TRUITY

DISC Assessment

TECHNICAL REPORT

Melanie Buford, Ph.D. and Molly Owens, M.A.

For Truity Psychometrics LLC
Roseville, CA, United States

Contents

Background

The DISC personality assessment measures fundamental aspects of behavior, in order to help individuals, leaders, and teams explore different styles of interaction in the workplace. The DISC model describes two axes of personality: How actively individuals engage with their environment, and the degree to which they expect their environment to be challenging rather than cooperative. Variations along these two axes produce four personality types, which are defined as follows:

Drive: Engaging actively with challenges. Drive types tend to be assertive leaders.

Influence: Engaging actively with support. Influence types tend to be enthusiastic leaders.

Support: Responding with support. Support types tend to be helpful contributors.

Clarity: Responding thoughtfully to challenges. Clarity types tend to be precise contributors.

This report discusses the reliability and validity of the DISC personality assessment published by Truity. The report explores the framework and development of the assessment, its statistical characteristics, and the relationship between DISC scores and real-world outcomes.

Conceptual Origins and Relevance

The theory of emotional behavior underlying the DISC assessment was developed in the 1920's by psychologist William Moulton Marston (1928). Marston proposed a system of personality based on four distinct types and their approaches to the world around them - which he called Dominance, Inducement, Submission, and Compliance. He created this framework to explain how typical people express emotions and behave in everyday situations (Bunn, 1997).

Marston developed the theory underlying the DISC, but he did not create an assessment to measure it. The first DISC assessment was created in 1956 by Walter Clarke, an industrial psychologist. Clarke's model was originally designed to help business leaders choose qualified employees, and he renamed the four factors Dominance, Influence, Steadiness and Conscientiousness. Many providers still use Clarke's terminology, but Truity's DISC model uses the words Drive, Influence, Support, and Clarity, as these terms are more readily understood today.

In the 21st century, DISC theory has been incorporated into a large number of personality assessments, for instance, the Everything DiSC® published by Wiley (2025) and Truity's own DISC assessment. Although different publishers have different approaches to this system, and continually revise their assessments to make them more accurate and relevant, the underlying concepts are similar for all DISC assessments.

Theoretical Framework and Test Design

The Truity DISC Personality Assessment is designed to allow precise yet flexible measurement of workplace behavior. It draws on a nuanced typing system that provides a broad categorization of behavior within the four main categories of DISC, as well as a more detailed classification into 12 subtypes. This allows the results to be easily understood, but also provides significant depth. Image 1 below describes the model, including the two axes upon which results are based.

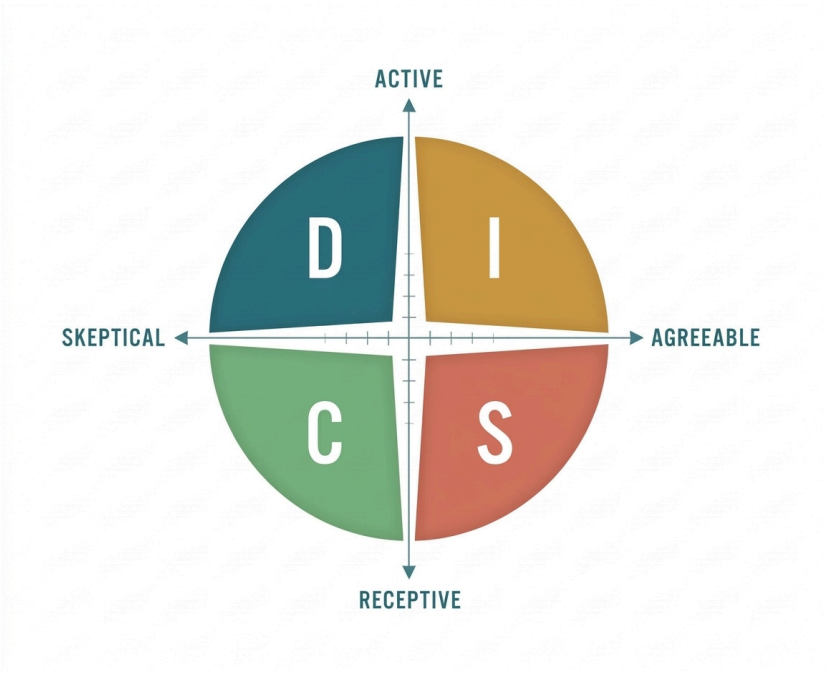

Image 1: Truity DISC Assessment Model

In total, four constructs are measured by the DISC assessment. Two of the constructs are represented by the vertical and horizontal axes of the circular DISC graphic:

- **Active vs. Receptive.** This axis runs vertically from top to bottom in the DISC graph, with D and I types representing the Active side of the spectrum, and S and C types representing the Receptive side.

