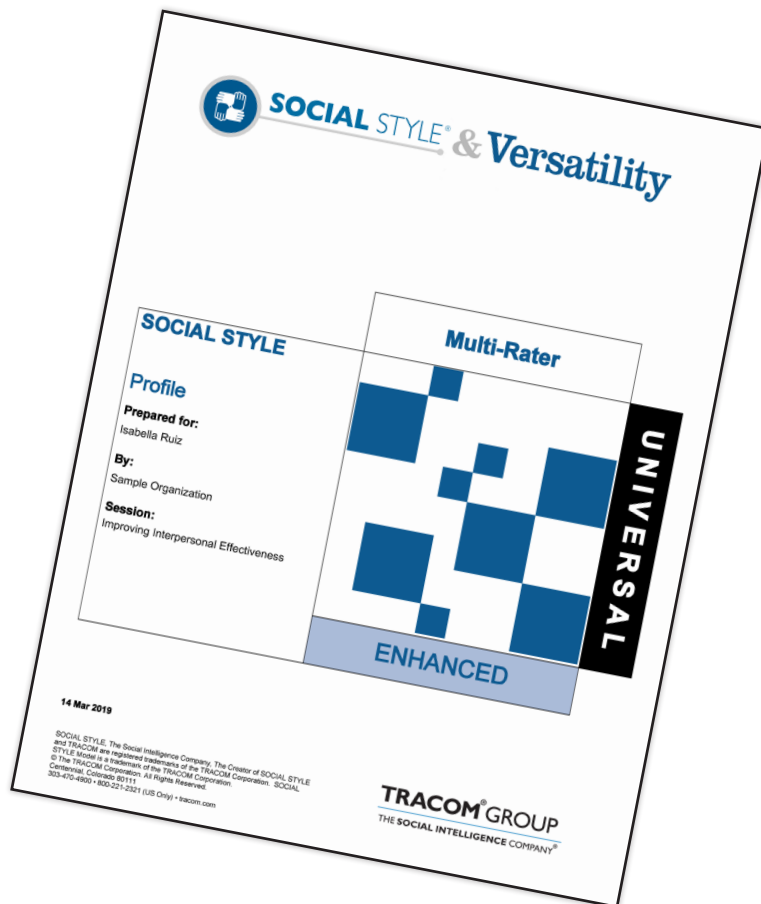**SOCIAL STYLE**® **& Versatility**

SOCIAL STYLE

Multi-Rater

Profile

Prepared for:
Isabella Ruiz

By:
Sample Organization

Session:
Improving Interpersonal Effectiveness

UNIVERSAL

ENHANCED

14 Mar 2019

SOCIAL STYLE, The Social Intelligence Company, The Creator of SOCIAL STYLE and TRACOM are registered trademarks of the TRACOM Corporation. SOCIAL STYLE Model is a trademark of the TRACOM Corporation.
© The TRACOM Corporation. All Rights Reserved.
Centennial, Colorado 80111
303-470-4900 • 800-221-2321 (US Only) • tracom.com

TRACOM® **GROUP**
THE SOCIAL INTELLIGENCE COMPANY®

TRACOM® **GROUP**
THE SOCIAL INTELLIGENCE COMPANY®

Casey Mulqueen, Ph.D.

Facilitator

Technical Report

Universal

Multi-Rater SOCIAL STYLE & Enhanced Versatility Profile

SOCIAL STYLE® & Versatility

SOCIAL STYLE

Profile

Prepared for:
Isabella Ruiz

By:
Sample Organization

Session:
Improving Interpersonal Effectiveness

Multi-Rater

UNIVERSAL

ENHANCED

14 Mar 2019

TRACOM GROUP
THE SOCIAL INTELLIGENCE COMPANY®

TRACOM® GROUP
THE SOCIAL INTELLIGENCE COMPANY®

Casey Mulqueen, Ph.D.

Contents

Acknowledgements

Introduction

Throughout our history, TRACOM has maintained a focus on research and development. We take pride in developing products that are grounded in meaningful and useful research. The SOCIAL STYLE Model™ was originally created through empirical research on how people behave at work. Today, we regularly conduct research on the Model and its applicability to people's work lives.

This Technical Report provides the most up-to-date research evidence for the reliability and validity of the SOCIAL STYLE® Profile – Enhanced (SSP-E) questionnaire. The SSP-E questionnaire is our flagship assessment instrument, and is used as the input to generate various profiles, including the Universal Multi-rater Profile, Managerial and Sales Profiles, and the Online Self-Perception Profile.

This report is intended to be a companion to other facilitator materials, notably the *SOCIAL STYLE & Versatility Facilitator Handbook*. By reviewing this report you will understand the important role that research plays in the SOCIAL STYLE Model. More importantly, you will gain an understanding of exactly how TRACOM insures that our profiles are accurate and dependable. This will help you to have confidence in our products and services, and also to explain these concepts to the people you teach.

We begin by discussing the history of the SOCIAL STYLE Profile and Model, along with the development of the SSP-E questionnaire. We then describe the research evidence for the reliability and validity of the questionnaire, followed by norm descriptions. Finally, we describe each of the TRACOM profiles that are based on the SSP-E questionnaire. For simplicity we refer to the questionnaire simply as the "SSP-E."

Glossary

This report is written to be understandable for people who will be facilitating and using SOCIAL STYLE programs. There are some key technical terms that are used frequently throughout the report, and it is helpful to define these upfront. Further clarification of these terms is given in the body of the report.

- Reliability – determines whether the SSP-E is consistent and dependable.

- Validity - determines whether the SSP-E measures accurately. In other words, does it truly measure the concepts (Style and Versatility) that it proposes to measure?

- Correlation - a correlation coefficient determines the extent to which two variables are related to each other. Values range from 0.0 (no relationship) to 1.0 (perfect relationship). For example, height and weight are proportional to one another and should be highly correlated. In fact, the correlation between height and weight among adults is 0.44, a strong relationship (Meyer et al., 2001).

- Item – an item is a behavioral statement on the survey, sometimes called a "survey question." An example of an item is "Interrupts others when they are talking."

- Scale – a scale is a collection of survey items that measures a single construct. For example, Assertiveness is a scale. It is measured by a group of items that are all related to the Assertiveness construct.

- Profile – a profile is the actual report that is given to participants. It includes a profile of SOCIAL STYLE and a profile of Versatility.

- Norms - Normative scores, or norms, are necessary for interpreting scores on scales. Norms provide context to an individual's scores by comparing them to a meaningful group (e.g., country, job group). Based on this, norms make it possible for us to determine a person's Style and Versatility positions.

History and Development

Since its inception in the 1960s, the SOCIAL STYLE Model has been applied to a variety of situations, including team effectiveness, sales training, Six Sigma training, and many others. Currently, TRACOM utilizes the SOCIAL STYLE Model for specific applications, such as managerial and sales effectiveness as well as the more universal SOCIAL STYLE products.

This section reviews the history of the SOCIAL STYLE Model and more modern development of the SSP-E.

The SOCIAL STYLE Profile

In 1964, Dr. David W. Merrill, an organizational psychologist and founder of Reed, Merrill, Brunson and Associates (RMBA), and Personnel Predictions and Research, Inc., (PPR), undertook research to explore ways to predict success in selling and management careers. (RMBA and PPR were joined to form TRACOM in 1978). He understood that people tend to behave in consistent ways that others can observe. He sought to find a method for measuring these behavioral observations through the use of descriptive adjectives. Using a technique that was unique for that time, Dr. Merrill measured behavior using a multi-rater approach, believing that people can agree about the behavior of a person they know. He utilized an empirical approach, meaning that the research was not designed to support any specific theory of behavior, and could be tested by other researchers.

TRACOM's original Adjective Checklist, the precursor to the SSP-E, was developed from an initial pool of more than 2,300 words. Work on this larger checklist had been done in the early 1960s by Dr. James W. Taylor, who at that time was staff psychologist at Martin Corporation (later Martin Marietta) in Denver.

Dr. Merrill obtained rights to use the checklist, and enlisted the participation of a major life insurance company that provided a pool of 600 people to study. These individuals had their co-workers complete an adjective checklist on them. The checklist was completed by answering "yes," "no" or "don't know" to whether or not the adjectives described the person being observed. Statistical analysis found that if a respondent felt a certain adjective described the individual's behavior, that same respondent would answer "yes" or "no" to certain other adjectives. In other words, some adjectives clustered together.

A statistical procedure called factor analysis was conducted on this data. Hundreds of adjectives were compared to each other to see which words clustered together. Adjectives that clustered together were considered to measure a dimension of human behavior.

A total of 150 adjectives measuring three scales were finalized. The three scales were labeled Assertiveness, Responsiveness, and Versatility.

Assertiveness: The way in which a person tries to influence others. Or in other words, it is the degree to which individuals tend to "ask" or "tell" in interactions with others.

Responsiveness: The way in which a person outwardly displays feelings and emotion. It is a measure of the degree to which a person tends to "control" or "emote" when interacting with others.

Versatility: A type of support and respect given to a person by others. Versatility is based, in part, on the extent to which others see the individual as interpersonally effective. It can be thought of as the extent to which a person appears to be working to make relationships mutually productive.

The scales that were discovered during this early research were used to develop the SOCIAL STYLE Model. By combining the two dimensions

of Assertiveness and Responsiveness, four patterns of behavior, or Styles, could be identified. The four Styles are:

- Driving (Tell Assertive + Control Responsive). These individuals are seen as strong-willed and more emotionally controlled.

- Expressive (Tell Assertive + Emote Responsive). These individuals are described as outgoing and more dramatic.

- Amiable (Ask Assertive + Emote Responsive). These individuals are seen as easy-going and supportive.

- Analytical (Ask Assertive + Control Responsive). These individuals are described as serious and more exacting.
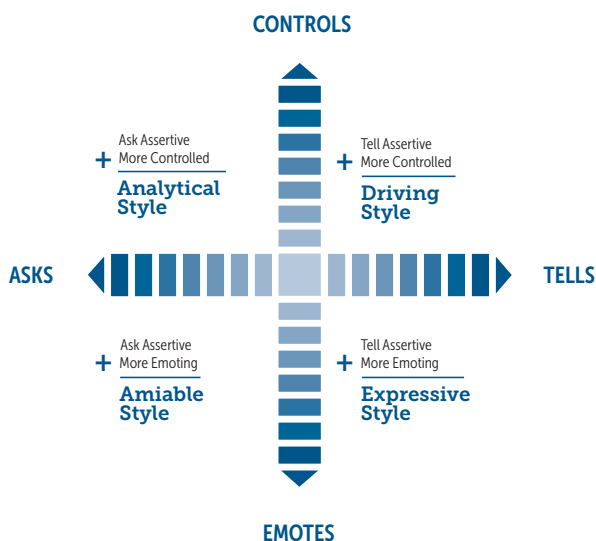
Versatility is an independent scale, reported separately from SOCIAL STYLE.

