savilleconsulting wave®

Saville Consulting Wave Professional Styles Handbook

PART 4: TECHNICAL

Chapter 19: Reliability

savilleconsulting

19.0 Reliability

When people are tested on different occasions or on different versions of the same test why do they get different scores? Because we cannot measure people's traits with perfect reliability.

Reliability of any test or assessment is concerned with how precisely the instrument measures particular characteristics or traits.

Reliability estimates provide an index of how precise and error free a tool is in measuring the desired constructs. The reliability of a test or assessment is an important prerequisite to allowing the test user to draw accurate inferences from assessment scores. The observed scores on the assessment are intended to provide an approximation of the individual's true scores. If test or profile scores are unreliable then they provide a less precise and less accurate reflection of the individual's true scores. The higher the reliability, the less the error and the more likely the observed scores are an accurate reflection of the individual's true scores.

Reliability is merely a stepping stone or prerequisite of test or questionnaire validity. If a test user is to draw a correct and meaningful inference from assessment scores, then the assessment must first be reliable. But that is not enough because the assessment should also be supported by appropriate validity data. In essence, a questionnaire must be measuring a construct reliably for it to go on to be a valid indicator from which a test user can then draw appropriate inferences and make accurate decisions. The greater the reliability, the greater the chance of high validity.

There are several methods of estimating test reliability. Three common approaches are detailed below:

Test-Retest Reliability

One estimate of reliability is to look at the stability of test scores over time. This can be accomplished by a group of individuals completing the test or assessment on one occasion and then sitting a test or assessment again on another occasion.

The (Pearson Product-Moment) correlation coefficient between how the group scores on a scale on one occasion and then on the second occasion provides this estimate of reliability.

A development aim of Wave Styles was that this form of reliability should be as high as possible.

Alternate Form Reliability

Where two or more versions of the test or assessment have been developed by the same developers, it is possible to estimate the reliability between the versions.

A group of people complete both versions of the test or assessment and a correlation coefficient (Pearson Product Moment) is calculated. This correlation provides an index of alternate form reliability. In other words, people who score high on one version also score high on the alternate version, and low scorers score low on both. When an assessment has high alternate form reliability, it means we can be confident that a person would achieve a similar score irrespective of which version of the assessment was used.

A development aim of Wave Styles was that this form of reliability should be as high as possible.

Internal Consistency Reliability

This form of reliability is an index of how the items in a test (or a personality scale) relate to one another. It carries the practical advantage that it can be computed without the need for a retest or an alternative form, but there are some drawbacks.

For self-report questionnaires it is important that internal consistency reliability is satisfactorily high without being artificially inflated. For instance, a personality scale with repetitive item content will have high internal consistency reliability estimates, but lack breadth of measurement. This narrowness of coverage of the content domain in a questionnaire may fall well short of what scales should be measuring and is likely to impact on the empirical validity of the test in forecasting effectiveness on independently assessed criteria. In the development of Wave Styles this problem of 'Bloated Specifics' was avoided by drawing on three distinct facet constructs for each Wave dimension. The selection of these facets was primarily based on their concurrent validity with internal consistency reliability being of secondary concern. This approach also ensured good construct separation between the dimensions measured by the Wave Styles questionnaires.

A development aim of Wave Styles was to have internal consistency reliability estimates of the Wave dimensions between .60 and .90. In essence, this form of reliability was seen by the authors as a measure of the breadth or narrowness of the scale. The results for alternate form and test-retest give a better indication of the reliability of Wave Styles questionnaires.

Sources of Error affecting Reliability

Assessment scores can contain errors of measurement from a number of sources, for example:

- **Questionnaire Design** – questions with negative phrasing or asking more than one question in an item tend to increase measurement error
- **Individual** - mood, temperament, motivation, well-being
- **Environment** - noise, temperature, presence of others
- **Administration** - degree and consistency of standardization
- **Scoring** – the accuracy of the scoring key and scoring process

Maximizing Reliability

The primary development aim of Wave was to develop a high validity instrument to predict performance outcomes at work. For an instrument to be highly valid it also needs to be reliable. To achieve this aim, specific steps were taken to ensure high reliability:

1. Negative phrased and keyed items were avoided. Negative items had less reliability in early trials
2. Questionnaire Instructions were standardized
3. A Normative Development Trial preceded the Full Standardization Trial that used the new Ra-Ra (Rate-Rank) response format
4. Questions were balanced in blocks of six to standardize the number of comparisons across different dimensions
5. Items were selected for blocks based on their mean endorsement value from the normative trial to ensure the items within a block were equally attractive to respondents
6. Items were written and reviewed against clear criteria (see Construction chapter)
7. Items were not included if they had low reliability as well as validity

Standard Error of Measurement

When test or assessment users receive a test score they make inferences, communicate and/or make decisions based on the test score. However, the observed score is subject to error and so to be in a better position to use the test score, it is important for a test user to have an appreciation of the band of error around the score and know how likely it is to contain the individual's true score. To do this the Standard Error of Measurement is computed (SEm).

