

# TECHNICAL REPORT



**SOCIAL STYLE**<sup>®</sup>  
for COMMUNICATION

**TRACOM**<sup>®</sup> GROUP  
THE SOCIAL INTELLIGENCE COMPANY<sup>®</sup>

Each facilitator is limited to a single download of these training materials.  
These materials are strictly limited for use by this facilitator.

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.

No part of this report may be reproduced, transmitted or stored in any form or by any means, electronic or mechanical, without prior written permission from The TRACOM Corporation.

© Copyright 2012, 2018, 2019, 2020, 2022, 2024 The TRACOM Corporation.  
(303) 470-4900 — (800) 221-2321 (U.S. only)

[tracom.com](http://tracom.com)

Version 4.1

# TABLE OF CONTENTS

|                                                         | PAGE |
|---------------------------------------------------------|------|
| Introduction.....                                       | 1    |
| Glossary .....                                          | 1    |
| History and Development.....                            | 2    |
| The SOCIAL STYLE Profile.....                           | 2    |
| Development of the SOCIAL STYLE Profile - Enhanced..... | 4    |
| Scientific Standards.....                               | 5    |
| Development of SOCIAL STYLE® v3 Profile .....           | 6    |
| Previous Research.....                                  | 7    |
| Psychometric Properties .....                           | 7    |
| Reliability.....                                        | 8    |
| Validity.....                                           | 14   |
| Norms .....                                             | 16   |
| Demographic Comparisons.....                            | 17   |
| Historical Research.....                                | 20   |
| Self and Other Perception.....                          | 20   |
| Interrater Reliability.....                             | 21   |
| Interrater Agreement.....                               | 22   |
| Versatility .....                                       | 23   |
| Versatility and Emotional Intelligence .....            | 23   |
| SOCIAL STYLE Compared to Myers-Briggs and DiSC .....    | 24   |
| Versatility and Performance .....                       | 25   |
| Impact of SOCIAL STYLE .....                            | 31   |
| Sales Impact .....                                      | 32   |
| Leadership Impact.....                                  | 32   |
| References.....                                         | 33   |

# Introduction

In 2020 TRACOM released SOCIAL STYLE® v3 Profile (SS v3), the latest version of our assessment. This report explains how the assessment was developed, along with reliability and validity evidence. We also include information from the previous assessment (SSP-E), which has a lengthy history of research. We include this historical research because evidence for the relationship between Versatility and job performance is still applicable to SS v3: the underlying model (not to mention many of the items on the two assessments) are the same.

This report is a companion to other materials, notably the SOCIAL STYLE & Versatility Facilitator Handbook. This report will help you understand the important role that research plays in the SOCIAL STYLE Model™ and how TRACOM ensures that the profile is accurate and dependable.

The first section describes the history of the SOCIAL STYLE Profile and Model, followed by the development and research on SS v3. We then describe the relationship between Versatility and job performance along with other psychometric research.

## GLOSSARY

Frequently used terms are necessary to understand some of this report.

- Reliability – consistency and dependability of the assessment.
- Validity –accuracy of the assessment: does it truly measure Style and Versatility?
- Correlation – extent to which two variables are related to each other. Values range from 0.0 (no relationship) to 1.0 (perfect relationship). For example, the correlation between height and weight among adults is 0.44, a strong relationship (Meyer et al., 2001).
- Item – a behavioral statement on the survey, sometimes called a “survey question.” An example of an item is *“Says what’s on their mind.”*
- Scale – a group of items that measures a single idea, for example, Assertiveness.
- Profile – output of the assessment given to participants. It includes two sections: SOCIAL STYLE and Versatility.
- Norms – provide context to results by comparing a person to a group, such as a country. Norms make it possible to determine a person’s Style and Versatility positions.

# History and Development

Since the 1960s the SOCIAL STYLE Model has been used for communication, team effectiveness, sales effectiveness, and leadership development. This section reviews the Model's history and development through the early 2000s.