- **Skeptical vs. Agreeable.** This axis runs horizontally from left to right in the DISC graph, with D and C types representing the Skeptical side of the spectrum, and I and S types representing the Agreeable side.

The DISC assessment also measures two additional constructs to help differentiate nuances between opposite types. These constructs do not appear directly in the report, but facilitate greater accuracy in scoring. They can be represented as contrasts between divergent types:

- **Drive vs. Support.** This construct contrasts the dominant, assertive character of the Drive type with the gentle, responsive character of the Support type.
- **Influence vs. Clarity.** This construct contrasts the relational, enthusiastic nature of the Influence type with the detail-oriented, reserved nature of the Clarity type.

These trait scores reflect key components of each DISC profile and are combined to create an individual's personal result - their DISC type and subtype. For instance, an individual who scores on the Agreeable side of the Skeptical vs. Agreeable axis, and scores on the Receptive side of the Active vs. Receptive axis would receive a primary result of Support, along with a corresponding subtype.

The test is a self-report assessment with a total of 38 items. Some items ask the test-taker to select between two words, such as "Helpful" and "Resolute." Others require test-takers to rate the extent to which a specific word describes them on a five-point likert-scale. Another set of items asks test-takers to complete a sentence describing themselves.

Image 2: Example Test Item

Reliability

In order to be considered psychometrically sound, assessments need to establish reliability. One statistical measure commonly used to assess the reliability of a test is Cronbach's alpha (Gerry, 2017; Vogt & Johnson, 2015). Cronbach's alpha is specifically a measure of internal consistency, or how well a test's items correlate with each other. It is possible to determine Cronbach's alpha values for each scale measured in the DISC assessment, and higher Cronbach's alpha values suggest greater internal consistency. While Cronbach's alpha values range from 0.0 to 1.0, a value of 0.70 or higher is generally considered sufficient to indicate internal consistency (Edelsbrunner et al., 2025).

Our research team calculated Cronbach's alpha - and our other measures of reliability and validity - using a sample of N = 43,516. This sample includes all Truity site users who completed the DISC assessment in February 2024. We chose to use a sample from a specific window in time to avoid seasonal demographic variation and small assessment revisions made over time.

Scale	Cronbach's Alpha
Agreeable	0.719
Receptive	0.722
Support	0.741
Influence	0.804

n= 43,516

Table 1: Cronbach's alpha Values for Each Scale

Cronbach's values are acceptable for each scale - above the 0.70 threshold - with the Influence scale demonstrating particularly strong internal consistency.

Validity

Establishing validity - that a test measures what it intends to measure - is a critical component of psychometric evaluation. Our research team used machine learning to conduct three forms of analysis to establish content validity – whether test items accurately represent the concept of interest.

First, we assessed the inter-item correlation of the test items – how well the items intended to measure the same constructs correlated with each other in practice. We also assessed factorial validity (determining what exactly a test is measuring) and conducted agglomerative clustering, which is a form of analysis that seeks to confirm that test items cluster into meaningful categories that reflect the theoretical model underlying the test (Vogt & Johnson, 2015).

Together these three methods provided justification for the validity of the DISC assessment, based on our sample of N = 43,516 test-takers.

Content Validity

Inter-item Correlation

We reviewed the test items to determine how scores on each item correlated with the others. To do this, we used two inter-item correlation matrices. Image 3 describes correlative patterns among the items we used to test the DISC scales, and Image 4 includes Pearson's correlations for each scale in the assessment.