The original Adjective Checklist has been administered to more than one million people over the course of the past four decades. Further information about the development of the SOCIAL STYLE Model and the Adjective Checklist can be found in Merrill and Reid's book, *Personal Styles and Effective Performance*, and in *SOCIAL STYLE and Versatility Facilitator Handbook*.

Development of the SOCIAL STYLE Profile – Enhanced

The SOCIAL STYLE Profile – Enhanced (SSP-E) was developed from 2001 through 2003. Three primary factors led to the decision to revise and expand the original profile instrument.

First, the original questionnaire is a list of adjectives that people use to describe themselves and others. These adjectives resulted from empirical research conducted in the early 1960s. The growing multiculturalism of American society and natural language evolution has made some of the original adjectives less frequently used in today's common language. In addition, the popular meanings of some of the adjectives have changed over the years. Since responses to the adjective checklist rely on a certain level of common vocabulary, without the use of a dictionary, newer generations of people may be challenged by some of the items on the original questionnaire.

Though the research that TRACOM has conducted over the years has never shown any indication that the validity of the Profile was affected by these events, there was some concern that a growing number of respondents were having difficulty with some of the adjectives on the questionnaire. For example, one of the adjectives, "religious," was at one time a relatively common way to describe someone who is conscientious and dutiful. However, some people now automatically interpret the word in its theological sense. Thus, the vocabulary level of the checklist was an issue.

Second, research in the area of emotional intelligence had led to a desire for updated research and expansion of the concept of Versatility. Psychologists have developed and expanded the theoretical framework of emotional intelligence for many decades (Gardner, 1983; McClelland, 1973; Sternberg, 1996). However, the use of the term "emotional intelligence" has only become popular since the mid 1990s (Salovey & Mayer, 1990; Goleman, 1995). Recent publications (Bar-On, 2002; Cherniss & Goleman, 2002; Goleman, 1998; Goleman, 2006; Goleman, McKee, & Boyatzis, 2002) have illustrated that TRACOM's concept of Versatility, originally developed in the 1960s, precedes and parallels many of the concepts of emotional intelligence.

For example, one central aspect of the SOCIAL STYLE Model is to "Know Yourself, Control Yourself, Know Others, and Do Something For Others." These correspond very closely to the four dimensions of emotional intelligence outlined by Goleman and his colleagues: Self Awareness, Self Management, Social Awareness, and Relationship Management. In addition, several of the competencies that compose emotional intelligence are very similar to those that define TRACOM's Versatility concept, such as Conscientiousness, Empathy, and Optimism.

Of the three constructs measured by the Model – Assertiveness, Responsiveness, and Versatility – Versatility is the most unfixed and changeable. Whereas Assertiveness and Responsiveness tend to be more consistent aspects of Style, Versatility can change across time and circumstances. Of the three constructs, Versatility is the most amenable to training and development, and the one that is most important for working effectively with others.

Because Versatility is such an important and trainable concept, there was a desire to expand upon this dimension by measuring its more

specific components. Thus, a central aspect of the SSP-E is its emphasis on measuring the four components of Versatility – Image, Presentation, Competence, and Feedback.

Finally, the third main reason for updating the measurement system was to allow the instrument to be more easily translated into other languages. When translating single adjectives, the original meaning of the words can be lost, affecting the validity of the profile. This is less of an issue when utilizing behavioral statements. In addition, during the translation process the statements are easily edited to insure their meaning remains stable across cultures.

Measurement Format

Unlike the original questionnaire, the SSP-E utilizes behavioral statements. These statements are responded to on an agreement continuum (a five-point scale ranging from "Strongly disagree" to "Strongly agree").

The decision to modify the format of the new instrument was based on several factors. First, behavioral statements are easier to respond to than descriptive adjectives. The original questionnaire is based on a behavioral model. That is, a person's Style is described in behavioral terms, and the adjectives were chosen because people could use them as descriptors of the behaviors of the people they were rating. However, most people find it easier to describe others based on actual behavior ("Easily adapts to unfamiliar situations"), rather than single words ("Changeable"). Behavioral statements have more inherent meaning than single adjectives when describing individuals.

Second, the original rating scale was essentially dichotomous. People responded either "yes" or "no", with a third "don't know" category. Respondents were discouraged from using the "don't know" category, and overuse of this

category resulted in an unscorable profile. A five-point rating scale was adopted because it is often difficult to describe people's behavior in such absolute terms as "yes" or "no". Behavior is a continuum, thus the continuous rating scale is ideal for allowing people to describe behavior.

Third, in order to measure more specific aspects of the Versatility dimension, a more explicit behavioral approach was necessary. Developing an expanded measurement model using the adjective checklist framework would have been psychometrically challenging; the behavioral statement model is more useful. By adopting the new model, we are able to measure and report important components of Versatility. Training program participants can now receive an enhanced report that describes specific components of Versatility (Image, Presentation, Competence, and Feedback). This more detailed level of feedback is instrumental in helping participants to identify the specific behaviors that they should focus on in order to leverage their overall effectiveness and impact on others.

Item Development

A deductive approach was used to develop items for the SSP-E questionnaire. The accumulation of research over the past several decades has resulted in a high level of understanding of the behaviors underlying the SOCIAL STYLE Model. This knowledge was utilized to develop the specific items for the questionnaire.

Descriptions of assertive and responsive behavior were used to define these concepts. Items corresponding to the range of assertive and responsive behaviors were then written. Likewise, a similar procedure was followed to develop Versatility items. However, the Versatility measure was expanded to include subdimensions of each of the four Versatility components.

The development of Versatility items relied not only on previous research on this concept, but also on research within the area of emotional intelligence. A comprehensive review of the emotional intelligence literature was conducted, and several key concepts were identified as useful expansions of the Versatility measure. These concepts were clearly defined in relation to the Versatility components, and items were written to measure them.

When reviewing the emotional intelligence literature, several criteria were established for selection of constructs that would correspond with and complement the Versatility dimensions. It was determined that each construct should:

- Have empirical support for its measurability.
- Have theoretical and empirical support for its relationship with interpersonal interactions and job performance.
- Theoretically fit into the existing conceptualization of Versatility.

Emphasis was put on adopting only those constructs that are most important for interpersonal skill and success within the workplace. The emotional intelligence framework is concerned with multiple facets of people's lives. While participants in TRACOM training programs often remark that the SOCIAL STYLE Model is relevant beyond the workplace, we wanted to maintain our emphasis on productive relations and functioning within the work environment. Figure 1 shows the breakdown of subdimensions that are measured under each of the four Versatility components. These subdimensions include both the pre-existing Versatility constructs as well as the newly adopted emotional intelligence constructs.

Image	Presentation	Competence	Feedback
• Dress and Grooming	• Effectiveness of Group Communication	• Conscientiousness • Flexibility • Innovation • Perseverance • Optimism	• Active Listening • Adaptive Communication • Empathy • Interpersonal Relations

Figure 1. Versatility Constructs Measured by the SSP-E Questionnaire.

The next section discusses the validity and reliability of the SSP-E.

Reliability and Validity

The SSP-E measures behavioral style. But behavior, like all psychological phenomena, is different from things that can be easily and accurately measured in the physical world, such as weight and height. So how do we know that we are accurately measuring behavior? In order to make this claim, the instrument has to adhere to standards that have been set forth by the scientific community. In particular, research evidence should correspond to criteria set forth in the *"Standards for Educational and Psychological Testing"* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), which provides benchmarks for developing psychological measurement instruments. This evidence comes in two primary forms: reliability and validity.

Reliability determines whether an instrument measures in a way that is *consistent* and *dependable.*

For example, imagine that you have a brick, and stamped on the side of the brick is the phrase "50 pounds." The brick feels light to you so you decide to verify its weight by weighing it yourself. You put it on a scale and it registers that it weighs 50 pounds. Just to be extra sure, you weigh it every day for a week, using the same scale, and every day the brick registers as weighing 50 pounds. Although by now you can be reasonably sure of the brick's weight, you wonder if maybe your scale is faulty. So you weigh the brick on 10 different scales, and every scale tells you that the brick weighs 50 pounds. You can now be sure that the brick weighs 50 pounds, and your measurement is reliable.

The example of the brick points out two unique but similar aspects of a reliable measurement system. First, is one of the measures dependable? The first scale was dependable because every day it indicated that the brick weighed 50 pounds.

We can be confident that the first scale is a dependable measure of weight. Second, if using more than one measure, are these measures consistent with one another? By using multiple scales, we showed that they were consistent in their measures; they all indicated that the brick weighed 50 pounds. Therefore we can be confident that all of these scales are consistently agreeing with one another about the weight of objects. You will see how these concepts relate to the measurement of SOCIAL STYLE soon, but first let's discuss the importance of validity.

Validity determines whether an instrument measures *accurately.* In other words, does it measure what it proposes to measure?

Let's return to the example of the brick. If, after determining that your weight scales were reliable using the procedure above, you placed the same brick on a brand new scale and it told you that the brick weighed 30 pounds, you would not be able to place faith in this new scale. In fact, you would conclude that this new scale does not measure "weight" accurately at all. It is giving you a measurement that is in pounds, but it is way off base in terms of its accuracy. Because of its unacceptable lack of accuracy, this new scale is not valid for its intended purpose of measuring weight.

If a psychological measure is to be used to make accurate inferences about a person's behavior, it must be both reliable and valid. Reliability is a prerequisite for validity, but is not sufficient by itself. An instrument can be highly reliable and still not be valid for a particular purpose. Crocker and Algina (1986, page 217) demonstrate the difference between reliability and validity with an analogy.

> Consider the analogy of a car's fuel gauge which systematically registers one-quarter higher than the actual level of fuel in the gas

tank. If repeated readings are taken under the same conditions, the gauge will yield consistent (reliable) measurements, but the inference about the amount of fuel in the tank is faulty.

This analogy underscores that determining the reliability of a test is an important first step, but not the only step, in determining the validity of a test.