## THE SOCIAL STYLE PROFILE

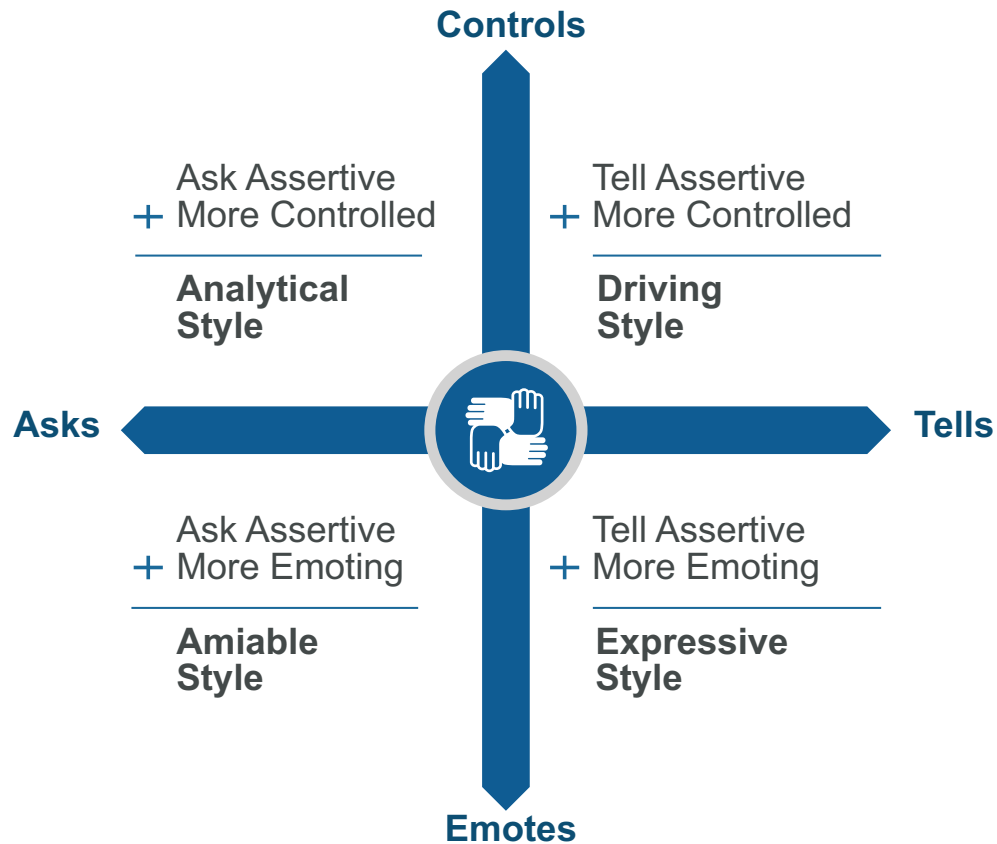
In 1964, psychologist Dr. David Merrill began researching ways to predict success in selling and management careers. He understood that people's behavior is consistent and observable and sought to measure these behaviors using adjectives. Using a technique that was unique for that time, he measured behavior using a multi-rater approach, believing that people can agree about the behavior of a person they know. It was empirical; the research was not designed to support any specific theory and could be tested by other researchers.

TRACOM's original Adjective Checklist was developed from a pool of more than 2,300 words. Work on this larger checklist had been done in the early 1960s by Dr. James Taylor (not the pop star) at the Martin Corporation (later Martin Marietta) in Denver. Dr. Merrill obtained rights to use the checklist and worked with a life insurance company that provided 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 the adjectives.

Analysis found that if a respondent felt a certain adjective described the individual's behavior, they 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 measured a dimension of 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. 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 emotions. How much a person tends to "control" or "emote" when interacting with others.
- **Versatility:** How a person adjusts their behavior to meet others' Style needs. People who are consistently versatile are seen as having good interpersonal effectiveness.

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:



Versatility is an independent scale, reported separately from SOCIAL STYLE. It consists of three elements: Presentation, Competence, and Feedback.

The SSP was used until 2003, when it was discontinued and replaced by the SOCIAL STYLE Profile - Enhanced (SSP-E).

## DEVELOPMENT OF THE SOCIAL STYLE PROFILE – ENHANCED

The SSP-E was developed from 2001 through 2003. Three factors led to the decision to revise and expand the original profile instrument.

First, the original questionnaire was a list of adjectives that people used to describe themselves and others. These adjectives resulted from empirical research conducted in the early 1960s. The growing multiculturalism of American society, along with natural language evolution, made some of the original adjectives less commonly used. In addition, the popular meanings of some of the adjectives had changed over the years.

Second, research in 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). TRACOM's concept of Versatility, originally developed in the 1960s, precedes and parallels many of the concepts of emotional intelligence.

Of the three constructs measured by the Model, 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, and therefore 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.

Finally, the third 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 ensure their meaning remains stable across cultures.

The SSP-E was used until the early 2020s, when it was revised and replaced by the SS v3.

## SCIENTIFIC STANDARDS

The profile measures behavioral style, and 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? To make this claim, the assessment has to adhere to criteria 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.

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

Psychological assessments must be both reliable and valid. Reliability is a prerequisite for validity, but is not sufficient by itself. An assessment can be reliable and still not be valid. Crocker and Algina (1986, page 217) demonstrate the difference between reliability and validity with an analogy:

*Consider 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, 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 an assessment is an important first step, but not the only step, in determining the validity of that assessment.

No psychological assessment 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 and valid an assessment is, rather than “if it is reliable and valid.” Evidence for reliability and validity is accumulated over time.

# Development of SOCIAL STYLE v3 Profile

In 2018, TRACOM began to develop the SS v3. Its predecessor, the SSP-E had been in use since 2003 and, like all psychological measures, it had reached a point where it was ready to be revised. Over time language evolves and the meaning of items can change from their original intent. Also, some items may become unnecessary because they no longer contribute to reliability or validity. For instance, items that measure Responsiveness can have too much redundancy, meaning items are too similar with one another and don't contribute unique information to the measure. In this case, some Responsiveness items can be dropped without affecting reliability or validity. TRACOM had been analyzing the items regularly over the years and had identified certain items that were ready to be revised or discarded.

In addition to these issues, the workplace has evolved. While working in different locations was not uncommon in the mid-2000s, working from home is now normal. This can make it challenging to assess one another's behavior. TRACOM needed an assessment that could account for the unique needs of virtual teams. Our goals for SS v3 were threefold:

1. Make the survey more relevant for virtual work.
2. Decrease the number of items.
3. Maintain reliability and validity.

To achieve these goals we first reviewed all items for wording concerns and especially for the ability of virtual teams to respond accurately to each item. Items that didn't meet these criteria were flagged, and we conducted item analyses to determine which of these items should be revised or dropped altogether. This was an iterative process to determine which items could be discarded, helping to shorten survey length, while still maintaining good reliability and validity (goals 2 and 3, respectively). Concurrent with this process, we wrote new items and tested them. This allowed us to collect data on new items and test their reliability and validity.