Formula

The Standard Error of Measurement (SEm) equals the Standard Deviation of a group multiplied by the square root of one minus the reliability coefficient.

$$SEm = SD\sqrt{(1 - r_t)}$$

Where:

SEm = Standard Error of Measurement

SD = Standard Deviation of the sample that the reliability coefficient was calculated from

r_t = the reliability coefficient (test-retest, alternate form, internal consistency)

If we take the average alternate reliability of the Wave Styles scales, which was r= .86 in the standardization trials, and want to calculate the Standard Error of Measurement for a sten score, then:

SD Sten Score = 2
Alternate Form Reliability = .86

$$SEm = 2\sqrt{(1 - 0.86)}$$
$$= 2 \times .37$$
$$= .74$$

A band of 1 SEm (i.e., .74 stens) either side of an individual's score results in a 68% probability that this band contains the true score for the individual. For instance, with a sten score of 6, we are confident that 68% of the time the person's true score will be between 5.26 and 6.78 - or 1 SEm to either side of the observed score.

By placing a band of 2 SEms (i.e., 2 x .74 stens or 1.48 stens) either side of the observed score gives a 96% probability that this band contains the true score for this individual.

+/- 1SEm – 68% Probability
+/- 2SEm – 96% Probability

In practice, stens are rounded to the nearest whole number between 1 and 10.

19.1 Reliability Overview

Alternate Form Reliability

The alternate form reliability of Saville Consulting Wave Professional Styles is based on two versions of Professional Styles; Invited Access (IA) and Supervised Access (SA). Tables 19.1 and 19.6 show the means and standard deviations for both versions of the Professional Styles questionnaire, along with their Normative, Ipsative and Total Score alternate form reliability coefficients (r_t).

Alternate form reliabilities of .70 and above are regarded by the authors as acceptable levels of reliability for a trait measure although higher levels than this are desirable. At the dimension level, the median reliability of the Total Score (combined Normative and Ipsative) scales was .87 and the minimum reliability estimate for any dimension was .78. At the section level, the median reliability of the Total Score scales was .92 and the minimum reliability estimate for any section was .86.

Normative and Ipsative scores of the questionnaire also had good alternate form reliabilities with median reliabilities of .86 (dimension level) and .91 (section level) for Normative and .83 (dimension level) and .89 (section level) for Ipsative and minimum reliability estimates of .78 (dimension level) and .87 (section level) for Normative and .72 (dimension level) and .82 (section level) for Ipsative.

Construct independence between the scales is demonstrated by the 'Other Highest Correlation' and 'Other Dimension/Section' columns, which show the highest correlation (other than that with the parallel version of the dimension/section of the same name) of one dimension/section in one version with the dimension/section in another (the off diagonals in a correlation matrix).

As can be seen from Tables 19.1 and 19.6, these correlations are substantially lower than the Alternate Form correlations between same scales, demonstrating good construct independence of the dimensions/sections at the individual dimension/section level. The highest correlation between different dimensions across the two versions is between Organized (SA) and Reliable (IA) with a correlation between the scales of .60. However, the respective alternate form reliability estimates of the two dimensions are .88 for Organized and .91 for Reliable. The highest correlation between the different sections across the two versions is between Driven (SA) and Assertive (IA), with a correlation between the scales of .60. However, the respective alternate form reliability estimates of the two sections are .93 in both cases.

Internal Consistency Reliability

Tables 19.2 & 19.3 provide the internal consistency (Cronbach's Alpha) of the 36 dimensions of Professional Styles for Invited Access (IA) (Table 19.2) and Supervised Access (SA) (Table 19.3). The dimensions of Wave Professional Styles were designed to have internal consistency estimates ranging from .60 to a maximum of .90. The median internal consistency (across the 72 dimensions across the two versions) is in the center of this desired range. Only one scale fell outside this – Insightful on Invited Access with an internal consistency of .58. However, Insightful has highly acceptable alternate form reliability and test-retest reliability estimates which are the fundamental reliability measures for Wave Styles. Tables 19.7 and 19.8 provide the internal consistency of the 12 sections of Professional Styles for Invited Access (IA) (Table 19.7) and Supervised Access (SA) (Table 19.8). No section fell outside the acceptable range of reliability estimates (.60 - .90).