No psychological measurement instrument is perfectly reliable or perfectly valid, since it is subject to various sources of error. Reliability and validity are a matter of degree, and it is more appropriate to ask how reliable/valid an instrument is, rather than "if the instrument is reliable/valid." Evidence for reliability and validity is accumulated over time.

Reliability

Several forms of reliability evidence are presented here; internal consistency, interrater reliability, interrater agreement, and retest reliability. These studies highlight that the SSP-E is consistent in the way it measures Style and Versatility, that different individuals agree with one another when using the SSP-E to rate a person, and that the SSP-E is consistent across time. Finally, we also discuss research on the difference between self-perception (participants) and the perceptions of others (their co-workers).

Internal Consistency

One of the most common and established methods for establishing the reliability evidence for instruments like the SSP-E is internal consistency. Internal consistency measures the relationship among survey items that claim to measure the same thing. This relationship is measured by a correlation coefficient.

Each scale on the SSP-E is measured using a set of behavioral statements, also referred to simply as items. If all items on a scale such as Assertiveness are truly measuring the same thing, then they should correlate with one another to a certain degree – they should be internally consistent.

The SSP-E survey scales were analyzed for internal consistency reliability using Cronbach's (1951) coefficient alpha. Alpha values range from 0.0 (no relationship among the scale items) to 1.0 (perfect internal consistency).

The appropriate value for alpha levels depends on the type of assessment, and there is much debate around this issue. A comprehensive review (Charter, 2003) found that personality scales have an average alpha value of .77 (the author did not review behavioral style scales). Alpha values that are too high can indicate items that are

redundant with one another, in which case they are not contributing uniquely to measurement of the behavioral scale. Thus, perfect internal consistency is not desirable.

Although there is disagreement about appropriate alpha values, general guidelines for evaluating the quality of a scale's internal consistency are helpful (Cichetti, 1994):

- Satisfactory: Alpha > .70

- Good: Alpha > .80

- Excellent: Alpha > .90

Table 1 displays the median alpha coefficients for the SSP-E scales, based on a random sample of cases that was randomly split into three equal subsamples.

Table 1. Alpha Reliability Coefficients for SSP-E Scales ($N = 14,343$)

Scale	Number of Items	Median Alpha Coefficient
Assertiveness	15	.93
Responsiveness	20	.78
Image	4	.83
Presentation	5	.77
Competence	26	.95
Feedback	18	.89

Interrater Reliability

The SSP-E is a multi-rater instrument; scores are generated based on the ratings of at least three colleagues (a self-score is reported separately). For multi-rater instruments, another form of reliability evidence comes from the consistency in ratings among raters, called interrater reliability. In other words, would my colleagues all view me as relatively similar, with "Ask" Assertive and "Control" Responsive tendencies (Analytical Style), or would some of them view me differently? This question is answered through analysis of interrater reliability.

Interrater reliability was calculated using the Intraclass Correlation Coefficient (ICC; Shrout & Fleiss, 1979). Like internal consistency, values range from 0.0 to 1.0 with higher values indicating greater consistency among raters.

Two forms of ICC were calculated, an average ICC for a single rater and an average ICC for all raters who evaluated each participant, described below:

- Average ICC for a single rater. This indicates the reliability (or relative consistency) for any given individual rater who is observing someone's behavior across all dimensions of the SSP-E. In other words, does a given individual evaluate a person consistently as he or she is responding to the SSP-E?

- Average ICC across raters. This indicates the reliability (or relative consistency) for all of the raters who evaluate any single individual. In other words, is there consistency among the group of individuals who is evaluating a person across all SSP-E scales?

These forms of ICC were calculated for a random sample of more than 9,000 participants. The average ICC for a single rater was .96, while the average ICC across raters was .99. These values indicate excellent consistency both for individual raters and for groups of individuals who evaluate a participant's behavior on the SSP-E. Table 2 displays detailed statistics for this study.

Self-ratings were not included in this analysis, and are discussed below under "Self and Other Perception."

Table 2. Interrater Reliability (*N* = 9,256)

	ICC for a Single Rater	ICC Average Over Raters
Mean	.96	.99
Standard Deviation	.03	.01
Median	.97	.99
Lowest Value	.67	.86
Highest Value	.99	1.0

Interrater Agreement

The third form of reliability evidence collected for the SSP-E is called interrater agreement, which is similar to interrater reliability but has an important distinction. Whereas interrater reliability indicates the consistency that raters have with one another across the entire instrument, interrater agreement is used to establish the *absolute* agreement among raters. In other words, it answers the question "to what degree do individuals rate a participant exactly the same?" This is calculated separately for Assertiveness, Responsiveness, and Versatility.

High levels of interrater agreement are more difficult to achieve than high interrater reliability. This is because although raters may rate consistently with one another, they won't necessarily be in perfect agreement. Obtaining high interrater agreement requires that raters assign virtually identical scores to an individual. Interrater agreement is a very stringent test, and is not often reported for psychological measures. However, because raters' SSP-E scores are aggregated to form a composite score for participants, it is necessary to show that raters agree with one another to an acceptable degree.

Agreement was evaluated using the within-group agreement statistic (r_{wg}; James, Demaree & Wolf, 1984). A random sample of 500 rater groups was chosen for this analysis (each group rated a single individual and r_{wg} was calculated for each group). The analysis discovered high levels of agreement. The average r_{wg} values were:

- Assertiveness (.96)
- Responsiveness (.97)
- Versatility (.99)

Table 3 displays detailed results from this study.

Table 3. Interrater Agreement (*N* = 500 rater groups)

	Assertiveness (r_{wg})	Responsiveness (r_{wg})	Versatility (r_{wg})
Mean	.96	.97	.99
Standard Deviation	.04	.02	.01
Median	.97	.98	.99
Lowest Value	.63	.63	.77
Highest Value	.99	.99	.99

Retest Reliability

An important aspect of assessment instruments is their stability across time, often called retest reliability. Retest reliability indicates the likelihood that a person's Profile results will remain the same or similar when profiled more than once over time. Multiple factors can affect a person's responses to the same questionnaire when taken more than once. For example, if a person is in a very good mood during the first administration, but in a very bad mood during the second administration a month later, she might respond differently. This type of unreliability is due to the individual.

Reliability can also be affected by environmental factors. There might be loud construction noises from outside a person's office window that make it difficult for her to concentrate. A more important environmental factor can affect multi-rater profiles: a different group of people might rate a person at two different administrations. This can affect results, though TRACOM's research on inter-rater reliability and agreement indicates that raters tend to evaluate individuals very similarly.

Unreliability due to the individual and due to the environment is outside of our control and can affect any given individual at any time. The study presented here focuses on the third source of retest reliability, the reliability of the instrument itself. For this type of study it is important to understand that the unit of analysis is not any given individual, but rather a large group of individuals. Statistical research virtually always applies to groups and not individuals. An instrument might demonstrate high reliability but some individuals will still score differently at different times. An analogy for this is the use of polls during elections. Based on a sample of only several thousand people, pollsters can predict the outcome of elections for entire nations within a

certain level of confidence. But of course these polls don't, or shouldn't, affect how any given individual votes.

To determine retest reliability, TRACOM analyzed data from 814 individuals who were profiled across time on our multi-rater profile. These individuals came from a variety of organizations, multiple occupations, and more than 25 industries. Seven percent of the group was from outside of North America, while the rest were from the U.S. or Canada. The time between administrations ranged from less than one month to over four years, with an average of 15.6 months.

Reliability was calculated based on individuals' multi-rater scores from co-workers, not on their own self evaluations. We based our analysis on others' ratings because the perception of others is integral to TRACOM's profiles and the lessons we teach in our materials and courses. This research design is unusual; in fact, in a literature review we found just one peer-reviewed study that examined personality retest reliability based on other-ratings (Connelly & Ones, 2010). Related to this, research has shown that others' perceptions of an individual are not only more accurate than self-perception, but are also better predictors of job performance. In a meta-analysis (an analysis of multiple research studies), researchers found that when personality profiles were based on others' perceptions, the relationship between personality and job performance was much greater than when personality profiles were based on self-perception (Oh, Wang, & Mount, 2011). In fact, using just one "other" rater made a significant difference, and the effect was magnified with multiple raters. The authors of this study concluded that the validity of personality for predicting job performance is much greater than previously believed, but this can only be shown when personality is evaluated by others who know the person.

In the current study we utilized a sample of convenience that included all individuals who had re-profiled in our database; therefore, we had no way to control whether the raters at time two were the same people who rated at time one. It is almost certain that many or most of the raters were different between the two administrations. As mentioned previously, the inability to empirically control for differences in rater groups can increase the amount of statistical "error" in ratings across time periods.

Like other forms of reliability, retest reliability is analyzed using a coefficient statistic. In general, correlations above .70 are considered reliable. Table 4 shows the correlations between the two time periods for Assertiveness, Responsiveness, and Versatility. The results show good consistency across time for the two scales that comprise SOCIAL STYLE; Assertiveness and Responsiveness. The correlation for Versatility and each of its components is lower, which is to be expected, since Versatility is less stable and can change across time and circumstances. In fact, this is one of the central principles of TRACOM's teachings and programs.

Table 4. Retest Correlations

Scale	Correlation between Times 1 and 2
Assertiveness	.73
Responsiveness	.76
Versatility	.55
Image	.59
Presentation	.53
Competence	.55
Feedback	.59

Because the time lapse between administrations varied widely among individuals, we ran partial correlations to statistically control for this effect. A partial correlation "partials out" the effects of a third variable that could be responsible for the initial correlation, insuring that the correlation between the two variables of interest is accurate and is not due to an uncontrolled variable. In this case the third variable is the amount of time between the two surveys. Controlling for time lapse did not change the correlations for any of the scales. This means that people who re-profiled years after their first profile were just as likely to maintain consistent scores as people who re-profiled only a few weeks after their first profiles.

Retest Reliability of Similar Instruments

To provide a baseline for these results, we reviewed retest reliability studies conducted on other personality and behavioral style measures.

Myers-Briggs Type Indicator®[1]

The Myers-Briggs Type Indicator (MBTI®) is an assessment of psychological type based on Carl Jung's theory of personality and is sold by CPP, Inc. Its typology is composed of four pairs of opposite preferences, called dichotomies:

- Extraversion (E) or Introversion (I)
- Sensing (S) or Intuition (N)
- Thinking (T) or Feeling (F)
- Judging (J) or Perceiving (P)

In a report released by CPP (Schaubhut, Herk, & Thompson, 2009), retest reliabilities on these four scales for the Form M assessment were calculated for time intervals ranging from less than three weeks to greater than a year. The reliabilities ranged from .67 to .73 (all time intervals combined).