The SS v3 consists of 42 items. This was achieved by increasing the efficiency of how each scale was measured by including only the most reliable and valid items. This allowed us to maintain reliability and validity while decreasing survey length. Style is measured by 17 items (9 Assertiveness, 8 Responsiveness) and Versatility is measured by 25 items (4 Presentation, 12 Competence, and 9 Feedback).

## PREVIOUS RESEARCH

While SS v3 is an update of SSP-E, the underlying SOCIAL STYLE and Versatility model remains unchanged. Also, many of the items are carried over from SSP-E. This is important because of the implications for previous research, in particular the relationship between Versatility and job performance. The section in this Technical Report, “Versatility and Performance,” describes this research, which is still applicable with SS v3. The fundamental findings of TRACOM’s ongoing research, that Versatility predicts higher job performance and effectiveness, are as valid now as always.

## PSYCHOMETRIC PROPERTIES

Over 115,000 people have profiled using SS v3, representing every continent and world region, and over 50 countries. Below is the research conducted on the assessment.

The behavioral statements are rated on a five-point scale ranging from (1) “Strongly Disagree” to (5) “Strongly Agree.” Descriptive statistics for each scale (and the three Versatility subscales) include the mean (average score) and standard deviation (amount of dispersion from the mean). 68% of the scores lie within one standard deviation of the mean, 95% lie within two standard deviations, and 99.7% lie within three standard deviations. For example, Assertiveness has a mean of 3.32 and a standard deviation of 0.39, meaning that 68% of the scores fall between 2.93 and 3.71 (3.32 +/- 0.39).

### ***Descriptive Statistics (N = 116,337)***

| Scale          | Mean | Standard Deviation |
|----------------|------|--------------------|
| Assertiveness  | 3.32 | 0.39               |
| Responsiveness | 2.70 | 0.28               |
| Versatility    | 4.18 | 0.29               |
| Presentation   | 4.19 | 0.38               |
| Competence     | 4.23 | 0.31               |
| Feedback       | 4.13 | 0.31               |

## RELIABILITY

Here we present three types of reliability: internal consistency, item subscale correlations, and retest reliability.

### Internal Consistency

The most common type of reliability is internal consistency, which measures the relationship among survey items that measure the same thing. For instance, people should respond to all the Assertiveness items consistently, rating themselves (or others) either Telling or Asking. This is measured by Cronbach's alpha (Cronbach, 1951), which ranges from 0.0 (no relationship among items) to 1.0 (perfect consistency). A widely accepted guideline for internal consistency is (Cicchetti, 1994):

- Satisfactory: 0.70 - 0.80
- Good: 0.81 - 0.90
- Excellent: 0.91 and higher

The alpha value should not be too high since this would indicate that the scale items are redundant with one another and are not measuring unique aspects of the concept.

For these analyses we used a stratified random sample that reflects the overall demographics of our worldwide sample but without over-representing any specific company. This was done because some of our larger clients outweigh other companies in our database and we wanted to reflect a more even distribution of companies in the analyses.

The table below shows alpha coefficients for each scale. Alpha values ranged from 0.71 to 0.95. As a point of comparison, the average alpha value of personality assessments is 0.77 (Charter, 2003).

### ***Internal Consistency Reliability (N = 45,020)***

| Scale          | Number of Items | Alpha |
|----------------|-----------------|-------|
| Assertiveness  | 9               | 0.85  |
| Responsiveness | 8               | 0.71  |
| Versatility    | 25              | 0.95  |
| Presentation   | 4               | 0.90  |
| Competence     | 12              | 0.91  |
| Feedback       | 9               | 0.88  |

## Item-Scale Correlations

In a reliable scale, all items correlate moderately or strongly with the total scale score. This indicates that each item is consistent with the psychological concept its overall scale is measuring. Item-scale correlations ranged from 0.43 to 0.81, which is very good.

## Retest Reliability

Retest reliability measures the likelihood that a person's profile results will remain similar when profiled more than once over time. Different factors can lower retest reliability. For example, a person might feel rushed to complete the survey the first time but relaxed the second time, which can affect their results. Reliability can also be affected by having a different rater group at the two times. Since people's circumstances change, most people will choose a different set of raters the second time around, and this can affect their multi-rater results. It is impossible to control for these two types of unreliability.

We examined retest reliability for both self-perception and multi-rater assessments. We also examined generational differences in retest reliability.

## Self-Perception Retest Reliability

We analyzed data from 1,813 people who completed the self-perception profile two times. These individuals came from a variety of industries, occupations, and job levels, and represented every world region and over 50 countries. The time between administrations ranged from less than one month to just over three years.

We estimated retest reliability using three methods: correlation, *t*-test, and Cohen's *d*. The table below shows these results, which are explained in the paragraphs following the table. Reading the table from left to right, for each scale the correlation between times 1 and 2 is shown. To the right of this are mean (average) scores at times 1 and 2 and the differences between these means. The final two columns provide *t*-value and Cohen's *d* tests, explained below.

