

PSYCHOMETRIC PROPERTIES OF CLEVRY'S ABILITY TESTS

clevry

1. Utopia (Expert)

Utopia is a high-level aptitude test series designed for top level managers and graduates. Utopia comprises verbal, numerical, and abstract critical-reasoning assessments. It provides rigorous psychometric assessments of ability with outstanding candidate friendliness. It is delivered by our Clevery online psychometrics platform.

Both the verbal and numerical tests relate to reference material in the form of Media Packs for an environmental magazine ('Utopia!'). Both have been designed to be interesting, topical, upbeat, and easy to relate to, with no prior knowledge of environmental issues or the media required.

Reliability

Verbal test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Graduate applicants for an accelerated career programme*	400	29.87	4.98	0.77	2.38
Graduate applicants for an internship in a financial services organisation	40	29.85	4.35	0.73	2.26
Graduate applicants to an International Bank	204	29.09	4.75	0.75	2.38

Numerical test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Graduate applicants for an accelerated career programme*	400	18.70	6.04	0.85	2.33
Graduate applicants for an internship in a financial services organisation	40	25.28	3.83	0.80	1.71
Graduate applicants to an International Bank	205	22.68	5.23	0.86	1.96

*These results are based on the samples used in the development of the Utopia series. The statistics reported are derived from those items which were selected for inclusion in the final forms of the instruments. Those items which were removed are excluded from the computation of these means, standard deviations, reliability coefficients, and standard errors of measurement.

Validity

Face validity

One of the key aims in the development of the Utopia series was to create a battery of tests with high face validity. Hence the themed format of the instruments, and the realistic time limits.

Much anecdotal evidence suggests that such an approach was successful. Positive feedback has been received from candidates, supervisors, line managers, senior managers, and personnel specialists within organisations which have used Utopia.

Detailed data on perceptions from undergraduates can be provided on request.

Content validity

The content validity of the Utopia series is assured on the basis that the nature, structure and content of the instruments were all determined on the basis of rigorous job analysis at the time the instruments were developed.

All items within the Utopia series were carefully developed to reflect tasks, activities and cognitive processes that differentiate between job incumbents in target populations.

Construct validity

Evidence of the construct validity of Utopia was established when candidates' scores on the Utopia verbal & numerical tests were found to correlate significantly with the verbal and numerical tests from SHL's *Managerial & Graduate Item Bank*.

	Utopia Verbal (N=64)	Utopia Numerical (N=59)
MGIB Verbal (VGM-3)	0.60*	0.34*
MGIB Numerical (NGM-3)	0.40*	0.69*

* $p < 0.01$

These results are based on a sample of 1999 undergraduates. Details regarding this study can be provided on request.

As part of the ongoing research into the Utopia series, Clevery are currently developing further research programmes aimed at assessing the validity of the Utopia Series.

Criterion-related validity

In 2003, a study was completed that demonstrated the criterion related validity of Utopia in a graduate assessment centre context. The results demonstrated the following significant positive relationships between candidates' Utopia test scores and assessment centre performance.

1. Correlation of Utopia tests with Overall Grade

The study confirmed the prediction that each of the Utopia tests (Verbal and Numerical) would correlate positively with the **overall** performance rating (the overall grade given, which indicated the overall spread of grades across the Capabilities and exercises).

	Utopia Verbal Test	Utopia Numerical Test
Overall grade	0.32*	0.22*

* $p < 0.05$

2. Correlation of Utopia tests with Overview of Capability on each exercise

The study partially confirmed the prediction that the Utopia aptitude tests (Verbal and Numerical) would correlate with overall performance on each of the five assessment centre exercises. Positive significant relationships were found between the Utopia tests and four out of the six assessment exercises.