---

[1] MBTI, Myers-Briggs and Myers-Briggs Type Indicator are trademarks or registered trademarks of the MBTI Trust in the United States and other countries.

DiSC® Model[2]

The DiSC model of human behavior was developed in the 1920s by William Moulton Marston, and the DiSC profile is sold by Inscape Publishing. The profile measures four dimensions of behavior: Dominance (D), Influence (i), Steadiness (S), and Conscientiousness (C). In a technical report (Inscape Publishing, 2008) one year retest reliabilities ranged from .71 to .80 on the four scales.

Big Five Personality Model

The Big Five personality model is one of the most popular and well-researched personality models in use today. It consists of five personality dimensions: Emotional Stability, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness. A meta-analysis of multiple studies that examined retest reliability on the Big Five model found reliability coefficients that ranged from .69 to .76 across the five personality dimensions (Viswesvaran & Ones, 2000).

A separate meta-analysis looked at personality trait retest reliability for people of different age groups (Roberts & DelVecchio, 2000). This study found that the consistency of personality traits increased from .31 in childhood to .54 during the college years, to .64 at age 30, and then reached a plateau around .74 between ages 50 and 70. Our research was not able to examine differences across age groups, but our findings are consistent with the highest range of reliability for personality that the meta-analysis found throughout the life span.

Finally, Connelly and Ones (2010) studied other-ratings of Big Five personality traits. In a meta-analysis they found that other-ratings of personality are measured at least as reliably as self-ratings. These authors concluded that for other-ratings to be accurate, however, other-raters

must have adequate opportunity to observe the target person. This accuracy is enhanced when other-raters have access to internal aspects of the target person's personality (thoughts, emotions, values, etc.) as a result of interpersonal intimacy.

---

[2] DiSC is a trademark of Inscape Publishing.

Distribution of Scores across Time

A helpful way to understand the consistency of scale scores across time is to visually plot scores from the two time periods against one another. Figures 2 through 4 show the frequency distributions for each scale across the two time periods. These graphs show the research group's distribution of scores on each scale for the two time periods. While these graphs do not directly plot each individual's scores across the two time periods, the consistency of the group's scores is clearly visible.
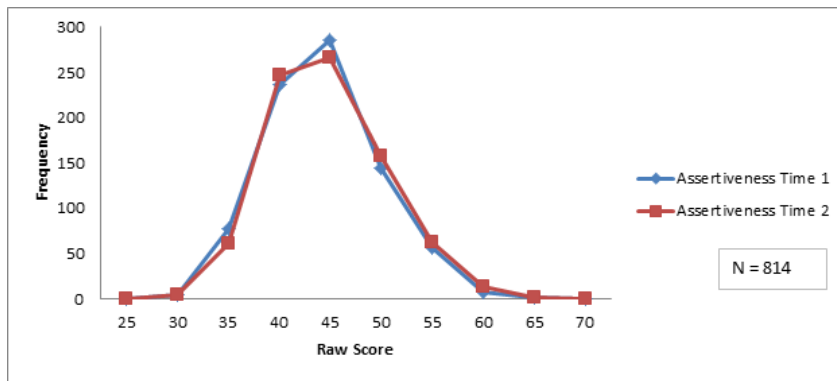

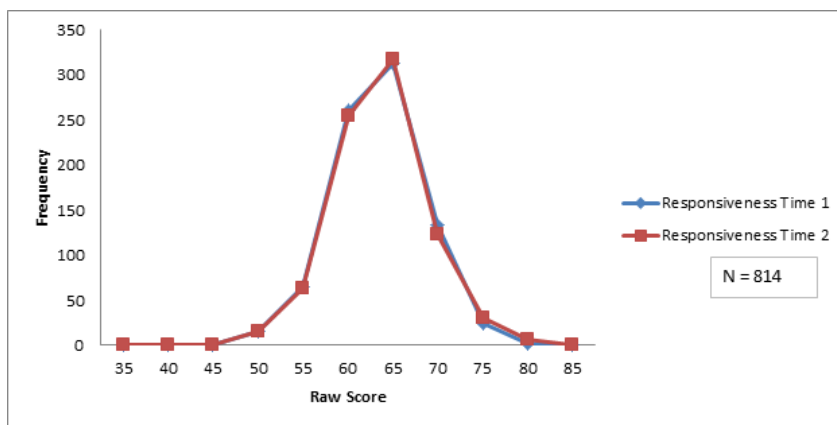Figure 2. Retest Reliability Distribution for Assertiveness

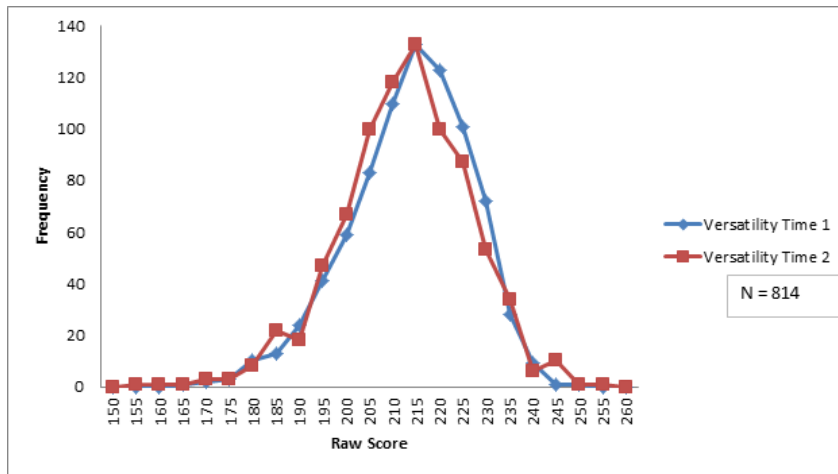Figure 3. Retest Reliability Distribution for Responsiveness.

Figure 4. Retest Reliability Distribution for Versatility.

Summary

This research study shows that TRACOM's Multi-Rater SOCIAL STYLE & Enhanced Versatility Profile has good retest reliability, specifically for the ratings of "others." While any given individual's profile results can change across time due to a variety of reasons, the measure itself has reliability that is comparable or better than other personality and behavioral style measures. Critically, this retest reliability information is based on the ratings of others, typically co-workers, showing that the behavior measured by the Profile is observable to others and remains reliably consistent over time. Other research from TRACOM has established the high degree of reliability that groups of raters have with one another when rating an individual at one point in time (i.e., inter-rater reliability and agreement). Also noteworthy is that Versatility showed lower retest reliability than the other scales. This corroborates the philosophy and design of this scale; Versatility is changeable across time and circumstances, whereas Assertiveness and Responsiveness are more stable.

Self and Other Perception

Research has shown that on multi-rater instruments, "self" ratings tend to be different from "other" ratings (Conway & Huffcutt, 1997). Some people tend to rate themselves more favorably on socially desirable traits, such as leadership abilities and interpersonal skills. This has also been found with the SSP-E.

In a study of more than 6,000 participants (plus their co-workers' ratings), it was found that self-perception of Style matched others' perceptions only 47% of the time. So, approximately half the time, people have a different view of their behavioral Style than their co-workers have of them.

What about Versatility, which includes socially desirable traits such as optimism and empathy? The results found that self-perception matches co-workers' perceptions only 35% of the time. This means that approximately 2/3 of people have a different view of their Versatility than their co-workers.

When broken down by levels of Versatility, low versus high, the study results are even more intriguing. When co-workers rated individuals as having the highest level of Versatility, a score

of "Z," only 46% of those individuals agreed with this assessment (54% rated themselves as having lower Versatility). It is possible that people with high Versatility are humble about their abilities, and also may be indicating that they still have room for improvement.

When the opposite phenomenon was examined – people whose co-workers rated as having lower Versatility ("W" score) – barely a quarter (28%) agreed with this assessment. More than 40% of these people rated themselves at least two quadrants higher ("Y" and "Z").

The next section of the report discusses validity.

Validity

Validity refers to the extent to which an instrument measures what it is supposed to measure, and also the interpretation of scores and the appropriateness of inferences drawn from those scores. The most rigorous model of validity is called construct validity. Construct validity requires several forms of evidence, including convergent/discriminant validity and factorial validity (Messick, 1989). Each of these, as well as additional forms of validity, is described below.

Convergent and Discriminant Validity

One method for establishing the construct validity of an instrument is to evaluate its relationships with other measurement instruments. An instrument should show meaningful relationships to other instruments that measure the same or similar constructs (convergent validity), and should display meaningless relationships with measures of unrelated constructs (discriminant validity).

Evidence of convergent validity is demonstrated by high correlations with similar measures, whereas discriminant validity is demonstrated by low correlations with measures that are not similar.

Two studies have shown this type of validity for the SSP-E.

Versatility and Emotional Intelligence

As previously discussed, the Versatility model is similar in some ways to emotional intelligence (EQ). Researchers at Colorado State University (Kraiger & Crane, 2009) tested this relationship by comparing Versatility to two different measures of EQ – the Trait Emotional Intelligence Questionnaire (TEIQue) and the Schutte Self Report Emotional Intelligence Test (SREIT).

The three measures were administered to 96 individuals. The TEIQue and SSP-E are multi-rater instruments; therefore, each participant also asked a group of co-workers to rate him or her using these two measures. This resulted in a sample size of 346 "other" raters. The SREIT is a self-report measure and therefore was completed by the 96 participants. A correlational analysis was performed on the scales of the instruments.

The study found that:

- Versatility self-report scores were highly and significantly correlated with TEIQue self-report scores (r = .83) and SREIT self-report scores (r = .78).

- Versatility other report scores were highly and significantly correlated with TEIQue other report scores (r = .78).

Table 5 displays these results.

Not surprisingly, correlations between "self" and "other" ratings tended to be much lower than correlations between "other" and "other" ratings on the multi-rater measures.

For example, the correlation between "self" and "other" ratings on Versatility was .43. The relationship between "self" and "other" scores on the TEIQue was .46.

This study indicates that Versatility and emotional intelligence are highly related to one another. While this is true, it is important to note that both Versatility and EQ models measure unique concepts. For example, Versatility measures Image and Presentation, two concepts that are unique to the Versatility model. Likewise, the TEIQue model assesses Emotional Perception and Happiness, two concepts that are not measured by Versatility.