### Self-Perception Retest Results (N = 1,813)

|                | Correlation<br>between<br>Time 1 and<br>Time 2 | Time 1<br>Mean | Time 2<br>Mean | Difference<br>between<br>Means | <i>t</i> -value | Cohen's <i>d</i> |
|----------------|------------------------------------------------|----------------|----------------|--------------------------------|-----------------|------------------|
| Assertiveness  | .69*                                           | 29.26          | 29.48          | -.22                           | -2.18           | -.05             |
| Responsiveness | .76*                                           | 21.79          | 22.06          | -.27                           | -3.75*          | -.09             |
| Versatility    | .73*                                           | 101.44         | 101.55         | -.12                           | -0.66           | -.02             |

\* Significant ( $p < .01$ )

The first estimate of retest reliability is a correlation coefficient where values above .70 are considered high. Responsiveness and Versatility had high correlations while Assertiveness was just below the .70 threshold.

The second estimate was to test the difference between average scores at times 1 and 2. We did this using a paired samples t-test that evaluates whether the average differences between times 1 and 2 differ meaningfully. The *t*-value for Responsiveness was statistically significant. Assertiveness and Versatility were not significant, meaning their average values at times 1 and 2 are essentially the same (high retest reliability).

The third estimate measured the *practical* significance of differences between the two time periods. When using large samples like ours, the chances of finding statistically significant differences are increased, even when these differences are meaningless on a practical level (Cohen, 1990). For instance, even though the mean difference on Responsiveness is only 0.27, this change over time was statistically significant. Practical differences were measured with Cohen's *d*, using professionally accepted guidelines:

- 0.1 to 0.5: The difference is small and does not have practical significance.
- 0.6 to 0.8: The difference is moderate and may have practical significance.
- 0.9 and higher: The difference is substantial and has practical significance.

For all three scales the Cohen's *d* values are less than 0.1 (shown in the far-right column) meaning the difference between times 1 and 2 is very small and is not practically meaningful. (The negative values indicate that the mean scores at time 1 were lower than the mean scores at time 2).

Some people re-profiled after less than one month while others re-profiled several years after their initial profile. To account for this time range, we ran "partial correlations" that control for the effects of the time lapse, ensuring that the correlation is accurate. 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 have consistent results as people who re-profiled only a few weeks after their first profile.

### Multi-Rater Retest Reliability

As a further test we went beyond people's self-perception and examined retest reliability of their raters' evaluations. This is a higher bar than self-perception reliability since it can be expected that individuals would rate themselves more consistently over time than their rater groups, especially when their raters are different at time 2. 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 multi-rater data (Connelly & Ones, 2010).

We analyzed data from 629 people who completed the multi-rater profile. Each person had at least three raters. Since people choose their raters and they did this twice across administrations, it's almost certain that many or all the raters were different between the two administrations. This would likely lower retest reliability since different groups of people are providing the ratings.

Like the self-perception data, these individuals came from a variety of industries, occupations, job levels, world regions, and countries. The time between administrations ranged from less than one month to just over three years. The table below shows the results.

***Multi-Rater Retest Results (N = 629)***

|                | <b>Correlation<br/>between<br/>Time 1 and<br/>Time 2</b> | <b>Time 1<br/>Mean</b> | <b>Time 2<br/>Mean</b> | <b>Difference<br/>between<br/>Means</b> | <b>t-value</b> | <b>Cohen's <i>d</i></b> |
|----------------|----------------------------------------------------------|------------------------|------------------------|-----------------------------------------|----------------|-------------------------|
| Assertiveness  | .66*                                                     | 29.25                  | 29.56                  | -.31                                    | -2.35*         | -.09                    |
| Responsiveness | .66*                                                     | 22.00                  | 22.07                  | -.07                                    | -0.85          | -.03                    |
| Versatility    | .53*                                                     | 103.52                 | 103.68                 | -.16                                    | -0.55          | -.02                    |

\* Significant ( $p < .01$ )

Our first estimate found that Assertiveness and Responsiveness had correlations near .70. The correlation for Versatility was lower, which was expected since the goal of training is to change Versatility over time and we expect a person's co-workers to notice this behavioral change and rate the person differently.

Our second estimate, the paired samples t-test, found that the mean difference on Assertiveness was statistically significant, while Responsiveness and Versatility were not significant (high retest reliability).

Our third estimate showed that all Cohen's *d* values were less than 0.1, meaning the difference between times 1 and 2 is very small and doesn't have practical significance for any of the three scales.

Finally, partial correlations to control for the time difference between the two administrations resulted in identical results, meaning time between the profiles didn't account for any changes.

## Retest Reliability by Age Group

Since research shows that personality can evolve over time (Roberts & DelVecchio, 2000), we wanted to examine whether this was occurring with SOCIAL STYLE and Versatility. We discovered that retest reliability did in fact differ across age groups, with higher reliability for older people.

### *Self-Perception Retest Correlations by Age Group*

|                | 30 and Under | 31 to 40 | 41 to 50 | 51 and Over |
|----------------|--------------|----------|----------|-------------|
| Assertiveness  | .69          | .75      | .70      | .81         |
| Responsiveness | .71          | .76      | .78      | .81         |
| Versatility    | .72          | .77      | .73      | .80         |
| Sample Size    | 586          | 501      | 374      | 233         |

All correlations significant ( $p < .01$ )

Correlations increased as age increased (with slight decreases in the “41 to 50” group on Assertiveness and Versatility). This was especially true for people 51 and older, which corroborates the research on personality change over time. Thus, Style and Versatility behaviors stabilize as one grows older.

We could not run this analysis on multi-rater data due to low sample sizes across the different age categories, which makes results unstable.

## Retest Reliability of Similar Instruments

To provide a baseline for these results, we reviewed retest reliability studies on other personality and behavioral style measures. It’s important to note that these results are all based on self-perception data since these profiles do not collect co-workers’ perceptions.