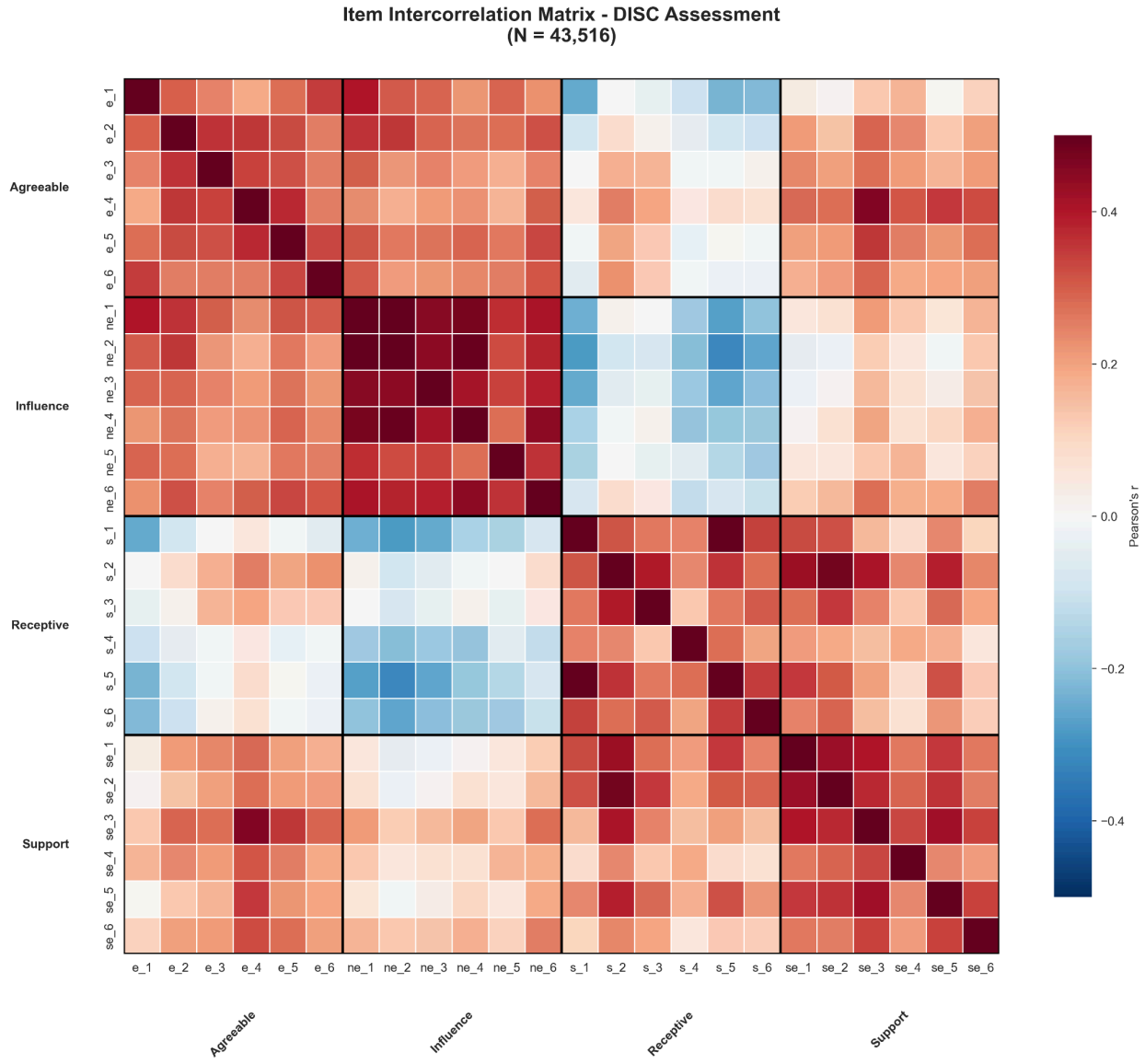

Image 3: Inter-Item Correlation Matrix

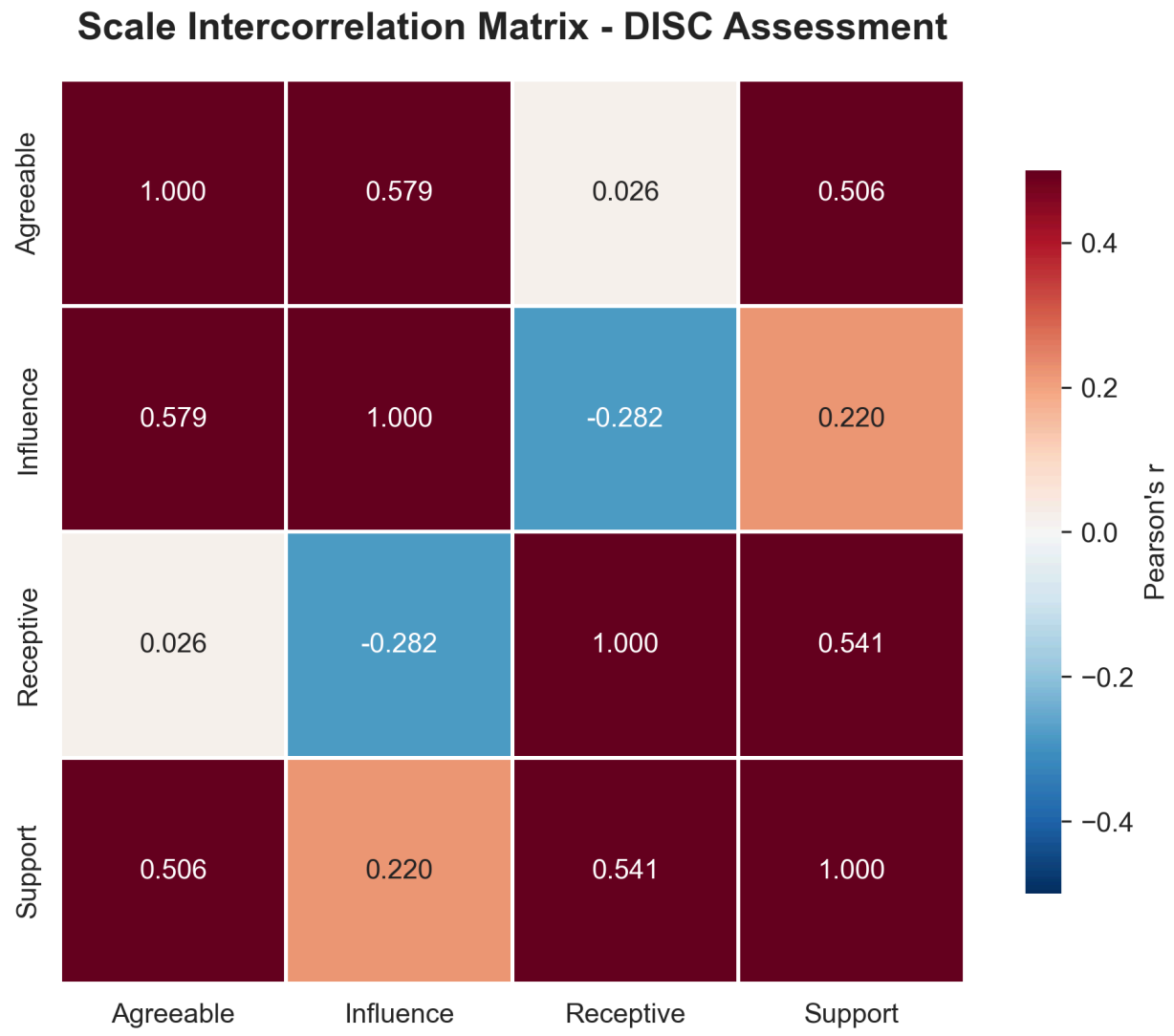

Image 4: Scale Intercorrelation Matrix (Pearson's r)

The matrices, represented as "heat maps," captured the correlative relationships for the test items, grouped by DISC facets. The red cells in the images represent higher correlations (these items tended to be answered in similar ways), and the blue cells represent lower or negative correlations (that is, these items tended to be answered dissimilarly or in opposite ways).

Thus, it is possible to see that items within each of the four categories have relatively robust correlations with each other. Items within each scale had average correlations of 0.333, while items between scales averaged 0.117, indicating that our scales hold together well. They demonstrated discriminant validity, in that items correlated more within a scale than across different scales.

Items in the Influence category and the Support category demonstrated especially strong correlations. Items in the Receptive category demonstrated a mild inverse correlation with items in the Influence category. The clustering of darker red cells along the central diagonal line indicates that items tended to loosely correlate in each category, indicating that items within categories measured similar constructs.

Factorial Validity and Agglomerative Clustering

Factorial validity is a statistical method that establishes relevant factors within data (Vogt & Johnson, 2015). In psychometric terms, it indicates the likelihood that a test is measuring something substantial. In this case, we used factor analysis to assess the strength of each of the DISC scales.

Factor analysis demonstrated that the test is consistently measuring legitimate constructs, with an overall Kaiser-Meyer-Olkin score of 0.921, which is considered “meritorious” (Vogt & Johnson, 2015). In terms of the factors themselves, we found that when asked to find four constructs, the analysis revealed that the Influence and Receptive scales loaded cleanly, that is, people who answered questions related to these scales tended to answer similarly, and these questions did not appear to unintentionally measure other constructs. We did find some cross-loading of the Agreeable and Support scales, indicating that questions in these categories had some overlap. Yet, it is not entirely unexpected that these characteristics may cluster together given their nature.

Agglomerative clustering was used to indicate the strength of connection between test items, and their corresponding scales. We sought to establish that items clustered into meaningful groups which reflected our intended theoretical scales

