

Q45. Test Evaluation Exercise

Conduct an evaluation of the use and psychometric properties of a test Manual and use the *Test Evaluation Summary Form* below to summarise the key features of the test.

Test Evaluation Summary

NAME OF TEST/PUBLISHER

CRTB2 - CRITICAL REASONING TEST BATTERY

Psytech International.

BACKGROUND TO THE TEST

Developed as a measure of general mental ability.

FACTORS ASSESSED

Verbal critical and numerical reasoning tests

REASON FOR USE

Employee selection

RELIABILITY

Test-retest (Rec. +0.70): Generally, reasoning tests are recognised as more reliable than personality tests and for this reason high standards of reliability are usually expected from such tests. While many personality tests are considered to have acceptable levels of reliability if they have reliability coefficients in excess of 0.7, reasoning tests should have reliability coefficients in excess of 0.8.

Internal consistency: Each of these reliability coefficients is substantially greater than 0.8, clearly demonstrating that the Verbal Critical Reasoning (0.88) and Numerical Critical Reasoning (0.83) are highly reliable across a range of samples.

VALIDITY

Construct Validity: As an evaluation of construct validity, the VCR2 and NCR2 were correlated with the APIL-B (Ability, Processing of Information and Learning Battery). The publisher reported these correlations as highly statistically significant, and substantial in size, providing strong support for the construct validity of the VCR2 and NCR2 but on review of the actual statistics .569 and .512 are actually moderate and one would expect them to be around .7

Construct Validity: Overall, comment on the evidence indicating that the test measure the constructs it proposes to:

Criterion Validity: A sample of 132 Insurance Sales Agents completed the CRTB as part of a validation study. The association between their scores on the VCR2 and NCR2 and their job performance was examined. Results indicate that, on average, the successful incumbents had significantly higher scores on these tests than did the non-successful incumbents. The difference in scores between these two groups reached statistical significance for the NCR2.

In addition, a group of MBA students completed the VCR2 and NCR2 prior to enrolling. Their scores on these.

NORMS

Sample size (Rec: N > 150)

Mixed sample of 365.

Male/female

Men = 245 Women = 119. Mean scores were provided for men and women on the verbal and numerical critical reasoning tests.

Age ranges

Mean age of 31.7. There was an undergraduate sample of 70 so the age distribution is skewed to the younger end of the age range of the general population.

Minorities

Nothing specified. Sample drawn from graduates, managerial and professional groups and undergraduates. Are minorities adequately represented in these groups? Test publisher would need to be queried.

Occupations

Insurance Sales Agents, MBAs and undergrads.

Education

MBAs, undergraduate and unclear re insurance sales agents.

Norms: Overall, comment on the appropriateness of the representativeness of the norms and sample sizes:

COMPLIANCE WITH LEGISLATION (Is there evidence of bias towards minorities, etc?)

Bias on gender differences but not significant.

Correlation between verbal and non-verbal is questionable.

ACCOMMODATIONS FOR DISABILITY (Are accommodations for disability specified?)

None in the manual Need to be asked to the test publisher.

RESTRICTIONS ON USE (who it is primarily recommended for use with)

Geared towards people of above average intelligence; ideally with graduates as this was referenced in the sampling. Aimed at people looking for senior-manager roles. Country of design and use is important to minimise cultural discrepancies.

PRACTICAL CONSIDERATIONS

TIME: 30 minutes to complete

EQUIPMENT: None needed.

MODE OF ASSESSMENT: Online or paper and pencil.

COSTS Available on <https://psytech.com/>