RELIABILITY AND VALIDITY OF TRUE COLORS

Judith A. Whichard, Ph.D.
True Colors Master Trainer
June 2006

Executive Summary

In an effort to ensure the ongoing quality of its programs and products, True Colors has conducted exploratory reliability and validity research. From 1998 through 2002, certain aspects of True Colors underwent the rigors of reliability and validity testing. The reliability and validity studies involved 416 participants and included individuals participating in True Colors Awareness Workshops or Level I Certification Training, as well as two (2) Psychology 100 college student groups. The participants represented a wide range of occupations, ranging from employees of public and private institutions, for profit and not-for profit organizations; had varied educational backgrounds, including individuals with and without post-secondary education; and were from all organizational levels.

The True Colors' Word Cluster proved to be highly reliable (.94) using a test-retest research format, leading to the conclusion that this assessment is understood in the same way each time it is administered, is appropriate to its intended audiences, is relatively free from measurement error, and has a high degree of overall usefulness.

True Colors showed significant content validity when measured against the MBTI and the DISC. With the exception of one dimension—the relationship between the DISC's Dominance type and True Colors' Orange type, all dimensions compared showed statistically significant content validity at the .05 confidence level.

Regarding construct validity, True Colors had significant construct validity when measured with key temperament theory concepts. The ten (10) concepts identified as underlying the True Colors program, assessments and products were judged as highly accurate by study participants in predicting and assessing behavioral characteristics and preferences.

In conclusion, True Colors showed considerable merit in precisely assessing and defining psychological types and temperament theory. As the culminating assessment tool of the True Colors processes and products, the Word Cluster proved to be highly successful in helping individuals determine their temperament types and personality preferences. Additionally, a wide range of assumptions and inferences about temperament theory based on the results achieved with the Word Cluster can be made with a high degree of credibility and correctness, particularly when relating temperament theory to individuals' behavioral preferences and characteristics in both employment and non-employment settings.

RELIABILITY AND VALIDITY OF TRUE COLORS

Judith A. Whichard, Ph.D.
True Colors Master Trainer

True Colors is committed to the ongoing quality of its training programs and products. Realizing that most endeavors are only improved by undertaking extensive evaluations and incorporating meaningful feedback into its policies and processes, True Colors has conducted exploratory reliability and validity research. The following information presents these initial, and promising, findings.

Purpose of Analyzing Measurement Instruments

True Colors recognizes that the exploration of temperament theory and psychological types has tremendous intuitive appeal, yet lacks substantive scientific empirical data on which to substantiate its theories. There are, however, ways that studies can link abstract concepts such as those of temperament theory to empirical, or numerical and measurable, indicators and thus allow for reasoned conclusions.

Such studies involve both theoretical and empirical considerations. From an empirical standpoint, the focus is on the observable response—for example, the actual cluster of words selected by the participant in the "Word Cluster" exercise. Theoretically, interest lays in the underlying unobservable (and directly unmeasurable) concept that is represented by the response—for example, an individual's color spectrum and its related temperament characteristics. These types of studies focus on the crucial relationship between the empirically grounded indicator(s)—that is, the observable response—and the underlying unobservable or directly unmeasurable concept(s).

When linkages or relationships (also called correlations) between the actual, observable (or empirical) responses and the theoretical constructs underlying assessment instruments are strong, one is able to use the empirical indicators with great confidence to measure the underlying theoretical constructs. Assumptions and inferences can be made with a high degree of credibility and accuracy.

Reliability and Validity

There are two primary ways that researchers can determine the extent to which particular empirical indicators (i.e., Word Cluster responses) represent a given theoretical concept (i.e., temperament theory as defined by True Colors). First, one can examine the reliability of an assessment instrument. Next, one can determine an instrument's validity.

It is important to note that the measurement of any phenomenon always contains a certain amount of chance error. Although ideal, the goal of error-free measurement is never attained in any area of scientific investigation.

> **Reliability Defined**: *Reliability* concerns the extent to which any measuring procedure yields the same results on repeated trials. The more consistent the results achieved by the same participants in the same repeated measurements, the higher the *reliability* of the measuring procedure; conversely the less consistent the results, the lower the *reliability*. An assessment instrument, for example, is quite reliable if an individual obtains approximately the same score or outcome on repeated examinations. *Reliability*

is an important indicator of an instrument's readability, understandability, and general usefulness.

Validity Defined: In a general sense, any measuring device is valid if it does what it is intended to do. More specifically, *validity* concerns the crucial relationship between concept and indicator. Unlike reliability that focuses on the performance of empirical measures, *validity* is usually more of a theoretically-oriented issue because it inevitably raises the question, "valid for what purpose?" *Validity* is crucial to an instrument's credibility; it is an indication that the instrument is indeed measuring what it was designed to measure and that it is measuring it accurately.