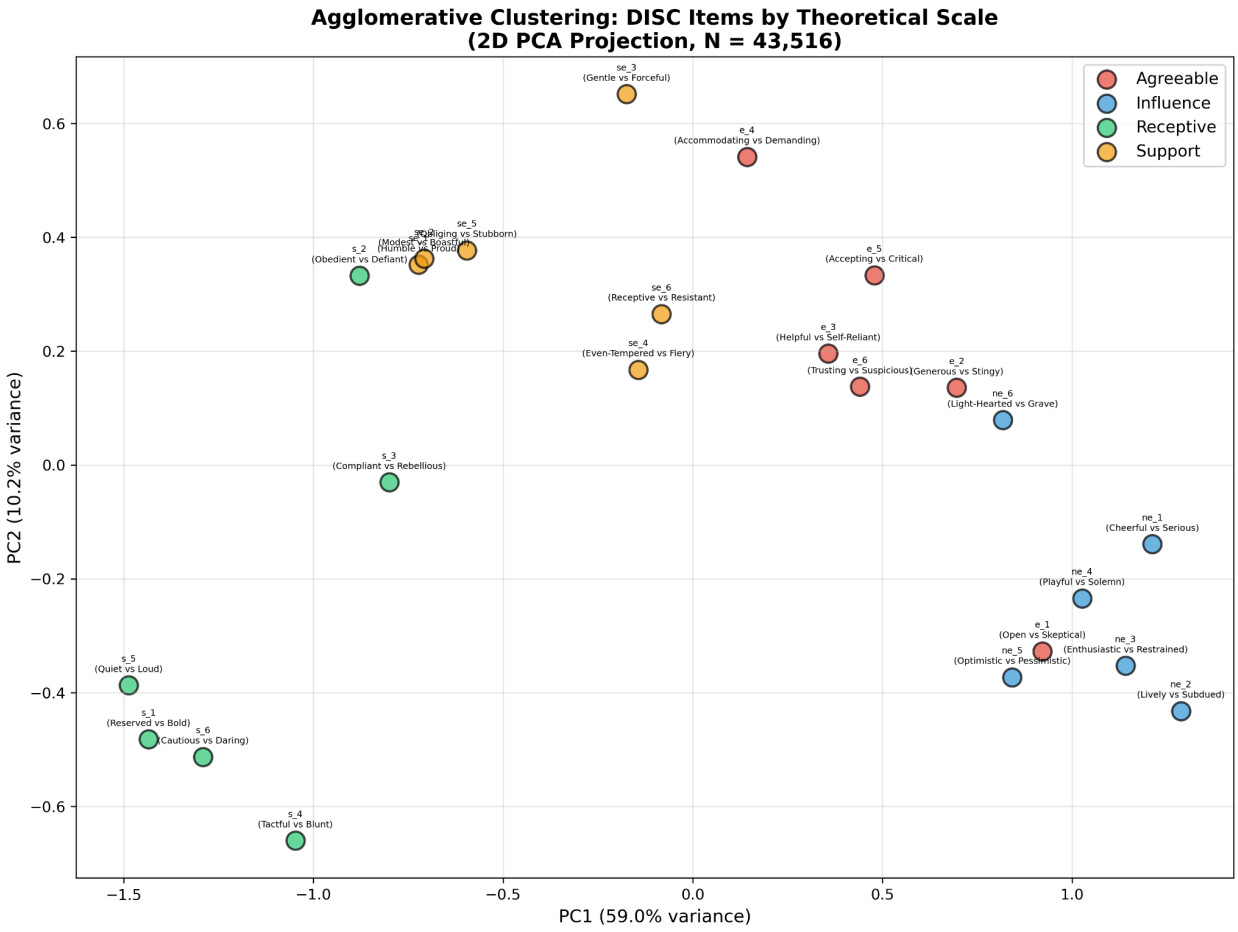

Image 5: Agglomerative Clustering for DISC Assessment Items

It is possible to discern in the agglomerative clustering results that scale items do group together in the analysis, supporting the test's content validity. Again, the Influence and Receptive scales were cleanly distinguished, while the Agreeable and Support scales had some overlap, consistent with our conceptual structure.

Predictive Validity

Predictive validity establishes that an assessment measures something relevant in the lives of test-takers (Vogt & Johnson, 2015). This is especially important for work-related psychometric assessments, and is one common way of capturing whether or not these assessments are useful and relevant in the real world (Gerry, 2017).

Demographic Statistics

First, our team gathered demographic information for test-takers in the sample. This information is included in Table 2, Table 3, and Table 4 below. Please note that in this 43,516 user sample, we received valid age responses from all users, valid gender responses from 30,930 users, and valid management responses from 29,818 users.

Age Ranges	Number of Users	Percentage
Under 18 years	12,501	28.73%
18 – 25 years	2,118	5.03%
26 – 35 years	7,119	16.36%
36 – 45 years	10,203	23.45%
46 – 55 years	5,991	13.77%
56 – 65 years	3,391	7.79%
65 – 70 years	1,661	3.82%
70+ years	462	1.06%
Total	43,516	100%

Table 2: Age of DISC Test Users (February 2024)

Identified Gender	Number of Users	Percentage
Woman	18,913	61.15%
Man	11,250	36.37%
Other/Non-binary	767	2.48%
Total	30,930	100%

Table 3: Identified Gender of DISC Test Users (February 2024)

Management Level	Number of Users	Percentage
Non-Manager	13,221	44.34%
Manages 1 – 5	8,529	28.60%
Manages 6 – 10	3,609	12.10%
Manages 11 – 50	3,245	10.88%
Manages 51+	1,214	4.07%
Total	29,818	100%

Table 4: Management Level of DISC Test Users (February 2024)

Correlative Relationships Between External Characteristics and DISC Results

We assessed the relationship between gender, age, and individual users' DISC results. Image 6 and Image 7 describe these relationships.