Test-Retest Reliability

Tables 19.4 and 19.9 provide the test-retest reliability of Saville Consulting Wave Professional Styles administered at an eighteen month interval. Test-retest reliabilities of .70 and above are acceptable levels of reliability. The 36 dimensions of Wave Professional Styles demonstrate acceptable test-retest reliabilities with coefficients ranging from .58 (Principled) to .85 (Activity Oriented) and a median reliability coefficient of .74. The 12 sections of Wave Professional Styles demonstrate high test-retest reliabilities with coefficients ranging from .76 (Structured) to .86 (Sociable) and a median reliability of .79.

19.2 Reliability Tables

Table 19.1 Professional Styles Alternate Form Reliability - Invited Access (IA) vs. Supervised Access (SA) - Normative, Ipsative and 36 combined Normative-Ipsative dimensions. *(N=1,153)*

Dimension	(IA) Mean	(IA) SD	(SA) Mean	(SA) SD	Norms r_t	Ips r_t	Total Score SEm (Stens)	Total Score r_t	Other Highest Correlation	Other Dimension
Analytical	62.87	9.78	62.82	9.95	.85	.79	.80	**.84**	.50	Abstract
Factual	65.89	8.86	64.93	9.56	.79	.79	.87	**.81**	.38	Analytical
Rational	53.82	13.09	52.80	13.55	.91	.88	.57	**.92**	.49	Analytical
Learning Oriented	62.43	11.93	64.42	10.99	.86	.84	.72	**.87**	.51	Abstract
Practically Minded	67.38	9.68	67.57	9.49	.85	.83	.75	**.86**	.26	Rational
Insightful	65.42	8.73	66.00	8.87	.82	.72	.92	**.79**	.45	Strategic
Inventive	52.83	13.41	52.36	13.32	.91	.87	.60	**.91**	.54	Strategic
Abstract	58.21	11.41	55.69	12.04	.85	.77	.82	**.83**	.51	Learning Oriented
Strategic	56.13	11.20	53.61	12.08	.84	.79	.80	**.84**	.54	Inventive
Interactive	53.75	13.63	52.31	12.50	.90	.85	.63	**.90**	.58	Engaging
Engaging	67.11	10.88	67.07	11.49	.87	.83	.72	**.87**	.58	Interactive
Self-promoting	45.80	12.32	43.85	12.34	.89	.84	.66	**.89**	.42	Interactive
Convincing	55.80	10.31	51.20	10.63	.85	.78	.80	**.84**	.54	Challenging
Articulate	57.59	12.25	56.56	11.89	.91	.86	.60	**.91**	.39	Interactive
Challenging	51.39	11.72	49.96	12.09	.86	.81	.75	**.86**	.54	Convincing
Purposeful	54.60	10.96	54.91	11.03	.87	.80	.72	**.87**	.53	Directing
Directing	58.81	12.86	56.22	12.65	.89	.84	.66	**.89**	.54	Empowering
Empowering	59.44	12.67	59.31	13.17	.90	.85	.66	**.89**	.54	Directing
Self-assured	59.70	11.06	61.62	12.28	.86	.78	.77	**.85**	.37	Positive
Composed	53.40	13.50	52.16	14.09	.90	.84	.66	**.89**	.48	Change Oriented
Resolving	58.65	11.90	55.50	13.05	.88	.84	.69	**.88**	.47	Attentive
Positive	65.93	10.92	64.91	10.17	.85	.81	.77	**.85**	.38	Change Oriented
Change Oriented	61.75	11.34	61.49	11.59	.85	.82	.75	**.86**	.48	Composed
Receptive	60.20	9.76	58.63	9.50	.81	.73	.94	**.78**	.24	Involving
Attentive	65.43	11.08	63.77	11.68	.83	.85	.75	**.86**	.52	Accepting
Involving	65.55	9.10	63.02	9.51	.79	.81	.87	**.81**	.51	Accepting
Accepting	63.60	12.01	65.51	10.84	.78	.82	.87	**.81**	.52	Attentive
Reliable	64.95	13.41	66.00	12.47	.89	.89	.60	**.91**	.60	Organized
Meticulous	64.48	13.69	64.65	13.04	.87	.87	.66	**.89**	.50	Organized
Conforming	53.48	14.52	54.12	14.65	.89	.90	.60	**.91**	.48	Reliable
Organized	64.94	11.40	65.44	11.43	.86	.88	.69	**.88**	.60	Reliable
Principled	71.84	9.66	74.38	9.77	.81	.77	.87	**.81**	.34	Accepting
Activity Oriented	65.35	10.63	64.47	11.07	.90	.86	.66	**.89**	.30	Reliable
Dynamic	57.83	10.53	57.09	10.92	.87	.81	.72	**.87**	.48	Directing
Enterprising	53.24	15.15	52.80	15.00	.93	.89	.53	**.93**	.53	Striving
Striving	61.82	10.50	62.10	10.63	.86	.79	.77	**.85**	.53	Enterprising
Mean	60.04	11.55	59.42	11.65	.86	.83	.73	**.86**	.48	
Median	59.95	11.37	60.40	11.64	.86	.83	.72	**.87**	.51	
Min	45.80	8.73	43.85	8.87	.78	.72	.53	**.78**	.24	
Max	71.84	15.15	74.38	15.00	.93	.90	.94	**.93**	.60	