Validity, like reliability, is a matter of degree. Attaining a perfectly valid indicator—one that represents the intended, and only the intended, concept—is unachievable. However, the higher an instrument's *validity* the higher the likelihood that it is measuring the theoretical constructs for which it is expressly designed.

Content Validity: Content (also called convergent) validity is a way to determine the extent to which similar instruments obtain similar results. For example, if one is interested in determining the relationship (or content validity) between the Myers-Briggs Type Indicator (MBTI) and True Colors, a group would be administered both the MBTI and True Colors assessments. Both scores for every individual would be compared through statistical correlation to determine the extent of sameness in results.

Establishing content validity allows conclusions about the extent to which the assessment compares to a widely accepted benchmark assessment purportedly measuring the same content. It helps answer the question, "How well does the assessment instrument in question compare to other like assessment instruments?"

An important factor in establishing content validity is that at least one of the instruments used must be commonly held as a highly accurate measure of the underlying theoretical concepts. In the preceding example, the MBTI would qualify as such an instrument as it has widespread support in yielding benchmark data on temperament theory and personality types.

Construct Validity: Construct validity, another type of validity that is appropriate to psychological, temperament theory tests and assessments, is concerned with the extent to which a particular instrument relates to the theoretically derived concepts (or constructs) on which the instrument is based. For example, if one is interested in measuring personality or temperament types for the purposes of predicting inter-communication preferences, one might administer True Colors to a particular group. Next, through self-disclosure and/or researcher observations, a relationship can be determined between True Colors' predictive outcome and an individual's inter-communication preferences.

Construct validity is achieved by identifying the primary concepts that an instrument professes to measure. In this case, True Colors claims to measure specific behavioral indicators and preferences that reflect an individual's unique temperament. Once these behavioral indicators and preferences are identified,

communication styles, learning preferences, personal values and needs and so on can be predicted.

After construct validity is determined, conclusions can be drawn and assumptions made about the extent to which the assessment instrument comprehensively and accurately measures its underlying theoretical constructs.

Research Findings

To establish the merit of the True Colors' program and products, certain aspects underwent the rigors of reliability and validity testing. Conducted over a period of four (4) years (from 1998 through 2002), both the reliability and validity studies involved 416 participants and included individuals participating in either a True Colors Awareness Workshop or Level I Certification Training, as well as two (2) Psychology 100 college student groups. The participants represented a wide range of occupations, ranging from employees of public and private institutions, for profit and not-for profit organizations; educational backgrounds, including individuals with and without post-secondary education; and from all organizational levels.

As you will note in Table 1 below, True Colors has proven highly reliable using a test-retest research format. One can safely conclude that the "Word Cluster" assessment is understood in the same way each time it is administered and is relatively free from measurement error.

Additionally, True Colors shows significant content validity when measured against the MBTI and the DISC (Tables 2 and 3, respectively). With the exception of one dimension—the relationship between the DISC's Dominance type and True Colors' Orange type, all dimensions compared showed statistically significant content validity at the .05 confidence level. These findings are illustrated in Tables 2 and 3.

Although the construct validity research was not designed to yield data for statistical testing, it is apparent that True Colors has significant construct validity when measured with key temperament theory concepts. The ten (10) concepts identified as underlying the True Colors program, assessments and products were judged as highly accurate by study participants in predicting and assessing behavioral characteristics and preferences. Table 4 shows the results of the construct validity research.

Reliability Findings

To determine the reliability—the consistency of results—of the True Colors' Word Cluster, seven (7) groups were asked to participate in a test-retest study. Each group participated in a True Colors Awareness Workshop presented by the researcher. In the initial Awareness Workshop, each participant scored his/her Word Cluster exercise as part of the training. Between five (5) and six (6) weeks after the awareness workshops (enough time had passed to reduce memorized responses, yet still capitalize on the learning that occurred in the workshops), the same participants reconvened and scored an identical "Word Cluster". To help eliminate some test-retest bias, the participants were given the same instructions prior to scoring their Word Cluster in their first awareness workshop and their retest experience.

Table 1 illustrates the reliability results of the Word Cluster instrument for each group, as well as its overall reliability. In general, the Word Cluster is a highly reliable assessment instrument with an overall reliability coefficient of .940 (1.00 is perfect reliability).