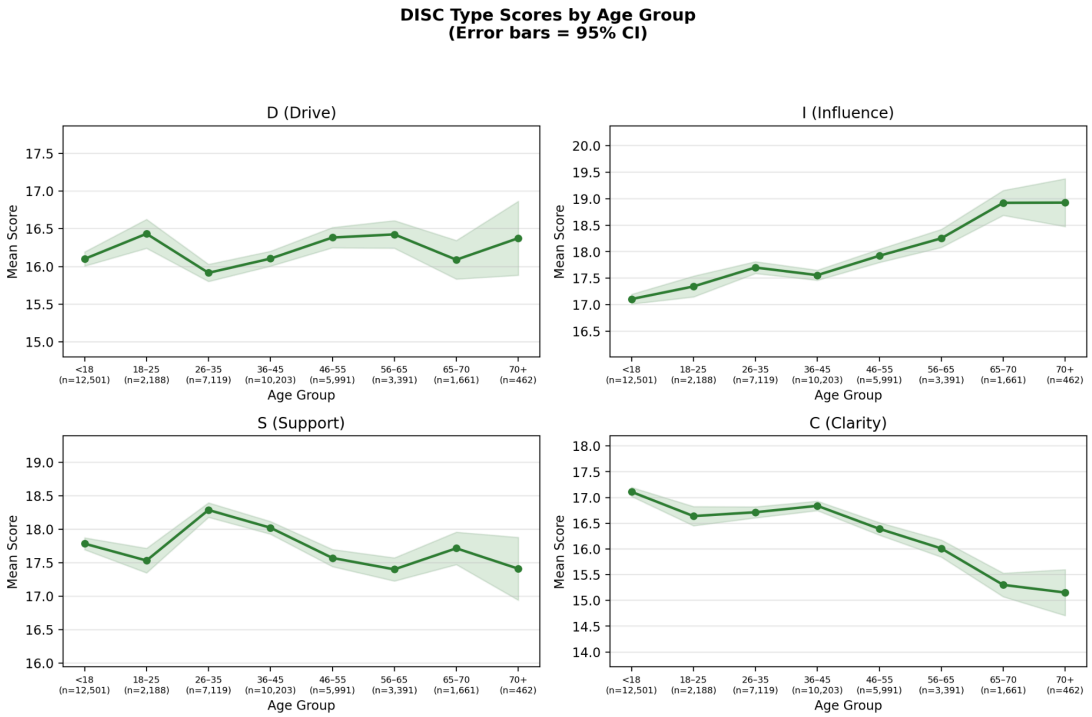

Image 6: DISC Type by Age Group

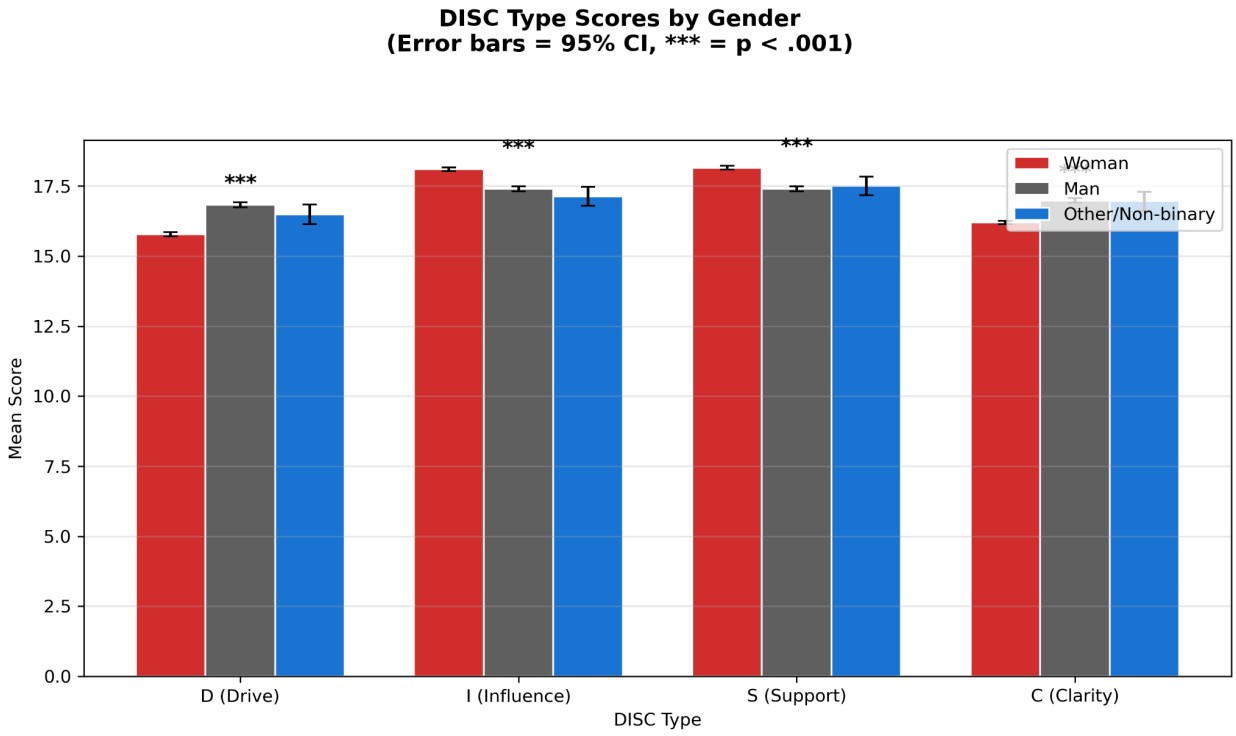

Image 7: DISC Type by Gender

These graphs indicate small differences among results for users of different ages and genders. For instance, Influence appears to increase with age and Clarity appears to decrease. The data also shows some significant differences in scores by gender for all four DISC dimensions. Each has a p value of less