The most obvious difference between Versatility and EQ is that Versatility focuses exclusively on outward behavior, while EQ measures people's internal thoughts, emotions, and beliefs.

The full report of this study is available from tracom.com.

Table 5. Correlations Between Versatility and EQ Measures

	SREIT (Self)	TEIQue (Self)	TEIQue (Other)
Versatility (Self)	.78	.83	-
Versatility (Others)	-	-	.78

All correlations significant (p<.01, 2-tail)
N = 96 for Self Measures
N = 346 for Other Measures

SOCIAL STYLE and Personality

Behavior, as measured by the SSP-E, is one aspect of our broader personalities. Because of this, we can compare the SSP-E to personality measures to determine areas where there are and are not overlaps. There should be overlaps in certain areas but not in others.

To test the relationship between SOCIAL STYLE and personality, the SSP-E and the Occupational Personality Questionnaire 32 (OPQ32) were administered to a sample of 306 undergraduate students. The OPQ32 is a personality assessment developed by SHL, an organizational consulting company, for use in human resources decision-making such as hiring. Some of the constructs measured by the OPQ32 are similar to those measured by the SSP-E, while others are distinct. Meaningful correlations with similar constructs would indicate convergent validity, while low correlations with irrelevant constructs would indicate discriminant validity. Because the SSP-E was in its initial stages of development, only a self-assessment version was administered for this study.

A correlational analysis examined the relationship of the SSP-E scales with the OPQ32 scales. Table 6 describes the relationships among these instruments. Only correlations equal to or greater than .35 are reported. In a sample this size many smaller correlations also reached statistical significance.

The Assertiveness and Responsiveness scales clearly correlate with OPQ32 scales that are similar to these constructs. The OPQ32 scales related to Assertiveness measure behaviors such as putting views across directly and not holding back in voicing criticism (Outspoken), talkativeness (Outgoing), directiveness and a desire to manage people (Controlling), and comfort with negotiation and changing others' points of view (Persuasive). Assertiveness is negatively correlated with Social Desirability, a scale that measures whether people are responding in a manner that is intended to make a good impression. The negative correlation indicates that respondents were not responding to the Assertiveness scale in an overly positive manner. There are a variety of reasons that can

Table 6. Correlations of SSP-E Survey and OPQ32 ($N = 306$)

SSP-E scale	Highest OPQ32 correlations
Assertiveness	Outspoken (.63); Outgoing (.56); Controlling (.50); Persuasive (.36); Social Desirability (-.36)
Responsiveness	Evaluative (-.40); Outgoing (.37); Data Rational (-.35)
Image	Detail Conscious (.43)
Conscientiousness	Conscientiousness (.60); Detail Conscious (.57)
Flexibility	Variety Seeking (.37)
Innovation	Innovative (.74); Conventional (-.64); Variety Seeking (.37); Independent Minded (.35)
Perseverance	Conscientiousness (.58); Controlling (.50)
Optimism	Optimistic (.82)
Empathy	Caring (.68)
Interpersonal Relations	Outgoing (.65); Affiliative (.50); Optimistic (.41); Controlling (.39); Emotionally Controlled (-.38); Caring (.35)
Communication	Controlling (.48); Outgoing (.46); Outspoken (.38)

underlie a high score on Social Desirability, including lack of self-insight, high self-esteem, and a high need to be liked. It should be noted that the final version of the SSP-E is a multi-rater instrument; therefore, socially desirable self-responses are countered by the more objective ratings of others.

The OPQ32 Outgoing scale is also related to Responsiveness. In addition to general talkativeness, the Outgoing scale measures behaviors such as being the center of attention, and viewing oneself as exuberant and fun to be with, concepts that are clearly related to the Responsiveness construct. Negative correlations with the Evaluative and Data Rational scales are a result of the direction of scoring for the SSP-E Responsiveness items (i.e., SSP-E items that are similar to these concepts are reverse scored). The Data Rational scale describes individuals who like to analyze information and base decisions on facts, while the Evaluative scale describes people who critically evaluate information and look for potential limitations in a piece of work or plan of action. Once again, these are very similar to attributes of the Responsiveness domain.

The various Versatility scales all correlate with OPQ32 scales that are similar to their SSP-E survey counterparts.

The SSP-E survey scales have near zero correlations with OPQ32 scales that are not similar in content (e.g., Assertiveness and Detail Conscious, Responsiveness and Decisive).

The combined results of these two studies provide evidence of convergent and discriminant validity for the SSP-E. The SSP-E scales are logically related to other similar measures, while also showing meaningless relationships to dissimilar measures.

Factorial Validity

Factor analysis is a statistical procedure that serves to uncover the underlying dimensions from a set of items. It is used as one form of construct validity evidence that items are measuring the constructs they are intended to measure.

In some respects factor analysis is similar to internal consistency reliability analysis described previously. However, a major distinction is that factor analysis groups items together without any *a priori* method for analyzing sets of items, whereas reliability analysis is specifically conducted on previously established scales. Therefore, the output of factor analysis is truly an "unknown" ahead of time.

The results of the factor analysis clearly displayed the scales and subscales of the SSP-E. Thirteen factors accounted for 69% of the variance in the dataset. These factors cleanly measured each scale and subscale on the instrument. For example, Assertiveness and Responsiveness, as well as the Versatility scales (Image, Presentation, and the subscales of Competence and Feedback) were all independent factors.

Scale Intercorrelations

Another form of construct validity evidence comes from examining the relationships of the scales within the SSP-E itself.

Scale intercorrelations indicate the extent to which different dimensions are related to one another. This supports construct validity by displaying that similar dimensions are more closely related to one another than dissimilar dimensions. For example, the scales that comprise Versatility (Image, Presentation, Competence, and Feedback) should be meaningfully related to one another, and have less of a relationship with Assertiveness and Responsiveness.

The SSP-E survey consists of a number of scales, and therefore they cannot be completely independent of each other. The survey is designed to provide a detailed picture of behavioral style, and like all multi-scale psychological assessments, inevitably there will be some statistical relationships among the scales.

Table 7 displays the intercorrelations among Assertiveness, Responsiveness, Versatility, and the four components of Versatility (for a random sample of cases). Assertiveness, Responsiveness, and Versatility have low intercorrelations, indicating the independence of these dimensions. Likewise, Assertiveness and Responsiveness have low correlations with each of the four

Table 7. Intercorrelations of SSP-E Survey Scales (N=8,551)

	A	R	V	I	P	C	F
A Assertiveness	1.00	.22	-.24	-.25	-.04	-.20	-.25
R Responsiveness		1.00	.19	.01	.15	.05	.38
V Versatility			1.00	.53	.75	.93	.87
I Image				1.00	.37	.41	.39
P Presentation					1.00	.66	.58
C Competence						1.00	.66
F Feedback							1.00

All coefficients significant (p<.01, 2-tail) except correlation between Image and Responsiveness

Table 8. Intercorrelations of SSP-E Survey Versatility Scales (N=8,551)

	IM	PR	C1	C2	C3	C4	C5	F1	F2	F3	F4
IM Image	1.00	.37	.40	.32	.25	.35	.35	.22	.11	.19	.25
PR Presentation		1.00	.51	.57	.57	.56	.49	.44	.35	.37	.49
C1 Conscientiousness			1.00	.56	.49	.70	.48	.35	.15	.26	.33
C2 Flexibility				1.00	.65	.62	.65	.44	.39	.42	.47
C3 Innovation					1.00	.61	.54	.42	.43	.36	.45
C4 Perseverance						1.00	.53	.36	.23	.30	.43
C5 Optimism							1.00	.33	.26	.46	.54
F1 Active Listening								1.00	.43	.39	.33
F2 Adaptive Communication									1.00	.48	.36
F3 Empathy										1.00	.64
F4 Interpersonal Relations											1.00

All coefficients significant (p<.01, 2-tail)

Versatility components, with the exception of Responsiveness and Feedback (r = 0.38), which has a modest correlation. Responsiveness measures the degree to which respondents outwardly display or control emotions. Feedback is a measure of interpersonal relations and the ability to successfully communicate and interact with others. The positive relationship between these scales may result because people who tend to be outgoing and who easily display their emotions (high Responsiveness) might also be perceived as more caring and skilled at interpersonal communication (high Feedback).

An examination of Table 7 also indicates that the four components of Versatility are meaningfully related with one another, demonstrating the construct validity of the Versatility dimension.

As a further test of the construct validity of Versatility, intercorrelations among its subdimensions were examined. Image and Presentation are each measured by single scales, while Competence and Feedback are split into various subdimensions (see Figure 1). Table 8 displays the intercorrelations among the Versatility subscales. As expected, there are meaningful correlations among the scales. This is evidence for the construct validity of the four Versatility components.

Face Validity

Face validity is the degree to which an assessment instrument appears to have relevance for a particular purpose (e.g., work performance and relationships). Face validity is most relevant for individuals who do not have training in the development of assessments, but who are the ultimate users of assessments and who are affected by assessment scores and interpretations. It is not necessary for an instrument to have face validity, but it makes its

use much more acceptable to those who receive feedback from the instrument.

The SSP-E has good face validity for its intended audience. The items and profile scales are clear descriptions of work-related behavior that most people can recognize as relevant to behavioral styles and interpersonal effectiveness. This straightforward behavioral perspective was a central aspect of the design philosophy for the SSP-E. This is not to say that all of the items are transparent to respondents, or that they would in any way know how the items are combined into the various scales.

Effectiveness Studies

In addition to the psychometric evidence given previously, another method for establishing the validity of an instrument is to determine its usefulness and applicability in the workplace. Throughout the years many studies have showcased the impact that SOCIAL STYLE has had in a variety of contexts. Many of these studies are described in David Merrill and Roger Reid's book "Personal Styles and Effective Performance" (1984). In this technical report we describe only those studies that have utilized the SSP-E and associated SOCIAL STYLE training programs.

Understanding and responding to the unique Styles of others is an important skill for working professionals. Extensive research has established that interpersonal skills are strong predictors of business and professional success in addition to cognitive ability and technical knowledge (Goleman, 1998; Goleman, Boyatzis, & McKee, 2006). In an analysis of job competencies at 286 organizations worldwide, it was found that 18 of the 21 competencies for distinguishing superior from average performers were interpersonal in nature (Spencer & Spencer, 1993). In a survey of 726 Human Resource and performance professionals, the top three most valued competencies in organizations were management leadership, technical knowledge, and people skills (BPM Forum & Success Factors, 2007). Intelligence rated near the bottom of the value scale.