Table 1. Word Cluster Test-Retest Reliability (n=167)

Group	No. of Participants	No. of Days Between Test-Retest	Reliability Coefficient
1	32	42	.945
2	19	38	.952
3	27	41	.916
4	21	36	.928
5	25	36	.933
6	29	38	.963
7	14	40	.946
	Total n = 167		Mean = .940

Content Validity Findings

To establish the content validity of True Colors, 67 participants from three (3) awareness workshops and 65 college students in two (2) Psychology 100 classes were given three (3) temperament instruments—the Myers-Briggs Type Indicator (MBTI), the DISC, and True Colors' Word Cluster. The results were scored (the MBTI was scored by a licensed test administrator) and then statistical correlations (Product-moment) were run to determine the following relationships:

1. The extent to which the MBTI and True Colors' Word Cluster results indicated similar personality, psychological, behavioral and temperament characteristics.

2. The extent to which the DISC and True Colors' Word Cluster results indicated similar personality, psychological, behavioral and temperament characteristics.

To interpret the statistical analyzes and to establish the degree of relationship between the assessments, a correlation coefficient was used. In this study, a perfect correlation, meaning a completely perfect relationship between the assessments, would be 1.00; no relationship, or no correlation, would be indicated by a coefficient of .00. Obviously, the closer the correlation coefficient is to 1.00 the stronger the relationship.

To establish statistical significance in the content validity study, a probability level of .05 was judged sufficient. This .05 probability level establishes that the results obtained from the correlation analysis would be wrong only approximately 5% of the time.

Table 3 reports the results of the correlations between the MBTI and True Colors. As you will note, all of the areas compared were statistically significant at or below the .05 probability level. These findings indicate a very strong relationship between the MBTI and the True Colors' Word Cluster.

Table 3. Relationship of the MBTI and True Colors (n=132)

Myers-Briggs MBTI	SP Perceptive	SJ Judging	NT Thinking	NF Feeling
Lowry True Colors	Orange	Gold	Green	Blue
Correlation Coefficient	*.751	*.776	*.861	*.834

*Note: Indicates those correlations that are statistically significant at or below the .05 probability level.

Table 4 represents the correlations between the DISC and the True Colors' Word Cluster. Similar to the MBTI-True Colors' findings, all but one dimension was correlated significantly at or below the .05 probability level. Again, a very strong relationship between the DISC and the True Colors' Word Cluster is evident.

Table 4. Relationship of the DISC and True Colors (n=132)

DISC	D Dominance	S Steadiness	C Compliance	I Influencing
Lowry True Colors	Orange	Gold	Green	Blue
Correlation Coefficient	**.734	*.807	*.787	*.819

*Note: Indicates those correlations that are statistically significant at or below the .05 probability level.

**Note: Indicates the correlation that is statistically significant at or below the .10 probability level.

Construct Validity Findings

To assess the construct validity of True Colors—the ability to accurately predict certain outcomes, traits or characteristics of individuals—several theoretical concepts purportedly measured by True Colors were defined. Two researchers, trained in both True Colors (both have achieved Master Trainer level) and research methodology, identified these concepts, which were subsequently validated by two independent evaluators and two independent researchers for accuracy.

Once defined, these concepts were articulated on a rating instrument that was completed by participants in True Colors Awareness Workshops and participants attending Level I Certification Training. The rating instruments asked participants for their judgments on the accuracy of True Colors in identifying the following ten (10) dimensions:

1. Communication Preferences: The way/s in which you like to receive information and the way/s in which you typically give information.

2. Learning Preferences: The types of environments and information delivery systems that encourage your learning.

3. Values: The attributes, characteristics, "things in life" that are of greatest importance to you and essentially guide your lives.

4. Strengths: The traits they possess that come to them readily and easily; the characteristics for which they are upon by others to exhibit.

5. Needs: What they need to have for positive interactions in all their relationships; life events that must occur for them to feel fulfilled.

6. Stressors: Those life events, either self or other induced, that create stress in their lives and cause them to exhibit "out-of-esteem" behaviors.

7. Diversity: The ability of True Colors to "depersonalize" events that were once ascribed to deficiencies of character rather than to legitimate differences in individual ways to achieve self-esteem.

8. Workplace Compatibility: The ability of True Colors to help you identify the reasons why you are both successful and challenged in your current workplace.

9. Self-esteem: The extent to which True Colors has identified the ways in which you achieve self-esteem.

10. Introversion and Extraversion: The ability of True Colors to help you identify where you get your "life energy"; that is, are you stimulated and energized by other people or are you prone to relying on self for energy renewal and sustenance.

Table 4 illustrates the ratings the participants from five (5) separate groups gave to True Colors on its ability to accurately predict the ten (10) identified dimensions. Note that the ratings ranged from 1 to 6 with 6 as the rating that most accurately predicted the identified dimensions and 1 as not at all accurate in predicting the dimensions. Thus, those ratings closest to 6 indicate those dimensions most accurately predicted by True Colors.