### **Myers-Briggs Type Indicator<sup>®1</sup>**

The Myers-Briggs Type Indicator (MBTI<sup>®</sup>) measures typology based on Carl Jung’s theory of personality and is distributed by the Myers-Briggs company. It 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)

<sup>1</sup> MBTI, Myers-Briggs and Myers-Briggs Type Indicator are trademarks or registered trademarks of the MBTI Trust in the United States and other countries.

Retest correlations on the Form M assessment ranged from .67 to .73 (time intervals ranging from less than three weeks to greater than a year) (Schaubhut, Herk, & Thompson, 2009). Again, these correlations are based on self-perception only. They did not analyze average score differences using *t*-tests or Cohen's *d*.

### Big Five Personality Model

The Big Five personality model is one of the most well-researched personality models in use, and it measures: Emotional Stability, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness.

A meta-analysis of multiple studies found coefficients that ranged from .69 to .76 (Viswesvaran & Ones, 2000). Like the MBTI, these results are self-perception only and did not analyze average score differences using *t*-tests or Cohen's *d*.

A separate meta-analysis looked at retest reliability for people of different age groups (Roberts & DelVecchio, 2000). This study found that the consistency of personality increased as people get older: from .31 in childhood, .54 during college, .64 at age 30, and plateauing around .74 between ages 50 and 70. Our research corroborates this research and indicates that personality continues to evolve throughout adulthood before stabilizing in middle age.

### Retest Reliability Interpretation

Retest reliabilities are in line with professional standards and are comparable with other behavioral and personality scales. We anticipated finding high self-perception reliability since people should be consistent when evaluating themselves at different times. This is true even for Versatility since people are likely to rate themselves similarly over time, even if they're trying to improve their Versatility.

Regarding multi-rater reliability, this is a much higher standard than self-perception since having raters increases the chances of change over time. Even with this caveat, we found that Assertiveness and Responsiveness stayed near the high level of reliability found on self-perception. Thus, although having different rater groups at the two times had some effect, it wasn't great. As predicted, Versatility had a lower correlation since co-workers are more likely to see changes in these behaviors. Another influence on reliability could be the increase in remote work since 2020 (the year SS v3 was released), which might affect people's opportunities to view one another's behavior. Normal day-to-day behaviors may be less obvious when interacting through video meetings.

Finally, retest reliability increases with age. This is important because it confirms that while reliability is good overall, it is higher for older people. This should be considered when people re-profile; if they are younger adults their Style could evolve as they mature, so having some Style change is normal.

## VALIDITY

Validity determines whether an assessment measures what it is supposed to measure. The core type of validation comes through factorial validity.

### Factorial Validity

Factorial validity is the degree to which the structure underlying a set of items actually appears in a data set. Essentially, this analysis confirms that the items are measuring what they're supposed to measure. It is determined using factor analysis, which reduces the data to its primary dimensions and shows which items fit under each dimension.

Results of the factor analysis aligned with expectations. Principal factor analysis was conducted on the three scales. For Assertiveness, two factors emerged and accounted for 49% of the total variance in the data set. Responsiveness also found two factors accounting for 63% of the variance. Both Assertiveness and Responsiveness scales consist of items that were written to tap two dimensions of each construct, so these solutions are exactly what was expected.

Versatility resulted in eight factors (71% of variance). Presentation was measured by a single dimension, Competence by four dimensions, and Feedback by three dimensions, as shown in the table below.

### ***Versatility Subscales***

| Scale        | Subscales                                                                                                                                            |
|--------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Presentation | <ul style="list-style-type: none"> <li>• Effectiveness of Group Communication</li> </ul>                                                             |
| Competence   | <ul style="list-style-type: none"> <li>• Conscientiousness/ Perseverance</li> <li>• Flexibility</li> <li>• Innovation</li> <li>• Optimism</li> </ul> |
| Feedback     | <ul style="list-style-type: none"> <li>• Active Listening</li> <li>• Adaptive Communication</li> <li>• Empathy/ Interpersonal Relations</li> </ul>   |

## Scale and Subscale Intercorrelations

Additional validity evidence comes with scale and subscale correlations. If an assessment is measuring accurately, then its scales should be related to each other. For example, we would expect Assertiveness, Responsiveness, and Versatility to have low correlations with one another because, theoretically, they are independent attributes. Additionally, we would expect Presentation, Competence, and Feedback to have moderate or high correlations with one another because they are subsets of Versatility and should be somewhat related.

As shown in the table below, analyses confirmed these hypotheses. Assertiveness, Responsiveness, and Versatility have low correlations with each other, meaning they are independent of one another, while the subscales of Versatility are strongly correlated. Of course, Versatility is highly correlated with its own subscales. For a point of reference with this type of analysis, 0.7 to 0.9 is a strong correlation, 0.4 to 0.6 is a moderate correlation, and 0.01 to 0.3 is a weak correlation (Dancey & Reidy, 2004).

### *Intercorrelations of Scales and Subscales (N = 45,020)*

| Scale/ Subscale   | 1    | 2    | 3    | 4    | 5    | 6    |
|-------------------|------|------|------|------|------|------|
| 1. Assertiveness  | 1.00 | .06  | .12  | .27  | .06  | .09  |
| 2. Responsiveness |      | 1.00 | -.09 | -.16 | -.18 | .10  |
| 3. Versatility    |      |      | 1.00 | .87  | .94  | .90  |
| 4. Presentation   |      |      |      | 1.00 | .74  | .75  |
| 5. Competence     |      |      |      |      | 1.00 | .71  |
| 6. Feedback       |      |      |      |      |      | 1.00 |

## Face Validity

Face validity assesses whether an instrument subjectively appears to measure what it purports to measure. In other words, face validity is the extent to which the instrument “looks valid” to respondents. While face validity is not technical, it does suggest that respondents accept the survey and profile.