	Utopia Verbal Test	Utopia Numerical Test
Inspires people	0.32*	0.25*
Focuses action	0.28*	0.28*
Seeks improvement	-0.16	-0.03
Drives results	0.22*	0.09
Takes ownership	0.25*	0.11

* $p < 0.05$

3. Correlation of Utopia tests with overall pass/fail decision

The study partially confirmed the prediction that the Utopia aptitude tests (Verbal and Numerical) would correlate with the pass/fail decision for the assessment centre.

	Utopia Verbal Test	Utopia Numerical Test
Overall Decision	0.14	0.26*

* $p < 0.05$

Further detail can be provided on request.

2. B2C (Enhanced)

B2C is a mid-level aptitude test series for customer-facing, junior management, and administrative roles. B2C comprises verbal, numerical checking, and abstract assessments. It provides rigorous psychometric assessments of ability with outstanding candidate friendliness. The verbal, numerical and checking tests are delivered by our Clevery online psychometrics platform with paper & pencil equivalents also available.

All tests relate to a fictitious music store and are suited to situations in which employees are required to perform a range of tasks. All have been designed to be interesting, upbeat, and easy to relate to, with no prior knowledge of music stores required.

Reliability

Verbal test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Composite group of applicants for Administrative positions	250	25.41	5.37	0.77	2.6

Numerical test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Composite group of applicants for Administrative positions	142	15.52	4.52	0.79	2.1

Abstract test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Composite group of candidates on Clevery	305	8.72	3.02	0.65	0.05

Checking test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Composite group of applicants for Administrative positions	128	19.88	6.80	0.87	2.5

Validity

Face validity

One of the key aims in the development of the B2C series was to create a battery of tests with high face validity. Hence the themed format of the instruments, and the realistic time limits.

Much anecdotal evidence has suggested that such an approach was successful. Positive feedback has been received from candidates, supervisors, line managers, senior managers and personnel specialists within organisations which have used the Business Administration Series. To maintain this face validity, we have up-dated the content and look of the tests so that candidates and users continue to view them positively, and as relevant to

modern work environments.

Content validity

The content validity of the B2C series instruments are assured on the basis that the nature, structure and content of the instruments were all determined on the basis of rigorous job analysis at the time when the instruments were originally developed.

Criterion-related validity

Predictive validity between verbal, numerical, and checking test scores at the recruitment stage, and performance data 18 months into employment for Dispensing assistants, Receptionists and Laboratory technicians were taken.

The RTMs asked questions about performance in the following areas:

- Interpersonal orientation
- Concern for order, quality, and accuracy
- Personal maturity
- Teamwork
- Impact and influence
- Organisational skills
- Information handling

It was predicted that each of the aptitude tests (Verbal, Numerical and Checking) would correlate positively with the **overall** performance rating (this was the average of the seven rating scales).

	B2C Verbal test	B2C Numerical test	B2C Checking test
Overall performance rating	0.40*	0.32*	0.24

* $p < 0.05$

The results indicated that all three tests correlated moderately with the overall performance rating. The validity coefficients were in the hypothesised range of 0.20-0.40. The Verbal Test showed the strongest correlation (0.40), followed by the Numerical Test (0.32). The Checking test did correlate with the index of overall performance but with a weaker validity coefficient of 0.24. This may indicate that general intellectual capability (as sampled by the Verbal and Numerical Tests) is more important in these particular roles than the attention to detail measured by the Checking Test. It was predicted that the aptitude tests would correlate with the ratings scales which reflect *Concern for order, Quality and accuracy, Organisational skills, and Information handling*.

The three aptitude tests did correlate positively with these task-related/cognitive performance rating scales. The strongest correlations were with the Information handling scale which correlated with the Verbal Test at 0.45, the Numerical Test at 0.41, and the Checking Test at 0.24. This confirms that the aptitude tests are particularly useful in predicting the task-related or cognitive aspects of the job roles.

This predictive validation study has revealed healthy correlations between each of the aptitude tests and the criterion measures of job performance, illustrating that the tests are not only reliable but also valid predictors of job success.