than 0.001, as indicated by the *** at the top of each scale. Yet, little peer-reviewed research exists on the relationship between DISC and gender, as of the time of this publication. Thus it is difficult to confirm any significant differences here.

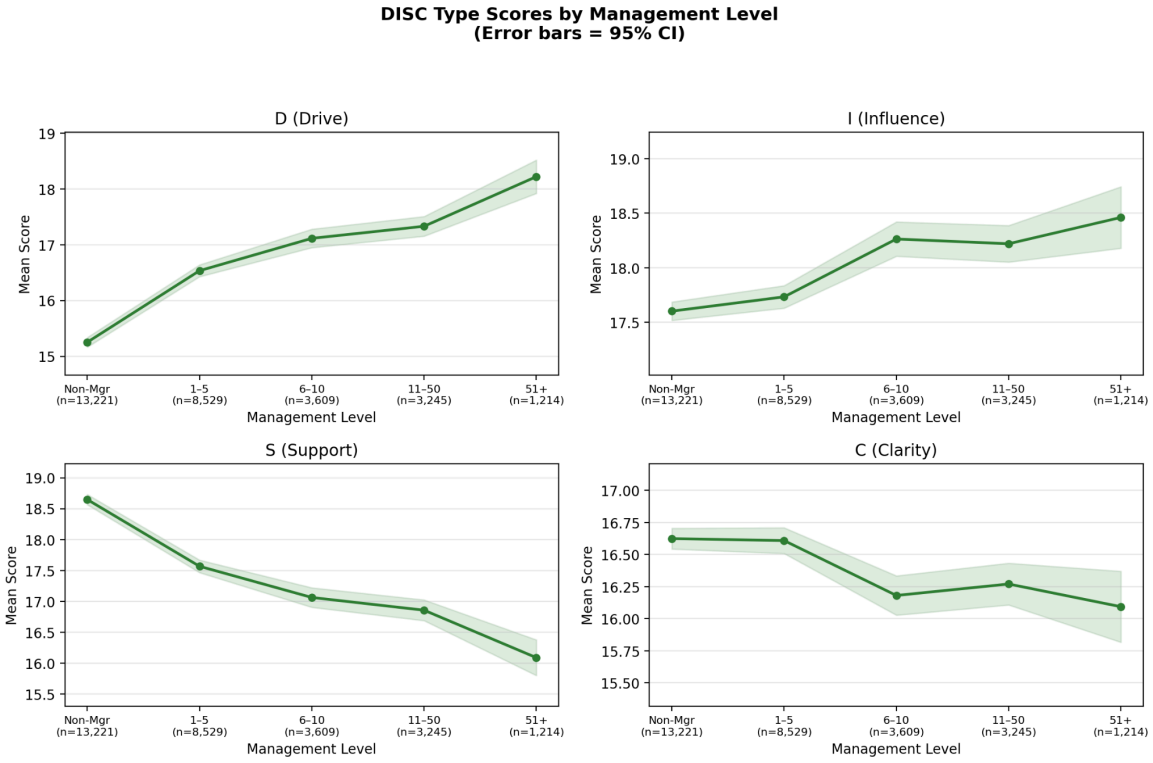

Image 8: DISC Type by Management Level

Image 9, Image 10, and Table 5 below describe the relationships between DISC scores and management status and level.