With the exception of the introversion/extraversion dimension (which True Colors introduces in workshops beyond Awareness and Level I Certification), the participants perceived a high level of accuracy by True Colors to predict temperament characteristics/behaviors on nine (9) of the ten (10) dimensions.

Table 4. Construct Validity of True Colors (n=117)

Construct	Group Ratings (n=117) (6=highly accurate to 1=no accuracy)					
	1	2	3	4	5	Mean
1. Communication preferences	5.2	5.3	4.9	5.2	5.3	5.18
2. Learning preferences	5.2	5.3	5.1	5.6	5.3	5.30
3. Values	5.1	5.1	5.2	5.3	5.4	5.22
4. Strengths	5.2	5.3	5.3	5.3	5.3	5.28
5. Needs	5.3	5.0	5.0	4.9	5.1	5.06
6. Stressors	4.9	4.8	4.8	5.1	5.2	4.96
7. Diversity	5.1	5.4	5.5	5.4	5.3	5.34
8. Workplace compatibility	4.8	4.9	5.0	5.0	5.1	4.96
9. Self-esteem	5.7	5.4	5.6	5.4	5.5	5.52
10. Introversion/extraversion	3.2	3.5	2.9	3.1	3.2	3.18

Conclusions

The purpose of this research was to determine the reliability and validity of True Colors as an accurate assessment of temperament theory, personality type, and behavioral characteristics. Based on the results obtained, the following conclusions can be drawn.

1. The findings indicate that the True Colors' Word Cluster is a highly reliable assessment instrument with an overall reliability coefficient of .94. Individuals administering and using this instrument are highly likely to obtain the same or very similar results on repeated usage of the Word Cluster.

   Additionally, users of True Colors can be assured that the Word Cluster is easily readable and understandable by the vast majority of workshop and certification training participants. Its language is appropriate to the intended audiences and meanings of the selected words are clear.

2. The research on True Colors' Word Cluster content validity with the MBTI was highly supportive of True Colors' ability to measure the same personality, psychological, behavioral and temperament characteristics as the MBTI. To a great extent, these instruments could be used inter-changeably and yield the same insights and results.

3. Similarly, the research comparing True Colors' Word Cluster content with that of the DISC indicated a strong relationship between the two instruments. With the exception of the Dominance-Orange dimension (which narrowly missed the .05 probability level with a probability level of .061), all dimensions achieved nearly the same outcomes on personality, psychological, behavioral and temperament characteristics.

4. Although not statistically based, the Word Cluster's construct validity appears to be quite high based on the reports of several study participants. This means that those characteristics and traits that True Colors claims to measure with its assessment tools and products are indeed measured through the Word Cluster. Moreover, the Word Cluster is highly likely to accurately predict individual behaviors and traits.

In summary, True Colors shows considerable merit in precisely assessing and defining psychological types and temperament theory. As the culminating assessment tool of the True Colors processes and products, the Word Cluster is proving to be highly successful in helping individuals determine their temperament types and personality preferences. Additionally, a wide range of assumptions and inferences about temperament theory based on the results achieved with the Word Cluster can be made with a high degree of credibility and correctness, particularly when relating temperament theory to individuals' behavioral preferences and characteristics in both employment and non-employment settings.

About the Researcher

Judith A. Whichard, Ph.D. received her Doctorate from Colorado State University in the School of Education. Her specialization areas include: research design and methodology, statistical design and analysis, evaluation and adult learning theory.

Professionally, she has been a Professor at Colorado State University, where she taught statistics, research methods, evaluation, and human relations at the graduate level. From Colorado State she moved to Aims Community College, serving as the Chief Information Officer. Prior to these college positions, Dr. Whichard was an international banker, inner-city secondary teacher, and a high school principal for high-risk youth. She is now President of Consulting Associates, a firm specializing in customized training in human relations, team building, leadership development, diversity, conflict resolution, mediation and personal and professional development.

Dr. Whichard began her affiliation with True Colors in 1995 as a Level I Certified Trainer. Currently, she has achieved Master Trainer status and continues to present True Colors' Basic Awareness Workshops and to train in Level I Certification. She conducted this research to satisfy her need to more fully understand the merits of the True Colors' program, and to provide empirical documentation that supported her ongoing use of True Colors' products.

Dr. Whichard continues to work in the research field as an independent evaluator, determining the effectiveness of institutional projects, policies and processes. Further, she designs policies, procedures and instruments for employee evaluations. She is also the co-author of two books—*Knowledge Nomads and the Nervously Employed: Workplace Change and Courageous Career Choices (2005)* and *The Manager as Facilitator (July 2006)*—as well as numerous journal articles.