Table 19.2 Internal Consistency Reliability of Professional Styles - Invited Access (dimension level). *(N=1,153)*

Dimension	Mean	SD	SEm (sten)	r_t
Analytical	62.87	9.78	1.23	.62
Factual	65.89	8.86	1.27	.60
Rational	53.82	13.09	1.04	.73
Learning Oriented	62.43	11.93	1.01	.75
Practically Minded	67.38	9.68	1.12	.68
Insightful	65.42	8.73	1.29	.58
Inventive	52.83	13.41	.75	.86
Abstract	58.21	11.41	.99	.76
Strategic	56.13	11.20	1.03	.74
Interactive	53.75	13.63	.96	.77
Engaging	67.11	10.88	1.02	.74
Self-promoting	45.80	12.32	.94	.78
Convincing	55.80	10.31	1.11	.69
Articulate	57.59	12.25	1.03	.74
Challenging	51.39	11.72	1.01	.75
Purposeful	54.60	10.96	1.23	.62
Directing	58.81	12.86	.89	.80
Empowering	59.44	12.67	.79	.84
Self-assured	59.70	11.06	1.17	.66
Composed	53.40	13.50	.93	.79
Resolving	58.65	11.90	.98	.76
Positive	65.93	10.92	.97	.76
Change Oriented	61.75	11.34	.93	.78
Receptive	60.20	9.76	1.22	.63
Attentive	65.43	11.08	.96	.77
Involving	65.55	9.10	1.10	.70
Accepting	63.60	12.01	.91	.79
Reliable	64.95	13.41	.88	.81
Meticulous	64.48	13.69	.74	.86
Conforming	53.48	14.52	.78	.85
Organized	64.94	11.40	.97	.70
Principled	71.84	9.66	1.12	.69
Activity Oriented	65.35	10.63	.98	.76
Dynamic	57.83	10.53	1.14	.67
Enterprising	53.24	15.15	.85	.82
Striving	61.82	10.50	1.18	.65
Mean	60.04	11.55	1.01	.74
Median	59.95	11.37	1.00	.76
Min	45.80	8.73	.74	.58
Max	71.84	15.15	1.29	.86

Table 19.3 Internal Consistency Reliability of Professional Styles - Supervised Access (dimension level). *(N=1,153)*

Dimension	Mean	SD	SEm (sten)	r_t
Analytical	62.82	9.95	1.17	.66
Factual	64.93	9.56	1.15	.67
Rational	52.80	13.55	.97	.76
Learning Oriented	64.42	10.99	1.04	.73
Practically Minded	67.57	9.49	1.08	.71
Insightful	66.00	8.87	1.13	.68
Inventive	52.36	13.32	.74	.86
Abstract	55.69	12.04	.90	.80
Strategic	53.61	12.08	.86	.82
Interactive	52.31	12.50	1.00	.75
Engaging	67.07	11.49	.87	.81
Self-promoting	43.85	12.34	.90	.80
Convincing	51.20	10.63	1.07	.71
Articulate	56.56	11.89	1.05	.72
Challenging	49.96	12.09	.94	.78
Purposeful	54.91	11.03	1.19	.65
Directing	56.22	12.65	.89	.80
Empowering	59.31	13.17	.74	.86
Self-assured	61.62	12.28	.96	.77
Composed	52.16	14.09	.83	.83
Resolving	55.50	13.05	.81	.84
Positive	64.91	10.17	1.06	.72
Change Oriented	61.49	11.59	.83	.83
Receptive	58.63	9.50	1.17	.66
Attentive	63.77	11.68	.82	.83
Involving	63.02	9.51	1.04	.73
Accepting	65.51	10.84	1.02	.74
Reliable	66.00	12.47	.89	.80
Meticulous	64.65	13.04	.71	.87
Conforming	54.12	14.65	.71	.87
Organized	65.44	11.43	.90	.80
Principled	74.38	9.77	.92	.79
Activity Oriented	64.47	11.07	.86	.82
Dynamic	57.09	10.92	1.07	.71
Enterprising	52.80	15.00	.83	.83
Striving	62.10	10.63	1.14	.67
Mean	59.49	11.62	.96	.77
Median	61.49	11.59	.94	.78
Min	43.85	8.87	.71	.65
Max	74.38	15.00	1.19	.87