Impact of SOCIAL STYLE

The importance of behavioral Style was underscored in a 2006 survey of 510 individuals who had recently completed SOCIAL STYLE programs. They were asked about their experiences and the impact that training has had in their workplaces. When asked about the impact that SOCIAL STYLE differences have on various aspects of work, there was a high level

of agreement – 87% stated they had seen conflict that was caused by Style differences. In addition, below are the percent of respondents indicating that Style differences in their workplaces had caused:

- Communication breakdowns (88%)
- Difficult relationships (76%)
- Low morale (62%)
- Negative performance (58%)

More important than the existence of these problems is participants' beliefs that applying SOCIAL STYLE concepts will help improve these situations. When asked "would applying SOCIAL STYLE help improve results in the following situations?", we found the following levels of agreement:

- Conflict (74%)
- Communication breakdowns (78%)
- Difficult relationships (75%)
- Low morale (68%)
- Negative performance (71%)

In addition to these findings, we asked about the value of using a multi-rater feedback profile. A majority (80%) felt that having a profile with "self" and "other" scores made them "more aware of challenges and opportunities that would not have otherwise been considered."

Comparison to Myers-Briggs and DiSC

Perceptions of effectiveness are important, but it is even more valuable to have actual performance data. A study published in 2010 compared the effectiveness of interpersonal skills training programs based on Inscape's DiSC model, TRACOM Group's SOCIAL STYLE Model, and CPP's Myers-Briggs Type Indicator (MBTI) (Kraiger & Kirkpatrick, 2010).

A total of 213 participants were trained on one of the three interpersonal skills training programs by a facilitator certified to deliver training on that specific program. The results showed that: 1) Participants in all three training programs held very positive reactions to the training; 2) participants in the SOCIAL STYLE training scored significantly higher than did participants in the DiSC and MBTI programs on a measure of retention of key knowledge covered in training; and 3) participants in the SOCIAL STYLE training scored significantly higher than did participants in the DiSC and MBTI programs on two measures of participants' skill at analyzing and responding to the interpersonal behaviors of others.

Thus, in terms of the effectiveness of the three programs for changing the knowledge and behavioral skills of participants, an independent evaluation of the three revealed a clear advantage for SOCIAL STYLE training. Specifically, training supporting the SOCIAL STYLE Model was found to be the most effective for improving interpersonal skills related to analyzing and responding to the behaviors of others.

Training participants, all working adults, were a diverse group in terms of gender, organizational level, work experience, and organizational setting. On average, participants were 43 years old and worked in 26 different industries, with the most frequently cited being Education or Adult Education (55), Government (37), and Manufacturing (28). Participants also reported a wide range of occupations, with the most common being Human Resources (34), Administration or Clerical (29), Information Services/Technology (16), Customer Service (13), Engineering (11), and General Manager (10).

Reaction measure. In terms of reactions to training, nearly all participants across the three programs were satisfied with the training and nearly all perceived the training they received as useful and easy to apply. This is not surprising—most people enjoy these types of classes since it gives them an opportunity to learn about themselves and how they interact with others.

Learning measure. On measures of learning, participants in the SOCIAL STYLE training scored significantly higher (80% on average) than did either participants in the DiSC training (67%) or MBTI training (60%). One possible explanation for this finding is that there are differences among measurement instruments in the extent to which supporting material is easy to grasp intuitively and encode to memory. If so, there is a clear advantage to participants receiving SOCIAL STYLE training. Participants receiving MBTI training typically accurately remembered their own profile, but struggled remembering many other key concepts covered in training.

Behavior measure. Regardless of what participants remember from the training, it is important that they be able to use the training to analyze and respond to the interpersonal behaviors of others. The researchers showed the same video (a clip from the film 12 Angry Men) to participants in each training program and measured their skill at labeling the interpersonal style or personality profile of characters in the video, and also their written answers as to how they would apply what they learned in training to work with other characters in the video given knowledge of their styles or personalities.

Again, there was a clear advantage on both measures to participants receiving the SOCIAL STYLE training. Participants in this program could identify more characters correctly (on average 2.8 of 5) than could participants in either the DiSC (1.9) or MBTI (.74) programs. Participants in the SOCIAL STYLE program also responded more accurately than participants in the other two programs when asked their strategies for working with other characters

in the video, knowing the characters' styles or personalities.

The full report of this study is available for download from tracom.com. MBTI, Myers-Briggs and Myers-Briggs Type Indicator are trademarks or registered trademarks of the MBTI Trust in the United States and other countries. DiSC is a product of Inscape Publishing.

Relationship to Job Effectiveness

As noted in a prior section, the SSP-E Versatility dimension is designed to measure behaviors that are important for interpersonal interactions that contribute to workplace effectiveness. A person's level of Versatility indicates his or her ability to interact effectively and gain the approval and support of co-workers, thereby increasing personal performance. In essence, Versatility indicates a person's ability to perform effectively regardless of his or her particular SOCIAL STYLE. Several recent studies have examined the relationship between Versatility and job performance.

Versatility and Managerial Performance

TRACOM partnered with an international publishing company to examine the relationship between Versatility and managers' job performance. We conducted this study to answer three primary questions:

1. Is Versatility related to managerial effectiveness?

2. Is there a meaningful difference in performance between managers with lower Versatility and managers with higher Versatility?

3. To what extent can Versatility and SOCIAL STYLE predict managerial performance?

Compared with managers lower in Versatility, we believed that managers higher in Versatility would perform at a higher level of effectiveness across a range of behaviors, from technical skill to coaching ability. We also believed that Versatility would contribute unique variance, or predictability, to job performance whereas SOCIAL STYLE would not. In the past we've found that SOCIAL STYLE is independent of job performance, and that individuals can succeed in their chosen fields regardless of their particular Style.

We found evidence for all three of these hypotheses. Main findings are briefly described below. The full report of this study is available for download from tracom.com.

Relationship between Versatility and Managerial Effectiveness

This research study found that Versatility is a strong indicator of workplace effectiveness. As Versatility increases, so do evaluations of job performance. Versatility was highly correlated with various important components of managers' jobs. For example, ability to coach others (.44), ability to work well within a team (.47), ability to establish effective relationships with direct reports (.51), and effectiveness as a team leader (.47), just to name a few. To put these numbers into context, it is helpful to examine some correlations among variables that are commonly understood by most people: taking aspirin daily and reduced risk of death by heart attack (.02), antihistamine use and reduced runny nose and sneezing (.11), SAT scores and subsequent college GPA (.20), effect of alcohol on aggressive behavior (.23), and relationship between weight and height among U.S. adults (.44) (Meyer et al., 2001). Thus the correlations of Versatility with managerial performance are strong and meaningful, indicating that the higher a manager's Versatility, the higher his/her performance was likely to be.

Difference in Effectiveness between Managers with Lower and Higher Versatility

Correlation analysis indicated that Versatility is positively and significantly related to workplace effectiveness. We wanted to examine specifically the differences in performance between managers with lower Versatility and managers with higher Versatility. We hypothesized that managers would differ significantly across job performance measures depending on their Versatility category.

We tested this hypothesis using analysis of variance (ANOVA). We found significant differences in job performance ratings between managers with lower Versatility and those with higher Versatility. Managers with higher Versatility had significantly higher job performance ratings on 46 of the 47 performance measures.

These findings indicate that managers' levels of Versatility are related to their effectiveness across many key indicators of job performance.

Versatility as a Predictor of Job Performance

Our third question was whether Versatility can predict job performance. We also wanted to test for the predictive effects of Assertiveness and Responsiveness, the two components that make up SOCIAL STYLE. Our hypothesis was that SOCIAL STYLE is independent of effectiveness, and that a person of any Style can be equally effective in a managerial position, whereas Versatility can to some extent predict performance.

Multiple regression analysis was used to test how well each of the three measures predicted overall job performance. Both Assertiveness and Responsiveness were found to be insignificant contributors to variance in job performance. However, Versatility accounted for 15% of the variance in overall job performance. This means that overall job performance is independent of a person's Assertiveness and Responsiveness, but is meaningfully affected by Versatility.

Versatility and Diversity Practices

In a study of 143 managers at a large multinational defense contractor, we found that managers with high Versatility were rated significantly more effective at promoting diversity and inclusiveness (D&I) than managers with lower Versatility. Importantly, these evaluations came from the managers' direct reports, those in the best position to determine D&I behaviors.

Managers with high Versatility were more likely to engage in pro-diversity behaviors, such as actively trying to understand others' experiences and perspectives, recognizing employees' contributions, fostering a welcoming environment for the team, and valuing different opinions. Highly Versatile managers were rated up to 17% more effective on these behaviors than low versatile managers.

We also found that SOCIAL STYLE has virtually no relationship to D&I practices. In fact, we discovered that Versatility accounted for 21% of the variance in D&I practices, whereas SOCIAL STYLE did not account for any of the variance.

This study is described below. To download the full report of this study, visit tracom.com.

Study Overview

Each manager's direct reports completed the multi-rater Profile and a 38-item D&I survey. This survey was developed for this study, tailored to the organization's D&I practices. It included items that measured individual manager behaviors, the impact that D&I practices have on the department and organization, and awareness of D&I initiatives. Each item was rated on a 5-point scale ranging from "strongly disagree" to "strongly agree."

Similar to the research on Versatility and managerial performance, we tested the data to answer three questions:

1. Is Versatility related to D&I practices,

2. Is there a meaningful difference in D&I practices between managers with lower Versatility and managers with higher Versatility, and

3. To what extent can Versatility and SOCIAL STYLE predict D&I practices?

Relationship between Versatility and D&I Practices

Versatility was positively and significantly correlated with 35 of the 38 D&I items, meaning that managers with high Versatility received higher ratings on these measures. Interestingly, the highest correlations were found with very specific behaviors that are directly under the control of managers, such as "My manager tries to understand others' experiences from their perspective" (r = .55) and "My manager treats me with respect" (r = .54). Although still significant, lower correlations were found between Versatility and aspects of D&I that are less directly influenced by a manager's specific behaviors, such as "The diversity and inclusion mission is directly linked to our division's strategic imperatives or business results" (r = .23) and "This organization is recognized outside of the company for its community outreach efforts" (r = .18). This indicates that employees see a strong relationship between their manager's Versatility and D&I behaviors, and these behaviors have a cascading effect on beliefs about D&I that are less directly under the influence of any individual manager.