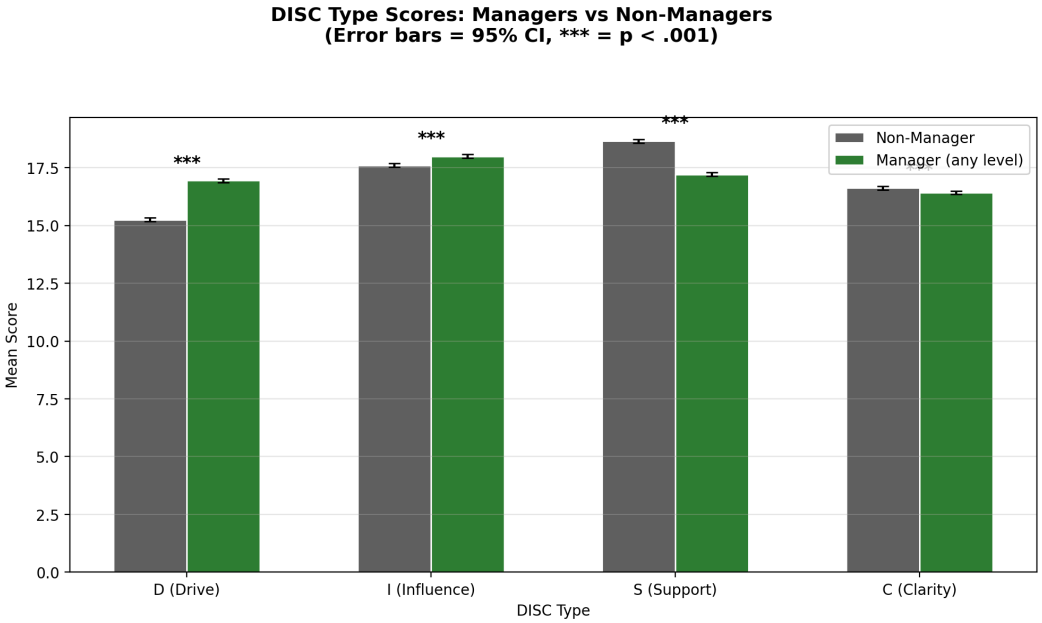

Image 9: DISC Scores by Management Status (Manager vs. Non-Manager)

DISC Type	Mean (M)	SD (M)	Mean (NM)	SD (NM)	Difference	Cohen's d
Drive	16.94	5.06	15.25	5.03	1.69	0.334
Influence	18.00	4.86	17.60	4.93	0.39	0.081
Support	17.21	4.85	18.65	4.81	-1.44	-0.297
Clarity	16.41	4.73	16.62	4.72	-0.21	-0.045

Table 5: Correlative Effect Size of DISC Scores for Managers and Non-Managers

In terms of management, our data revealed that managers had a small increase in Drive scores (Cohen's d = 0.33), and a small decrease in Support scores relative to non-managers (Cohen's d = -0.30). Influence and Clarity were not clearly correlated with management or non-management.

In summary, this report evaluated the DISC assessment created by Truity Psychometrics using multiple measures consistent with standards provided by Gerry (2017) and Vogt and Johnson (2015). We found support for the conceptualization, reliability, and validity of the assessment as an effective psychological measure.

References

Bunn, G. C. (1997). The lie detector, Wonder Woman and liberty: The life and work of William Moulton Marston. *History of the Human Sciences*, 10(1), 91–119. https://doi.org/10.1177/095269519701000105

Edelsbrunner, P. A., Simonsmeier, B. A. & Schneider, M. (2025). The Cronbach's Alpha of Domain-Specific Knowledge Tests Before and After Learning: A Meta-Analysis of Published Studies. *Educ Psychol Rev* 37(4). https://www.doi.org/10.1007/s10648-024-09982-y

Gerry, D. (2017). The practical application of test user knowledge and skills. In *Psychometric Testing*, B. Cripps (Ed.). https://www.doi.org/10.1002/9781119183020.ch5

Marston, W. M. (1928). *Emotions of normal people: by William Moulton Marston*. Harcourt, Brace & Company.

Vogt, W. P., & Johnson, R. B. (2015). Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences (5th ed.). SAGE.

Wiley. (2025). *What Is the DiSC® Model?* Everything DiSC. https://www.everythingdisc.com/what-is-disc/

Wallace, S. Rains, Clarke, Walter V., Dry, Raymond J. (September 1956). "The Activity Vector Analysis as a Selector of Life Insurance Salesmen." *Personnel Psychology* 9: 337–345. https://dx.doi.org/10.1111%2Fj.1744-6570.1956.tb01072.x