Table 19.4 Test-Retest Reliability of Professional Styles - Invited Access (dimension level). *(N=100)*

Dimension	$Mean_{t1}$	SD_{t1}	$Mean_{t2}$	SD_{t2}	SEm (sten)	r_t
Analytical	61.90	9.51	61.26	8.59	1.17	**.66**
Factual	67.12	9.12	66.44	9.96	1.11	**.69**
Rational	51.75	14.70	53.52	14.29	.85	**.82**
Learning Oriented	66.81	10.60	64.73	11.84	.88	**.81**
Practically Minded	66.50	11.78	66.50	11.12	.87	**.81**
Insightful	64.96	8.98	65.62	8.26	1.19	**.65**
Inventive	54.44	13.31	55.88	13.81	.97	**.76**
Abstract	58.37	11.39	56.98	12.36	1.07	**.71**
Strategic	57.06	10.74	57.58	12.19	1.07	**.71**
Interactive	54.30	13.23	52.47	14.39	.82	**.83**
Engaging	64.83	11.95	63.34	12.11	.85	**.82**
Self-promoting	46.58	12.61	45.99	11.82	1.03	**.74**
Convincing	53.52	10.85	53.21	10.17	1.02	**.74**
Articulate	57.67	12.76	56.62	13.03	.87	**.81**
Challenging	49.83	12.40	49.89	12.42	1.11	**.69**
Purposeful	53.30	11.87	53.74	12.22	1.09	**.70**
Directing	57.39	14.63	58.52	15.31	.83	**.83**
Empowering	58.86	11.69	60.94	12.12	1.20	**.64**
Self-assured	63.28	10.82	62.29	10.43	0.94	**.78**
Composed	51.66	13.76	52.45	14.01	1.03	**.73**
Resolving	58.15	12.27	56.27	12.57	.96	**.77**
Positive	64.09	11.43	61.58	12.98	.99	**.75**
Change Oriented	58.48	11.78	58.57	12.78	.92	**.79**
Receptive	60.58	10.84	58.43	10.79	1.07	**.72**
Attentive	64.80	12.74	63.68	13.14	.82	**.83**
Involving	63.15	9.69	62.80	9.92	1.12	**.69**
Accepting	61.20	10.48	59.56	10.57	1.06	**.72**
Reliable	66.77	11.49	66.58	11.90	.93	**.78**
Meticulous	66.39	12.45	66.88	12.06	1.02	**.74**
Conforming	55.36	13.58	55.87	14.89	.96	**.77**
Organized	67.74	10.48	67.68	10.22	1.09	**.70**
Principled	71.49	9.24	70.13	10.11	1.30	**.58**
Activity Oriented	66.51	11.58	66.15	12.19	.77	**.85**
Dynamic	59.70	10.56	60.75	11.08	1.06	**.72**
Enterprising	50.52	14.52	49.92	14.68	1.01	**.75**
Striving	62.77	11.22	61.41	11.46	.93	**.78**
Mean	59.29	11.88	58.87	12.26	1.00	**.75**
Median	58.67	11.73	58.55	12.19	1.02	**.74**
Min	46.58	9.24	45.99	9.92	.77	**.58**
Max	71.49	14.63	70.13	15.31	1.30	**.85**

Note: Subjects completed Wave twice at an 18 month interval.