The SS v3 assessment demonstrates good face validity. Since it shares the same format as its predecessor, but has been updated to reflect more modern language and usage, the items are clear, measure with precision, and meaningfully link back to their intended constructs.

## NORMS

Norms are important for interpreting scores. Norms provide context to individuals' scores by comparing them to a meaningful group, such as people from their same country. For example, norms make it possible to say that a person is either more Ask or Tell Assertive. 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 Assertiveness, the four quartiles 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.

Responsiveness 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 people and those in the "4" category are less emotionally controlled than 75% of people.

Versatility is divided into quartiles labeled W, X, Y, and Z, where those in the "W" quartile have lower Versatility than 75% of people and those in the "Z" quartile have higher Versatility than 75% of people. The three sources of Versatility—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.

TRACOM develops unique norms for dozens of countries and world regions, and regularly updates these norms.

## DEMOGRAPHIC COMPARISONS

We tested whether significant differences exist across demographic categories. This was done using the United States sample because this is the only country where this type of demographic information is collected. As noted previously, large samples increase the likelihood of finding significant differences, even when these differences are meaningless on a practical level (Cohen, 1990). Therefore, these analyses are based on effect size.

Two types of effect sizes were calculated: partial eta-squared ( $\eta_p^2$ ) and Cohen's  $d$ . Partial eta-squared shows the proportion of variance in scores that is explained by the demographic variable. For instance, how much difference in Assertiveness scores is due to a person's age? The following guidelines were used:

- 0.01 to 0.05: The demographic variable explains a small amount of the difference.
- 0.06 to 0.10: The demographic variable explains a moderate amount of the difference.
- 0.11 and higher: The demographic variable explains a substantial amount of the difference.

### Ethnicity

When different ethnic groups were compared, no meaningful differences were found. In practical terms this means that knowing a person's ethnicity will tell you nothing about their profile results. The table below shows these results.

#### **Mean Comparisons by Ethnicity in United States (N = 31,742)**

| Ethnicity              | Assertiveness |        |        | Responsiveness |        |        | Versatility |        |        |
|------------------------|---------------|--------|--------|----------------|--------|--------|-------------|--------|--------|
|                        | M             | SD     | N      | M              | SD     | N      | M           | SD     | N      |
| Native American        | 28.16         | 3.87   | 97     | 21.85          | 3.87   | 97     | 102.54      | 7.97   | 97     |
| Asian                  | 28.81         | 3.17   | 4,809  | 21.21          | 2.16   | 4,809  | 105.60      | 7.18   | 4,809  |
| African American       | 28.19         | 3.19   | 2,414  | 21.88          | 2.25   | 2,414  | 104.81      | 8.11   | 2,414  |
| Hispanic               | 28.73         | 3.35   | 2,596  | 21.92          | 2.33   | 2,596  | 105.10      | 7.53   | 2,596  |
| Hawaiian               | 29.12         | 3.50   | 194    | 22.00          | 2.18   | 194    | 103.66      | 8.49   | 194    |
| White                  | 28.65         | 3.37   | 20,699 | 21.76          | 3.37   | 20,699 | 104.72      | 7.61   | 20,699 |
| Other                  | 28.79         | 3.34   | 933    | 21.77          | 2.42   | 933    | 104.88      | 7.59   | 933    |
| F <sub>(6,31741)</sub> |               | 11.03* |        |                | 44.98* |        |             | 11.43* |        |
| $\eta_p^2$             |               | 0.00   |        |                | 0.01   |        |             | 0.00   |        |

\* Significant ( $p < .01$ )

## Organization Level

We examined people's organization level (individual contributors, managers, and executives) and found no significant differences. These results are shown below.

### Mean Comparisons by Organization Level in United States (N = 31,742)

| Level                         | Assertiveness |      |        | Responsiveness |      |        | Versatility |      |        |
|-------------------------------|---------------|------|--------|----------------|------|--------|-------------|------|--------|
|                               | M             | SD   | N      | M              | SD   | N      | M           | SD   | N      |
| Ind. Contributor              | 27.96         | 3.35 | 10,659 | 21.91          | 2.42 | 10,659 | 104.55      | 7.69 | 10,659 |
| Manager                       | 28.90         | 3.16 | 18,524 | 21.51          | 2.23 | 18,524 | 104.95      | 7.46 | 18,524 |
| Executive                     | 30.29         | 3.76 | 1,840  | 21.51          | 2.52 | 1,840  | 105.00      | 7.99 | 1,840  |
| F <sub>(4,31020)</sub>        | 519.78*       |      |        | 103.04*        |      |        | 9.92*       |      |        |
| Eta <sup>2</sup> <sub>p</sub> | 0.03          |      |        | 0.01           |      |        | 0.00        |      |        |

\* Significant (p < .01)

## Generation (Age Group)

Many people like to talk about differences in Style and Versatility across generations. However, our analysis found no meaningful differences across age categories. While people's behavior becomes more stable as they get older (as discussed in our retest reliability results), there are no meaningful differences between average Style behaviors or Versatility by generation.