Differences in Diversity Practices between Managers with Lower and Higher Versatility

The correlations showed that Versatility is positively and significantly related to D&I practices. The next step was to examine how

meaningful the differences in D&I practices were between managers with lower Versatility and managers with higher Versatility. Our hypothesis was that managers would differ significantly depending on their Versatility category, and this hypothesis was supported. We found that high Versatility managers did in fact measurably outperform low Versatility managers on D&I practices.

To test these differences we conducted ANOVA. We calculated an average score across the 38 D&I items, and examined differences between low and high Versatility managers on this score. The ANOVA found significant differences between managers with "W" Versatility and managers in the upper half of Versatility, those scoring "Y" or "Z". We also found that the difference in D&I scores between "X" and "Z" Versatility was significant (all mean differences were significant at the .05 level).

This indicates that there is a noticeable difference in the D&I practices between lower Versatility and higher Versatility managers.

Versatility as a Predictor of Diversity Practices

Our third hypothesis was that Versatility could, to a certain extent, predict D&I practices. We used the average D&I score to test this hypothesis. We used multiple regression analysis to test how well Versatility predicted D&I practices. We found that Versatility accounted for 21% of the variance in D&I practices. This is comparable to the influence that measures such as intelligence, education or personality have on job performance. Assertiveness and Responsiveness were included in the regression analysis but did not meaningfully predict D&I practices. We should note that the regression analysis only included the three variables of Versatility, Assertiveness, and Responsiveness, because these were the only

three variables we measured in this study. If we had accounted for other relevant variables such as age, years of tenure with the organization, or other demographic variables, this could have affected the results somewhat, though it is likely that Versatility would still have been a significant predictor.

Versatility and Police Sergeant Performance

In a study with police sergeants in the Douglas County Colorado Sheriff's Office, it was found that two important components of performance – Leadership and Employee Development – were significantly related to Versatility (Nicholson-Kluth, 2004).

Overall performance measures were collected for a very small sample of 32 sergeants who volunteered for the study. The performance evaluation was used for administrative purposes, and was therefore a "high stakes" evaluation. Although Versatility was not significantly related to this overall measure, it was significantly correlated with the Leadership ($r = .46$, $p<.01$) and Employee Development ($r = .41$, $p<.05$) components of the measure. Higher scores on Versatility were related to higher performance scores.

Leadership was defined as "promotes and influences cooperation to achieve success and effect change," and included behaviors such as:

- Builds a team with complementary strengths and abilities

- Positively changes opinions and actions of others in a desired direction

- Understands people, political dynamics and the organizational culture in order to promote change

- Sets a positive example and environment for peers and staff members

- Leads past status quo to achieve new levels of excellence or change; challenges "way it has always been done"

Employee development was defined as "plans, coaches and supports growth and development of employees' skills and abilities," and included behaviors such as:

- Creates an effective learning environment by providing tools, knowledge and opportunities for staff development

- Provides timely, honest and relevant feedback

- Recognizes and rewards hard work and achievements

- Inspires actions and opinions of others by providing a supportive environment for risk-taking

Also noteworthy is that although a very small sample of sergeants was studied, these supervisors were evenly distributed across the four SOCIAL STYLES. This supports other research showing that leaders come from all Styles, with Versatility as the distinguisher in effectiveness.

Identifying Style Through Email

The proliferation of global business and virtual teams has led people to wonder if Style can be determined through communication methods other than face-to-face, such as email. Many people work in virtual teams where they may never meet other team members face-to-face, and where the primary methods of communication are telephone and email. Virtual teams are at a higher risk of misunderstanding and conflict. In fact, one study found that people correctly interpret email messages less than 50% of the time (Kruger, Epley, & Parker, 2005).

Because misunderstandings can result from Style-related differences, and these misunderstandings can be exacerbated by email, the ability to

identify a person's Style through email can benefit individuals and teams. Correctly diagnosing people's Styles through email can help individuals understand how to interact and communicate more effectively.

An independent researcher examined the ability to correctly identify Style through email (Firari, 2007). Thirty-four managers from ten companies participated in the study, representing eight unique industries. Each participant completed the multi-rater SSP-E. In addition, each manager submitted ten business emails that he or she had written.

The study utilized a sophisticated neural networking algorithm to determine the Style of each manager. The neural network accounted for hypothesized Style-related email characteristics, such as email length, type of greeting, use of specific words (e.g., "think," "feel"), and type of salutation, if any. In addition, the program accounted for the relationship between the manager and the recipient of the email (e.g., subordinate, supervisor, peer, client).

Using the neural network strategy to determine Style resulted in the correct prediction of managers' Styles 56% of the time. While this may not seem like a high success rate, it is actually quite impressive considering that the technique did not involve any human judgment; determining each manager's Style was done entirely by a computer program. In the field of neural networks, this degree of accurate prediction is considered highly successful.

The important lesson of this study is that by paying attention to critical cues in email messages, a person can reasonably determine other people's Styles.

Sales Impact

A 2006 survey of salespeople (N = 107) who had recently completed SOCIAL STYLE training found that:

- 94% said they are more conscious about how their behavior impacts their customers
- 92% said they developed more positive relationships with their customers
- 73% said that customers are more willing to disclose relevant information
- 76% said they built relationships with prospects more quickly

In terms of direct impact on performance, respondents reported the following:

- 60% increased the speed of their sales process
- 79% improved their ability to gain ongoing sales
- 68% converted prospects to customers more quickly
- 58% closed sales they otherwise might not have

Leadership Impact

In a similar study of managers (N = 79) who had completed SOCIAL STYLE training, we found that:

- 87% of managers said that learning about SOCIAL STYLE will help them be more effective when working with others
- 86% of managers indicated that as a result of training, they were better able to determine the behavioral style of others
- 81% of managers said that the multi-rater profile made them more aware of challenges and opportunities they would not have otherwise considered

- 94% of managers had seen communication breakdowns in the workplace that were due to Style differences

- 75% of managers said that when difficulty in relationships has occurred, applying SOCIAL STYLE would improve the result

Leader Skill and Contributor Perceptions

It is well known that self-awareness is critical for improving leadership effectiveness, and that many leaders lack such self-awareness. TRACOM administered separate surveys to groups of individual contributors, managers, and executives that asked related questions. We found some interesting common perceptions as well as gaps in what leaders believe versus how they are perceived to behave.

The three groups agreed to a high degree about the causes of poor productivity in the workplace. Both individual contributors (N = 377) and managers (N = 337) rated "poor communication within or across work teams" as the number one cause of poor productivity (84% for individual contributors, and 77% for managers).

Paradoxically, when executives (N = 166) were asked to list the greatest deficiencies among first-level managers, their number one choice, agreed to by 85% of respondents, was "ineffective communication." So even though managers felt that poor communication was a problem, they were also seen as ineffective at solving the problem.
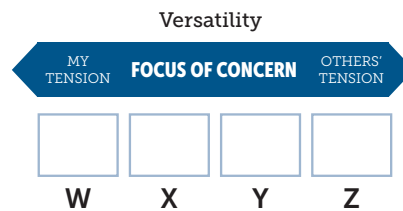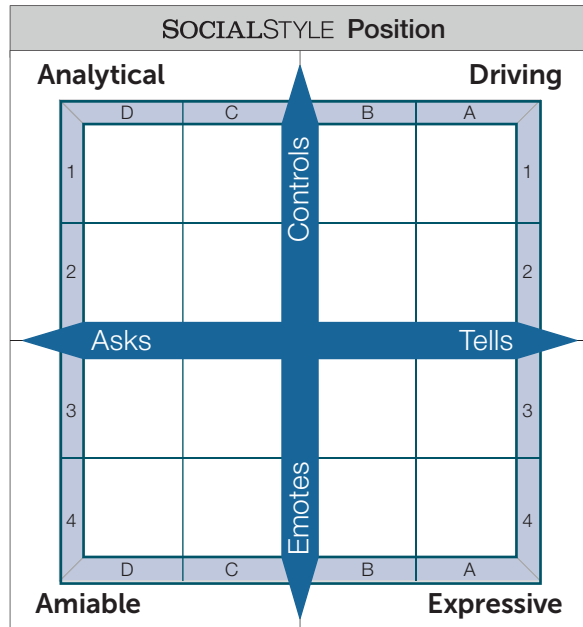
Norms

Normative scores, or norms, are important for interpreting scores. Norms provide context to individuals' scores by comparing them to a meaningful group (e.g., country, job group). Based on this, norms make it possible for us to understand what scores are low, medium and high. For the SSP-E, norms are developed by dividing the raw scale scores into quartiles (25% of the sample falls within each score range). These quartiles are used to plot profiles.

On the Assertiveness scale, the four quartile categories are labeled A, B, C, and D. Those in the "A" quartile are seen as more Tell Assertive than 75% of the norm group, while those in the "D" quartile are seen as less Tell Assertive than 75% of the norm group.

The Responsiveness scale is divided into quartiles labeled 1, 2, 3, and 4, where those in the "1" quartile are seen as more emotionally controlled than 75% of the norm group. Those in the "4" category are less emotionally controlled than 75% of the norm group.

The Versatility scale is divided into quartiles labeled W, X, Y, and Z, where those in the "W" quartile have lower Versatility than 75% of the norm group and those in the "Z" quartile have higher Versatility than 75% of the norm group. The four sources of Versatility—Image, Presentation, Competence and Feedback—are also normed in this way.

Respondents' scores on Responsiveness and Assertiveness are combined to form the SOCIAL STYLE Profile. Versatility and its components are reported separately.

The next section describes the various norm groups that exist for the SSP-E.

Norm Groups

Geographic Regions

Norms have been created for various parts of the world. It is helpful to have these norms because culture influences the way in which SOCIAL STYLE and Versatility display themselves. Style and Versatility exist in all cultures, but their relative levels often differ across cultures. In other words, every culture has its own unique attributes, and this is sometimes reflected in the way Style and Versatility are displayed.

For example, in China the ways in which people display Assertiveness are generally more direct than in western societies. When we compare China and the U.S. on Assertiveness, the Chinese

average is higher (more Tell Assertive) than the U.S. average. This does not mean that people in the U.S. do not possess Tell Assertiveness behavior, they obviously do. But on average it is at a lower level than it is in China. Therefore, we develop norms for China that adjust for their cultural attributes. This provides a more meaningful measure of Style and Versatility for people in China because they are being measured relative to others within their culture.

TRACOM develops norms for many cultures throughout the world. Check tracom.com for the most recent updates to regional and country norms.