Table 19.5 Internal Consistency Reliability of Professional Styles with a US sample (dimension level). *(N=2,102)*

Dimension	Mean	SD	SEm (sten)	r_t
Analytical	67.46	8.70	1.28	.59
Factual	66.91	8.52	1.28	.59
Rational	55.44	11.88	1.09	.71
Learning Oriented	66.26	9.90	1.07	.71
Practically Minded	68.52	8.21	1.21	.63
Insightful	69.04	7.72	1.39	.52
Inventive	57.02	11.37	.89	.80
Abstract	59.97	10.48	.99	.75
Strategic	65.23	10.25	1.05	.72
Interactive	53.88	12.08	.99	.76
Engaging	66.89	10.53	.94	.78
Self-promoting	41.66	11.17	.94	.78
Convincing	56.05	9.74	1.19	.65
Articulate	64.56	10.84	1.05	.73
Challenging	47.60	11.03	1.03	.73
Purposeful	57.92	10.33	1.20	.64
Directing	66.70	10.66	1.02	.74
Empowering	65.54	11.31	.84	.82
Self-assured	65.58	8.75	1.33	.56
Composed	55.32	11.96	1.04	.73
Resolving	57.52	10.82	.97	.76
Positive	66.86	9.73	1.03	.74
Change Oriented	64.58	10.59	1.01	.74
Receptive	61.85	9.20	1.26	.60
Attentive	64.23	10.27	1.00	.75
Involving	65.88	9.81	.99	.75
Accepting	64.71	9.97	1.06	.72
Reliable	67.63	11.41	.93	.78
Meticulous	66.08	11.49	.90	.80
Conforming	52.37	13.14	.86	.81
Organized	68.61	10.36	1.00	.75
Principled	78.54	7.45	1.10	.70
Activity Oriented	66.55	9.99	.95	.77
Dynamic	64.31	9.61	1.22	.63
Enterprising	59.80	15.53	.79	.84
Striving	71.64	8.97	1.24	.62
Mean	62.74	10.38	1.06	.71
Median	64.97	10.34	1.03	.74
Min	41.66	7.45	.79	.52
Max	78.54	15.53	1.39	.84

Table 19.6 Professional Styles Alternate Form Reliability - Invited Access (IA) vs. Supervised Access (SA) - Normative, Ipsative and 36 combined Normative-Ipsative sections *(N=1,153)*

Section	(IA) Mean	(IA) SD	(SA) Mean	(SA) SD	Norms. r_t	Ips. r_t	Total Score SEm (stens)	Total Score r_t	Other Highest Correla-tion	Other Section
Evaluative	182.58	24.02	180.56	25.35	.91	.90	.57	.92	.43	Investigative
Investigative	195.24	19.14	197.99	18.62	.87	.83	.75	.86	.46	Evaluative
Imaginative	167.18	28.53	161.67	31.11	.92	.88	.57	.92	.41	Investigative
Sociable	166.66	29.14	163.23	28.71	.93	.91	.53	.93	.37	Impactful
Impactful	164.78	25.69	157.73	26.06	.91	.88	.60	.91	.47	Assertive
Assertive	172.84	29.08	170.44	29.53	.93	.90	.53	.93	.59	Driven
Resilient	171.75	24.02	169.28	25.08	.90	.83	.66	.89	.39	Assertive
Flexible	187.88	21.80	185.03	20.98	.89	.82	.72	.87	.41	Resilient
Supportive	194.57	26.65	192.31	26.38	.87	.90	.63	.90	.16	Resilient
Conscientious	182.91	33.26	184.77	32.27	.92	.94	.49	.94	.52	Structured
Structured	202.13	20.91	204.29	21.37	.89	.88	.66	.89	.55	Conscientious
Driven	172.89	29.05	171.99	29.39	.93	.90	.53	.93	.60	Assertive
Mean	180.12	25.94	178.27	26.24	.91	.88	.60	.91	.45	
Median	177.74	26.17	176.27	26.22	.91	.89	.58	.92	.45	
Min	164.78	19.14	157.73	18.62	.87	.82	.49	.86	.16	
Max	202.13	33.26	204.29	32.27	.93	.94	.75	.94	.60	

Table 19.7 Internal Consistency Reliability of Professional Styles - Invited Access (section level). *(N=1,153)*

Section	Mean	SD	SEm (sten)	r_t
Evaluative	182.58	24.02	.94	.78
Investigative	195.24	19.14	1.15	.67
Imaginative	167.18	28.53	.72	.87
Sociable	166.66	29.14	.77	.85
Impactful	164.78	25.69	.87	.81
Assertive	172.84	29.08	.75	.86
Resilient	171.75	24.02	1.00	.75
Flexible	187.88	21.80	.98	.76
Supportive	194.57	26.65	.72	.87
Conscientious	182.91	33.26	.66	.89
Structured	202.13	20.91	1.00	.75
Driven	172.89	29.05	.80	.84
Mean	180.12	25.94	.86	.81
Median	177.74	26.17	.84	.83
Min	164.78	19.14	.66	.67
Max	202.13	33.26	1.15	.89

