

VCA Assessment Exercises-Summary report on internal reliability

RELIABILITY ANALYSIS

ELEANOR RUTHERFORD, C.PSYCHOL.

Contents

Contents	1
Index of tables	2
Introduction	3
Types of Reliability	3
Test-retest reliability	3
Inter-rater reliability.....	4
Internal reliability.....	4
Level of Reliability	4
Method	5
Results	6
Fire Fighter Community Exercise	6
Crew level set one.....	6
Crew level set two	7
Crew level set three	7
Watch manager set one	8
Watch manager set three.....	8
Station manager set one	9
Station manager set two	9
Group manager Case Study	10
Conclusion and Recommendation	10
References	11

Index of tables

Table 1: Summary of candidate numbers.....	3
Table 2: Guide to level of alpha	4
Table 3: Competencies measured by each exercise	5
Table 4: Fire Fighter Community Exercise Cronbach's alpha	6
Table 5: Crew manager set one Cronbach's alpha	6
Table 6: Crew manager set two Cronbach's alpha	7
Table 7: Crew manager set three Cronbach's alpha.....	7
Table 8: Watch manager set one Cronbach's alpha	8
Table 9: Watch manager set three Role play Cronbach's alpha.....	8
Table 10: Station manager set one Cronbach's alpha	9
Table 11: Station manager set two Cronbach's alpha	9
Table 12: Group manager Case study Cronbach's alpha	10

Introduction

Data has been collected on the performance of candidates for all levels of role, across several fire services in England. In line with Standards for The Design and Delivery of Assessment (British Psychological Society, Division of Occupational Psychology) reliability analysis has been conducted on exercises where numbers were sufficient. Numbers requirements for reliability analysis vary, however for the purposes of this report, Samuels' (2015) recommendation of a minimum size of 30 was followed. Analysis was conducted on the following data sets: -

Table 1: Summary of candidate numbers

	level set	Case study numbers	Role play numbers
<i>Fire fighter Crew</i>	NA	103	
	1	57	75
	2	30	62
<i>Watch</i>	3	30	31
	1	62	89
<i>Station</i>	3	NA	36
	1	63	53
<i>Group</i>	2	59	54
	1	34	NA

Types of Reliability

There are three main types of reliability used within psychometric analysis. These are outlined below for reference, but it should be noted that not all of these are appropriate within the quality assurance process for assessment centre exercises as standards for these are separate from psychometric instruments such as intelligence tests and personality questionnaires.

Test-retest reliability

This is used within psychometric tests which measure stable characteristics which are not expected to change over time, for example personality. Test-retest reliability is not appropriate for characteristics which may be open to change due to practice effects or change due to personal development by the individual. Furthermore, under normal

circumstances assessment centre exercises are not generally used twice on the same cohort or where some but not all, of the cohort may have previously completed the exercise. This ensures fairness for all candidate but prevents the collection of data for test-retest reliability. For all these reasons assessment centre exercises are not evaluated for test-retest reliability.

Inter-rater reliability

This is used where two or more raters are used to assess a trait, behaviour, or skill. Since the current exercises are assessed using one assessor this analysis is not possible.

Internal reliability

Internal reliability is used to assess the internal consistency of a scale. It is widely used within psychometrics and research. It allows the consistency of behavioural indicators to be measured and ensures that levels of random error within a measurement are low. It should be noted that reliability is a property of the tests scores rather than the measurement itself (Pedhazur and Schmelkin, 1991). There are multiple measures of internal reliability including split half, Kuder Richardson and Cronbach's Alpha. Cronbach's alpha is one of the most popular measurements used as it is robust and easily applied to this type of data.

Level of Reliability

Cronbach's alpha can produce a coefficient anywhere between -1 and 1 (Streiner, 2003), though it would be extremely unusual to see results below zero. Bartram (1995) provides the following guide for reliability results: -

Table 2: Guide to level of alpha

correlation	Verbal label
.85	Excellent
.8	Good
.7	Reasonable
.6	Adequate
Under .6	Inadequate

Method

Two exercises are used for the assessments, case study and a role-play. Each exercise assesses a number of competencies as per table below. Each competency is assessed by three behavioural indicators.

Table 3: Competencies measured by each exercise

Case study	Personal Impact
	Outstanding Leadership
	Service Delivery
	Organisational Effectiveness
Role play	Personal Impact
	Outstanding leadership
	Service Delivery
	Organisational Effectiveness

In order to evaluate the overall reliability of each exercise and competency, Cronbach's alpha has been calculated for each exercise (with all competencies) (12 behavioural indicators). The exception to this is the firefighter exercise which has 13 behavioural indicators.

Results

Fire Fighter Community Exercise

Table 4: Fire Fighter Community Exercise Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Community exercise	.802	Good

The overall level of Cronbach's alpha is high at .802 for the community exercise, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Crew level set one

Table 5: Crew manager set one Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.890	Excellent
Role Play	.873	Excellent

The overall level of Cronbach's alpha is high within each exercise at .890 for the case study and .873 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Crew level set two

Table 6: Crew manager set two Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.905	Excellent
Role Play	.830	Excellent

The overall level of Cronbach's alpha is high within each exercise at .905 for the case study and .830 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Crew level set three

Table 7: Crew manager set three Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.898	Excellent
Role Play	.855	Excellent

The overall level of Cronbach's alpha is high within each exercise at .898 for the case study and .855 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Watch manager set one

Table 8: Watch manager set one Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.835	Good
Role Play	.857	Excellent

The overall level of Cronbach's alpha is high within each exercise at .835 for the case study and .857 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Watch manager set three

Table 9: Watch manager set three Role play Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Role Play	.859	Excellent

The overall level of Cronbach's alpha is high within each exercise at .859 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Station manager set one

Table 10: Station manager set one Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.926	Excellent
Role Play	.924	Excellent

The overall level of Cronbach's alpha is high within each exercise at .926 for the case study and .924 for the role play, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Station manager set two

Table 11: Station manager set two Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.689	Adequate
Role Play	.887	Excellent

The overall level of Cronbach's alpha is adequate within the case study at .689 and high with the roleplay at .887, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Group manager Case Study

Table 12: Group manager Case study Cronbach's alpha

Competency	Result Cronbach's Alpha Coefficient	Level (as per Bartram, 1995)
Case study	.760	Reasonable

The overall level of Cronbach's alpha is reasonable within the case study at .760, as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP). This indicates there is a high level of consistency in performance across competency areas.

Conclusion and Recommendation

The exercises are currently demonstrating an excellent level of internal reliability indicating accuracy as per standards required by the British Psychological Society (BPS), Division of Occupational Psychology (DOP).

In line with BPS DOP standards for long term evaluation, data should continue to be collected and reviewed as a minimum on an annual basis. Where possible performance data should also be requested on candidates to allow long term validation work to be conducted to supplement the reliability studies.

References

Bartram, D. (1995) Review of Personality Assessment Instruments (level B). British Psychological Society

British Psychological Society, Division of Occupational Psychology (2015) The Design and Delivery of Assessment Centres: A Standard produced by the British Psychological Society's Division of Occupational Psychology. British Psychological Society

Pedhazur, E.K., & Schmelkin, L.P. (1991) Measurement Design and Analysis: An Integrated Approach Lawrence Erlbaum Associates Ltd.

Samuels, P. (2015). Statistical methods: Scale reliability analysis with small samples. Birmingham: Birmingham City University, Centre for Academic Success.

Streiner, D.L., (2003) Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. Journal of Personality Assessment. 80 (1) 99-103