Job Level and Occupation

TRACOM has developed norms for specific job levels and occupations. We have done this only for Versatility, not Style. This is because we have found differences in Versatility scores and this is the focus of our programs to increase personal effectiveness.

Norms have been developed for three distinct job levels: individual contributor, manager, and executive. We have found that the level of Versatility increases as job level increases. Since TRACOM has programs that are designed specifically for different job levels such as managers, we created these norms in order to provide the most representative comparisons possible in terms of Versatility.

In addition to job level, Versatility norms have been created for the sales profession. TRACOM has programs specifically tailored for sales people, and these norms provide the most meaningful reports for these purposes.

Demographic Comparisons

Norms are important for pointing out meaningful differences that exist across cultures and geographic regions. However, when various demographic characteristics are compared within a culture, there are generally few, if any, meaningful differences.

We tested whether significant differences on scale scores exist across demographic categories. This was done using the United States sample because this is where we have adequate sample sizes for this type of analysis, and also because more demographic information is collected from our U.S. sample.

Statistically speaking, large samples like those we used in the analyses increase the likelihood of finding significant differences, even when these differences are meaningless on a practical level (Cohen, 1990). Therefore, these analyses are more appropriately evaluated based on effect size.

Two types of effect sizes were calculated based on the analysis including partial eta-squared

(Eta_p^2) and Cohen's d. Partial eta-squared provides an estimate of the proportion of variance in the dependent variable (scale scores) that is related to the independent variable (demographic distinctions). Cohen's (1988) guidelines were used to interpret the effect sizes for Cohen's d: shared variance between 1% and 9% was considered as a small effect; between 10% and 24%, a medium effect; and 25% or greater, a large effect. A percentage less than 1% was considered no effect. The interpretation of effect size should be exercised with caution because rules of thumb in statistics are rarely applicable for all situations, and opinions about the magnitude of effect sizes are often in the eyes of the beholder.

When different ethnic groups within the United States were compared, no meaningful differences were found. In practical terms this means that knowing a person's ethnicity will tell you nothing about his or her score on any of the SSP-E scales.

Table 9 displays these results.

Similar results were found when comparing males to females. Differences on Assertiveness and Versatility were trivial, while a larger difference was found on Responsiveness. As a group, women scored as more Emote Responsive than men.

Table 9. Mean Comparisons for SSP-E Scales by Ethnicity based on United States Sample ($N = 21,350$)

Ethnicity	Assertiveness			Responsiveness			Versatility		
	M	SD	N	M	SD	N	M	SD	N
Native American	41.86	6.53	116	61.42	4.71	116	208.46	14.02	116
Asian	40.85	5.64	794	59.78	4.61	794	209.38	14.38	794
African American	40.62	5.34	1,470	60.86	4.39	1,470	210.53	14.06	1,470
Hispanic	40.91	5.35	1,079	61.38	4.76	1,079	210.39	14.01	1,079
Hawaiian	40.15	5.36	103	61.84	4.36	103	212.55	14.42	103
White	41.69	5.54	17,444	61.25	4.84	17,444	210.39	14.34	17,444
Other	41.44	5.61	371	61.19	4.42	344	210.64	13.71	344
$F_{(6,21343)}$		10.97*			13.95*			1.44	
Eta_p^2		0.00			0.00			0.00	

Note - In order to minimize capitalization on chance, a p-value less than .01 was used to determine statistical significance (denoted by *).

Though this difference is small, it is born out when we compare the numbers of women and men who profile in the different Style categories. Slightly more women profile as Amiable and Expressive (26% and 32%, respectively) than as Analytical or Driving (27% and 15%, respectively).

Table 10 displays these results.

Table 10. Mean Comparisons for SSP-E by Gender based on North America Sample ($N = 22,751$)

Gender	Assertiveness			Responsiveness			Versatility		
	M	SD	N	M	SD	N	M	SD	N
Male	41.79	5.45	11,174	60.11	4.67	11,174	209.13	14.43	11,174
Female	41.14	5.77	11,577	62.23	4.73	11,577	211.40	13.97	11,577
$t_{(22749)}$	8.72*			34.05*			12.05*		
Cohen's d	0.12			0.45			0.16		

Note - In order to minimize capitalization on chance, a p-value less than .01 was used to determine statistical significance (denoted by *).

We also analyzed differences across regions of the U.S. Again, no meaningful differences were found. Table 11 displays these results.

Table 11. Mean Comparisons for SSP-E by Region based on United States Sample ($N = 21838$)

Region	Assertiveness			Responsiveness			Versatility		
	M	SD	N	M	SD	N	M	SD	N
Northeastern	41.49	5.50	4,779	61.23	4.52	4,779	210.88	13.84	4,779
Southeastern	41.39	5.76	5,107	61.55	4.66	5,107	210.52	13.55	5,107
North Central	41.54	5.49	3,560	61.19	4.92	3,560	209.86	14.41	3,560
South Central	41.30	5.73	5,219	60.83	4.89	5,219	209.61	14.98	5,219
Western	41.58	5.58	3,173	61.02	5.09	3,173	210.74	14.95	3,173
$F_{(4,21833)}$	1.81			15.41*			6.89*		
Eta_p^2	0.00			0.00			0.00		

Note - In order to minimize capitalization on chance, a p-value less than .01 was used to determine statistical significance (denoted by *).

In sum, these analyses mean that there are no meaningful differences across major demographic categories in the U.S. This means that it is highly unlikely for individuals' profile results to be impacted by their demographic characteristics.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational & psychological testing.* Washington, D.C.: American Educational Research Association.

Bar-On, R. (2000). *Emotional and social intelligence: Insights from the Emotional Quotient Inventory (EQ-I).* In R. Bar-On and J.D.A. Parker (Eds.), Handbook of Emotional Intelligence. San Francisco: Jossey-Bass.

BPM Forum and Success Factors. (2007). *Performance & talent management trend survey 2007.* [White paper]. Palo Alto, CA: Author.

Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *Journal of General Psychology*, July, 2003.

Cherniss, C. & Goleman, D. (2002). *The emotionally intelligent workplace: How to select for, measure, and improve emotional intelligence in individuals, groups, and organizations.* San Francisco: Jossey-Bass.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6, 284-290.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45(12), 1304-1312.

Connelly, B. S. & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136(6), 1092-1122.

Conway, J., & Huffcutt, A. (1997). Psychometric properties of multisource performance ratings: A meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance*, 10(4), 331-360.

Crocker, L. & Algina, J. (1986). *Introduction to classical & modern test theory.* Belmont, CA: Wadsworth Group.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Firari, F. (2007). Email in style, improving corporate email communications with employees at remote locations: A quantitative study. Doctoral dissertation, Capella University.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences.* New York: Basic Books.

Goleman, D. (1995). *Emotional intelligence.* New York: Bantam Books.

Goleman, D. (1998). *Working with emotional intelligence.* New York: Bantam Books.

Goleman, D. (1998). What makes a leader? *Harvard Business Review*, 76(1), 93-102.

Goleman, D. (2006). *Social intelligence: The new science of human relationships.* New York: Bantam.

Goleman, D., Boyatzis, R.E., & McKee, A. (2002). *Primal leadership: Learning to lead with emotional intelligence.* Cambridge, MA: Harvard Business School Press.

Inscape Publishing. (2008). *Everything DiSC® Classic Validation Report.* Inscape Publishing, Inc.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85-98.

Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels*. San Francisco, CA: Berrett-Koehler.

Kraiger, K. & Crane, E. (2009). *The relationship among three measures of workplace interpersonal effectiveness*. [Whitepaper]. Fort Collins, CO: The Center for Organization Effectiveness at Colorado State University.

Kraiger, K. & Kirkpatrick, S. (2010). An empirical evaluation of three popular training programs to improve interpersonal skills. *Journal of Psychological Issues in Organizational Culture*, 1(1), 60-73.

Kruger, J., Epley, N., & Parker, J. (2005). Egocentrism over e-mail: Can we communicate as well as we think? *Journal of Personality and Social Psychology*, 89, 925-936.

McClelland, D. C. (1973). Testing for competence rather than intelligence. *American Psychologist*, 28, 1-14.

Merrill, D. W. & Reid, R. H. (1984). *Personal styles & effective performance*. Boca Raton, FL: CRC Press.

Messick, S. (1989). *Validity*. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). New York: MacMillan.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American Psychologist*, 56(2), 128-165.

Nicholson-Kluth, H. (2004). *Predicting the effectiveness of first line supervisors in law enforcement with the use of social styles concepts*. Unpublished master's thesis, Regis University, Denver, Colorado.

Oh, I., Wang, G. & Mount, M. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*. 96(4), 762-773.

Roberts, B. W. & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1), 3-25.

Salovey, P. & Mayer, J. D. (1990). *Emotional intelligence. Imagination, Cognition and Personality*, 9(3), 185-211.

Schaubhut, N. A., Herk, N. A., & Thompson, R. C. (2009). *MBTI® Form M Manual Supplement*. CPP, Inc.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.

Spencer, L.M., Jr., & Spencer, S. M. (1993). *Competence at work*. New York: Wiley.

Sternberg, R. J. (1996). *Successful intelligence: How practical and creative intelligence determine success in life*. New York: Simon and Schuster.

Viswesvaran, C. & Ones, D. S. (2000). Measurement error in "Big Five factors" personality assessment: Reliability generalization across studies and measures. E*ducational and Psychological Measurement*, 60(2), 224-235.

Your SOCIAL STYLE and Versatility certification experience doesn't have to end ...

In fact you can continue learning more about STYLE by following us on Facebook, Twitter, LinkedIn and YouTube. Each week we cover topics such as:

- Which SOCIAL STYLEs are most prevalent in different industries and occupations?
- Does gender affect a person's SOCIAL STYLE or Versatility?
- What impact do cultural differences have on SOCIAL STYLE and Versatility?
- Does Versatility vary by age?

Keep up with the latest SOCIAL STYLE news, join the discussions on TRACOM's Performance Blog, or sign up to follow us via social media at **tracom.com.**

Don't Stop Here… There's More to Discover

A wide variety of SOCIAL STYLE® & Versatility products are available to meet the needs of your organization, department, or team! SOCIAL STYLE & Versatility skills apply to all workplace interactions including management, leadership, team development, coaching, conflict, communication, customer service, and sales. Never underestimate the power of your behavior or your ability to respond appropriately to the behavior of others.

To learn more about these and other TRACOM products, visit tracom.com.