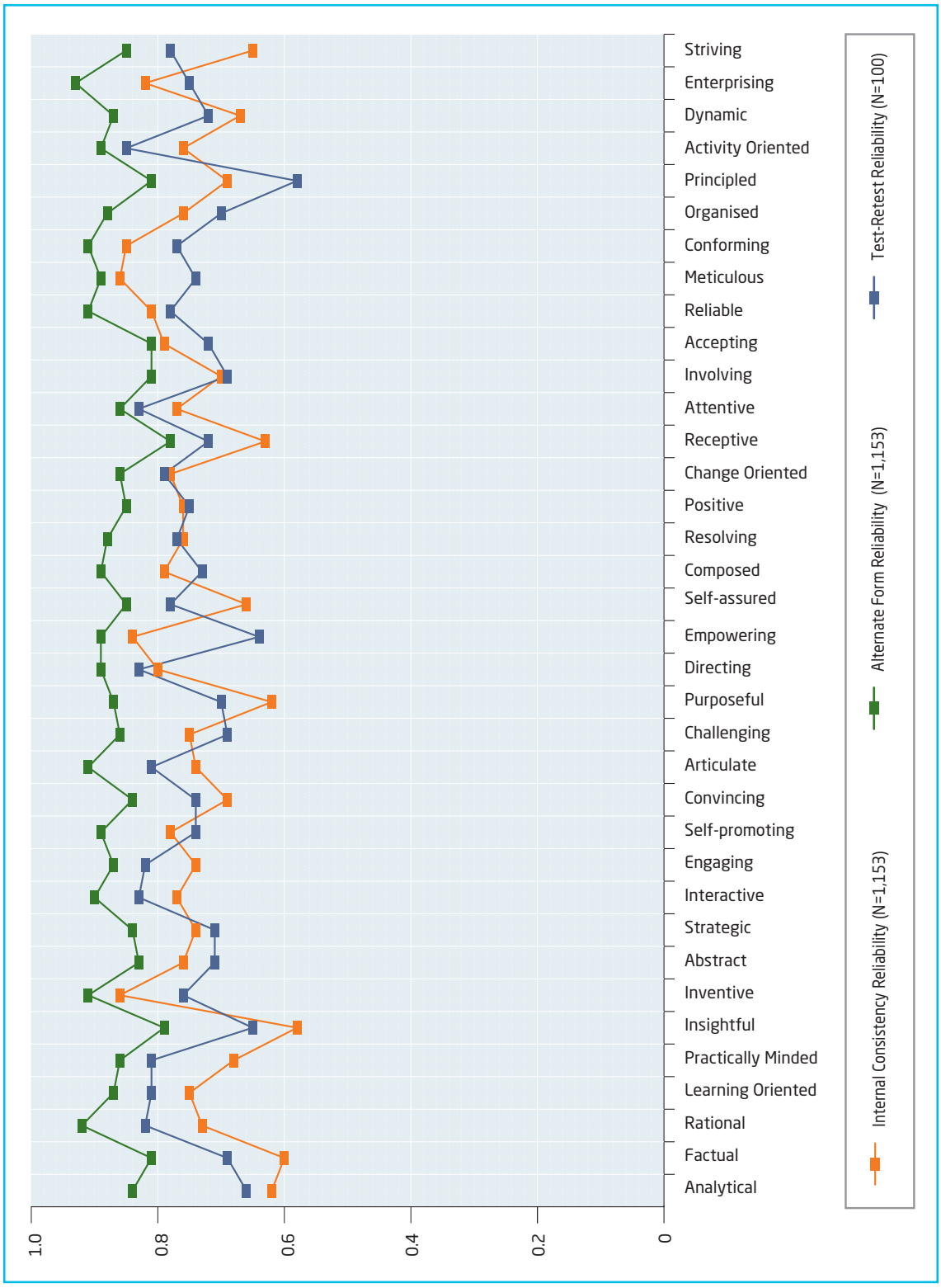
Table 19.8 Internal Consistency Reliability of Professional Styles - Supervised Access (section level). *(N=1,153)*