### Mean Comparisons by Age Category in United States (N = 31,742)

| Age                           | Assertiveness |      |        | Responsiveness |      |        | Versatility |      |        |
|-------------------------------|---------------|------|--------|----------------|------|--------|-------------|------|--------|
|                               | M             | SD   | N      | M              | SD   | N      | M           | SD   | N      |
| 30 or under                   | 28.44         | 3.20 | 12,568 | 21.82          | 2.27 | 12,568 | 106.18      | 7.11 | 12,568 |
| 31 to 40                      | 28.63         | 3.31 | 9,050  | 21.49          | 2.33 | 9,050  | 104.70      | 7.40 | 9,050  |
| 41 to 50                      | 29.00         | 3.45 | 6,006  | 21.62          | 2.35 | 6,006  | 103.99      | 7.83 | 6,006  |
| 51 and up                     | 28.82         | 3.51 | 4,623  | 21.79          | 2.45 | 4,623  | 102.90      | 8.31 | 4,623  |
| F <sub>(4,32243)</sub>        | 43.87*        |      |        | 40.88*         |      |        | 262.14*     |      |        |
| Eta <sup>2</sup> <sub>p</sub> | 0.00          |      |        | 0.00           |      |        | 0.02        |      |        |

\* Significant (p < .01)

## Sex

We also tested for differences on sex. For this analysis, since there are only two categories, Cohen's *d* was used to test significance. The following guidelines were used:

- 0.1 to 0.5: The difference is small and does not have practical significance.
- 0.6 to 0.8: The difference is moderate and may have practical significance.
- 0.9 and higher: The difference is substantial and has practical significance.

Differences on Assertiveness and Versatility were trivial, and though women scored slightly more Emote Responsive than men, the effect size is still small.

### **Mean Comparisons by Sex in United States (N = 31,742)**

| Gender           | Assertiveness |      |        | Responsiveness |      |        | Versatility |      |        |
|------------------|---------------|------|--------|----------------|------|--------|-------------|------|--------|
|                  | M             | SD   | N      | M              | SD   | N      | M           | SD   | N      |
| Male             | 28.86         | 3.31 | 16,531 | 21.28          | 2.30 | 16,531 | 104.28      | 7.96 | 16,531 |
| Female           | 28.43         | 3.36 | 15,788 | 22.13          | 2.29 | 15,788 | 105.52      | 7.14 | 15,788 |
| <i>t</i> (32317) | 11.55*        |      |        | 33.18*         |      |        | 14.71*      |      |        |
| Cohen's <i>d</i> | 0.13          |      |        | 0.37           |      |        | 0.16        |      |        |

\* Significant ( $p < .01$ )

In sum, there are no meaningful differences across major demographic categories in the U.S. sample. It is highly unlikely for individuals' profile results to be impacted by their demographic characteristics.

# Historical Research

The reliability and validity studies described above were conducted on the latest SS v3 assessment. Previous research, described below, highlights other important aspects of the model.

## 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 SOCIAL STYLE assessment.

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 SOCIAL 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”).

## INTERRATER RELIABILITY

Since the assessment is multi-rater (profiles are based on co-worker ratings with a separate self-profile), another form of reliability evidence comes from the consistency in ratings among raters, called interrater reliability. In other words, would your colleagues all view you as relatively similar in your behavior? 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 consistency for any given individual rater who is observing someone's behavior across all dimensions of the assessment. In other words, does a given individual evaluate another person consistently?
- Average ICC across raters. This indicates the consistency for all of the raters who evaluate any single individual. In other words, is there consistency among the group of individuals evaluating a person across all 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 person's behavior. The table below displays detailed statistics for this study.

### ***Interrater Reliability (N = 9,256)***

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

## INTERRATER AGREEMENT

Interrater agreement 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 assessment, interrater agreement is used to establish the absolute agreement among raters. In other words, it answers the question “to what degree do raters rate a person exactly the same?” This is calculated separately for Assertiveness, Responsiveness, and Versatility.

High levels of interrater agreement are more difficult to achieve than 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’ 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  $r$  was calculated for each group). The analysis discovered high levels of agreement. The average  $r_{wg}$  values were:

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

### *Interrater Agreement (N = 500 rater groups)*

|                    | <b>Assertiveness</b><br>( $r_{wg}$ ) | <b>Responsiveness</b><br>( $r_{wg}$ ) | <b>Versatility</b><br>( $r_{wg}$ ) |
|--------------------|--------------------------------------|---------------------------------------|------------------------------------|
| Mean               | .96                                  | .97                                   | .99                                |
| Standard Deviation | .04                                  | .02                                   | .01                                |
| Median             | .97                                  | .98                                   | .99                                |
| Lowest Value       | .63                                  | .63                                   | .77                                |
| Highest Value      | .99                                  | .99                                   | .99                                |

# Versatility

## VERSATILITY AND EMOTIONAL INTELLIGENCE

Versatility 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 SOCIAL STYLE Profile are multi-rater instruments; therefore, each participant also asked a group of co-workers to rate them 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 ( $r = .78$ ).

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.

### ***Correlations Between Versatility and EQ Measures***

|                      | <b>SREIT (Self)</b> | <b>TEIQue (Self)</b> | <b>TEIQue (Other)</b> |
|----------------------|---------------------|----------------------|-----------------------|
| 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 COMPARED TO MYERS-BRIGGS AND DISC

One of the most important questions to consider is how effective a training program is, especially when comparing to other programs. Peer-reviewed research (Kraiger & Kirkpatrick, 2010) has compared the effectiveness of Inscape's DiSC model, TRACOM Group's SOCIAL STYLE Model, and CPP's Myers-Briggs Type Indicator (MBTI). A description of this research can be downloaded at [tracom.com/resources](https://tracom.com/resources).