3. CWS (Essential)

CWS, or the Criterion Workforce Series, are tests which simulate the demands of production, manufacturing and engineering roles. CWS comprises verbal, numerical, and mechanical reasoning. It provides rigorous psychometric assessments of ability, with outstanding candidate friendliness. The verbal, numerical and mechanical reasoning tests are delivered by our Clevery online psychometrics platform with paper & pencil equivalents also available. The fault detection and SPC1 plotting data test are available in paper and pencil format.

Reliability

Verbal test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Apprentice applicants to the steel industry*	93	18.95	4.86	0.89	1.61
Hourly paid operative applicants to a car manufacturer*	486	15.91	6.02	0.85	2.33
Apprentice incumbents employed by a car manufacturer	98	17.22	4.60	0.80	2.06
Craft and operator incumbents employed in the steel industry	148	17.10	4.79	0.81	2.09

Numerical test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Apprentice applicants to the steel industry*	93	21.76	7.69	0.91	2.31
Hourly paid operative applicants to a car manufacturer*	486	13.34	5.53	0.91	1.66
Apprentice incumbents employed by a car manufacturer	99	19.56	5.86	0.86	2.19
Craft and operator incumbents employed in the steel industry	148	17.24	5.64	0.86	2.11

Mechanical test

Sample	N	Mean	SD	Cronbach's alpha	SEm
Apprentice applicants to engineering roles in engineering company	139	32.37	3.49	0.60	2.21

*These results are based on the samples used in the development of the CWS series. The statistics reported are derived from those items which were selected for inclusion in the final forms of the instruments. Those items which were removed are excluded from the computation of these means, standard deviations, reliability coefficients, and standard errors of measurement.

Validity

Face validity

One of the key aims in the development of the CWS workforce series was to create a battery of tests with high face validity. Hence the themed format of the instruments and the realistic time limits.

Much anecdotal evidence has been collected in support of the face validity of the CWS instruments. This has taken the form of positive feedback from candidates, supervisors, line managers, senior managers, and personnel specialists within organisations which use the CWS. In addition, a simple study was undertaken with 93 apprenticeship applicants to the Steel Industry.

These applicants had completed three tests from the CWS, and the Morrisby Differential Test Battery. The candidates were asked to make a direct comparison between the two batteries. They were asked:

Which of the tests (CWS or Morrisby) did you feel was more relevant to a real working environment?

- 86% indicated that the Criterion Workforce Tests were more relevant.
- 14% indicated that the Morrisby Tests were more relevant.

A large majority therefore favoured the face validity of the CWS.

Content validity

The content validity of the CWS instruments for use in the Manufacturing Sector is assured on the basis that the nature, structure, and content of the instruments were all determined on the basis of rigorous job analysis at the time when the instruments were developed.

Construct validity

90 apprentice incumbents completed a battery of SHL aptitude tests at the time of their selection and subsequently (1-2 years later) completed CWS instruments in the context of a standardisation exercise. Relationships between the CWS and the SHL tests are shown below:

Test	SHL: NT2 Numerical Computation (N=28)	SHL: MT4 Mechanical Comprehension (N=90)	SHL: VTS1 Following Instructions (N=90)	SHL: DA5 Diagramming (N=90)	SHL: FTS4 Fault Finding (N=60)
CWS: Verbal	.34	.31	.67	.46	.33
CWS: Numerical	.60	.50	.63	.65	.54
CWS: Fault Detection	.48	.32	.37	.36	.38
CWS SPC1	.46	.36	.43	.48	.35

Criterion-related validity

Two concurrent separate validation studies have been conducted on the CWS. Both of these involved the testing of job incumbents within the Steel Industry. Data was gathered for the CWS verbal, numerical and fault detection test and assessed in relation to job performance data (including time keeping, attendance record, reliability, motivation and more). Based on the statistical analysis of such data, the CWS series tests have been found to be statistically significant in relation to external criteria, and have been found to add an objective assessment appropriate in selection processes.

Further details about the validation studies can be provided on request.