Section	Mean	SD	SEm (sten)	r_t
Evaluative	180.56	25.35	.87	.81
Investigative	197.99	18.62	1.10	.70
Imaginative	161.67	31.11	.63	.90
Sociable	163.23	28.71	.75	.86
Impactful	157.73	26.06	.85	.82
Assertive	170.44	29.53	.72	.87
Resilient	169.28	25.08	.94	.78
Flexible	185.03	20.98	.98	.76
Supportive	192.31	26.38	.72	.87
Conscientious	184.77	32.27	.63	.90
Structured	204.29	21.37	.92	.79
Driven	171.99	29.39	.77	.85
Mean	178.27	26.24	.82	.83
Median	176.28	26.22	.81	.84
Min	157.73	18.62	.63	.70
Max	204.29	32.27	1.10	.90

Table 19.9 Test-Retest Reliability of Professional Styles Normative - Invited Access (section level). *(N=100)*

Section	Mean $_{t1}$	SD $_{t1}$	Mean $_{t2}$	SD $_{t2}$	SEm (Sten)	r_t
Evaluative	179.31	25.09	181.22	24.77	.94	.78
Investigative	198.84	21.80	196.84	20.54	.89	.80
Imaginative	169.53	28.45	170.44	29.98	.97	.77
Sociable	167.21	32.28	161.79	31.92	.76	.86
Impactful	161.87	25.86	159.71	26.68	.91	.79
Assertive	169.60	27.46	173.20	29.50	.96	.77
Resilient	173.86	24.34	171.01	23.50	.89	.80
Flexible	182.92	21.53	178.58	24.20	.96	.77
Supportive	189.35	25.83	186.04	27.77	.81	.84
Conscientious	187.65	29.86	189.33	28.98	.83	.83
Structured	204.83	20.69	203.95	21.76	.99	.76
Driven	171.30	29.41	172.08	30.45	.92	.79
Mean	179.69	26.05	178.68	26.67	.90	.80
Median	176.58	25.84	175.89	27.22	.92	.79
Min	161.87	20.69	159.71	20.54	.76	.76
Max	204.83	32.28	203.95	31.92	.99	.86

Graph 19.1 Internal Consistency, Test-Retest Alternate Form Reliability of Professional Styles



19.3 Reliability of Facet Scales

Wave Professional Styles is composed of 108 different two item facet scales. While the individual 108 facet scales are not individually plotted on a profile, a Wave user's attention is drawn to facet ranges, where there is a difference of three of more sten scores between the three facet scales within each dimension. Internal Consistency is not an ideal method of reliability estimation for the facet scales of Wave Professional Styles as the two items of each facet are designed to measure different content (i.e., one motive and one talent item). Alternate Form Reliabilities range from .50 to .90 for two item facet scales (Ra-Ra) with median of .78 (N=1,153). This compares with Alternate Form Reliabilities of Wave Professional Styles six item dimension scales of r= .86 (composed of three facet scales – six items).

36 of the facet concepts of Wave Professional Styles have also been subject to test-retest in two item facet scales in Wave Focus Styles for over six months and the figures ranged from .58 to .84 for two item facet scales with median of .72 (N=214).

19.4 Summary of Reliability

No measure of human traits has perfect reliability, yet good reliability of measurement is an important property of any assessment. This chapter highlights in particular, given the design of Wave Professional Styles, the importance of alternate form reliability as an appropriate method for the estimation of reliability.

The method of development of Wave Professional Styles targeted scales to have internal consistencies (Cronbach's Alpha) between .60 and .90. The reason for targeting this level of internal consistency is that internal consistency provides a measure of scales' breadth of content measurement.

Wave Professional Styles was designed by selecting facets/items with varied content within each of the dimensions. The internal consistency of the dimensions (Cronbach's alpha) ranged at standardization from .58 to .86. Graph 19.1 indicates the scales at Factual and Insightful have internal consistency reliabilities of less than .60, however both of these scales display good alternate form reliabilities: Factual .81; Insightful .79. This suggests that despite their breadth of measurement these dimensions are reliable and reproducible. Information on the validity of these dimensions can be found in the Validity chapter.

Alternate form median at standardization was .87 (no corrections applied) for the dimensions and the reliabilities ranged from .78 to .93. A Test-Retest was conducted with a month's interval between original test and retest during development and achieved a median of .80 for the dimensions.

Alternate form also provides a method of investigating construct separation and the Wave Professional Styles dimensions provide clear evidence supporting this separation.

19.5 Further Reference Material

Further information can be found in the Norms, Fairness chapters and appendices of this handbook.