A total of 213 people participated in one of the three programs. The results showed three key findings:

- 1) Participants in all three programs had positive reactions to the training.
- 2) Participants in the SOCIAL STYLE training scored significantly higher on a measure of key knowledge covered in training.
- 3) Participants in SOCIAL STYLE training scored significantly higher on two measures of 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, SOCIAL STYLE had a clear advantage. Specifically, SOCIAL STYLE training was found to be the most effective for improving skills related to analyzing and responding to the behaviors of others, which are the stated goals of the program.

**Reaction measure.** Nearly all participants across the three programs were satisfied with the training and thought training was 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.** SOCIAL STYLE participants scored significantly higher (80% on average) than DiSC participants (67%) or MBTI participants (60%). One possible explanation 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 SOCIAL STYLE training. MBTI participants typically remembered their own profile accurately, but struggled to remember other key concepts.

**Behavior measure.** Regardless of what participants remember, it is important that they be able to analyze and respond to the interpersonal behaviors of others. The researchers showed the same video to participants in each program and measured their skill at labeling the interpersonal style or personality profile of characters, 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 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.

## VERSATILITY AND PERFORMANCE

Research has examined the relationship between Versatility and job performance. Descriptions of these studies can be downloaded at [tracom.com/resources](https://tracom.com/resources).

### Versatility and Leadership Effectiveness

TRACOM partnered with an international publishing company to examine the relationship between Versatility and leaders' job performance. We answered three 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 predict 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 described below.

### ***Relationship between Versatility and Managerial Effectiveness***

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 correlations among common variables: 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 their performance.

### ***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 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 and 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 dimensions of SOCIAL STYLE. Our hypothesis was that SOCIAL STYLE is independent of effectiveness, and that a person of any Style can be an equally effective leader, whereas Versatility can 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 Style, but is meaningfully affected by Versatility.

## Versatility and Equity, Diversity, and Inclusiveness (EDI)

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 equity, diversity, and inclusiveness (EDI) than managers with lower Versatility. Importantly, these evaluations came from the managers' direct reports, those in the best position to determine EDI behaviors.

Managers with high Versatility were more likely to engage in pro-EDI 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 EDI practices. In fact, we discovered that Versatility accounted for 21% of the variance in EDI practices, whereas SOCIAL STYLE did not account for any of the variance.

### ***Study Overview***

Each manager's direct reports completed the SOCIAL STYLE assessment and a 38-item EDI survey. This survey was developed for this study, tailored to the organization's EDI practices. It included items that measured individual manager behaviors, the impact that EDI practices have on the department and organization, and awareness of EDI 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 answered three questions:

- 1) Is Versatility related to EDI practices?
- 2) Is there a meaningful difference in EDI practices between managers with lower Versatility and managers with higher Versatility?
- 3) To what extent can Versatility and SOCIAL STYLE predict EDI practices?

### ***Relationship between Versatility and EDI Practices***

Versatility was significantly correlated with 35 of the 38 EDI 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 EDI that are less directly influenced by a manager's 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 EDI behaviors, and these behaviors have a cascading effect on beliefs about EDI that are less directly under the influence of any individual manager.

### ***Differences in EDI Practices between Managers with Lower and Higher Versatility***

The correlations showed that Versatility is related to EDI practices. The next step was to examine how meaningful the differences in EDI practices were between managers with lower and 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 measurably outperformed low Versatility managers on EDI practices.

To test these differences, we conducted ANOVA. We calculated an average score across the 38 EDI 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 EDI 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 EDI practices between lower Versatility and higher Versatility managers.

### ***Versatility as a Predictor of EDI Practices***

Our third hypothesis was that Versatility could, to a certain extent, predict EDI practices. We used the average EDI score with multiple regression to test how well Versatility predicted EDI practices. We found that Versatility accounted for 21% of the variance in EDI 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 EDI 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).

Performance measures were collected for 32 sergeants who volunteered for the study. Versatility 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 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 Text/Email

The proliferation of virtual teams has led people to wonder if Style can be determined through email/text. Virtual teams are at a higher risk of misunderstanding and conflict, with one study finding that people correctly interpret text 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 text, the ability to identify a person's Style through email can benefit individuals and teams. Correctly diagnosing Style through email can help individuals understand how to interact and communicate more effectively.

Research by Firari (2007) examined the ability to correctly identify Style through email. Thirty-four managers from ten companies participated in the study, representing eight unique industries. Each participant completed the multi-rater profile and submitted ten business emails that they 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.

This study shows that by paying attention to critical cues in email messages, a person can reasonably determine others' Styles.

# Impact of SOCIAL STYLE

The importance of SOCIAL STYLE was underscored in a survey of 510 individuals who had recently completed learning 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%)

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.”

## SALES IMPACT

A 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.

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.

Together, these studies indicate that participants see a high degree of value and real-world impact from their SOCIAL STYLE training.

# 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.
- Dancey C., Reidy J. (2004). *Statistics without maths for psychology: Using SPSS for Windows*. London: Prentice Hall.
- 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.
- Viswesvaran, C. & Ones, D. S. (2000). Measurement error in "Big Five factors" personality assessment: Reliability generalization across studies and measures. *Educational and Psychological Measurement*, 60(2), 224-235.
- 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.