

APPENDIX F: APPEAL FORM

Should submitters feel the need to dispute the findings of the review process, the final outcome and the certification decision reached can be appealed.

An appeal shall be considered valid if any of the following criteria are met:

- It can be substantively motivated, with research findings, where possible.
- A misinterpretation or misunderstanding of the psychometric information provided can be shown to have occurred which could have led to erroneous conclusions being reached by the reviewers.
- In the case of the review outcome assigned to the test during the evaluation process, substantive reasons are provided why this was erroneous.

Please complete Sections 1 and 2.

1 General information

1.1	Instrument name (local version)	
1.2	Short name of the test <i>(if applicable)</i>	
1.3	Original test name <i>(if the local version is an adaptation)</i>	
1.3.1	Authors of the original test	
1.3.2	Authors of the local adaptation	
1.4	Local test distributor/publisher	
1.5	Publisher of the original version of the test <i>(if different to current distributor/publisher)</i>	
1.5.1	Date of publication of current revision/edition	
1.5.2	Date of publication of adaptation for local use	
1.5.3	Date of publication of original test	

2 Evaluation

In the following section: **ONLY for the sections under appeal**, give the rating allocated for the item in the review, in the first column. For any arguments/further information supplied/ or further studies that are relevant to the item, please indicate what these are in the column on the right, with dates and page references to your manuals where appropriate. Do not indicate what you would like the new rating to be.

ORIGINAL REVIEW REFERENCE	DESCRIPTION OF SCALE	RATING IN LAST REVIEW	REASON FOR APPEAL AND EVIDENCE SUPPLIED
2.1	RELIABILITY INFORMATION		
2.1.1	Data provided about reliability		
2.1.2	Interpretation of reliability evidence		
2.1.3	Reliability coefficients are reported with samples that		
2.1.4	Internal consistency		
2.1.4.1	Sample size		
2.1.4.2	Kind of coefficients reported		
2.1.4.3	Size of coefficients		
2.1.5	Test-retest reliability – Temporal stability		
2.1.5.1	Sample size		
2.1.5.2	Size of coefficients		
2.1.5.3	Data provided about test-retest interval		
2.1.6	Equivalence reliability (Parallel or Alternative forms)		
2.1.6.1	Sample size		
2.1.6.2	Are assumptions for parallelism met?		
2.1.6.3	Size of coefficients		
2.1.7	IRT based method		
2.1.7.1	Sample size		
2.1.7.2	Kind of coefficients reported		
2.1.7.3	Size of coefficients		
2.1.8	Inter-rater reliability		
2.1.8.1	Sample size		
2.1.8.2	Kind of coefficients reported		
2.1.8.3	Size of coefficients		
2.2	VALIDITY INFORMATION		
2.2.1	Content validity		
2.2.1.1	Content validity processes followed		
2.2.2	Construct validity		
2.2.2.1	Information about construct validity presented		
2.2.2.1.1	Factor analysis		
2.2.2.1.2	Mean score differences for relevant groups		
2.2.2.1.3	Correlations with similar constructs (convergent validity).		
2.2.2.1.4	Correlations with different constructs (discriminant validity)		
2.2.2.1.5	Rasch analysis		
2.2.2.2	Adequate sample sizes		
2.2.2.3	How old are the studies?		
2.2.3	Criterion validity		

2.2.3.1	Description of the type of criterion study / information presented (concurrent / predictive)		
2.2.3.2	Sample sizes		
2.2.3.3	Quality of the criterion measure used		
2.2.3.4	Strength of the relation between test and criterion scores.		
2.2.3.5	How old are the criterion validity studies?		
2.2.4	Reviewers comments, evaluation & recommendation on validity		
2.3	Bias and equivalence information		
2.3.1	Evidence of factor structure invariance across relevant groups		
2.3.2	Investigation into differential item functioning for different sample groups		
2.3.3	Evidence of similarities of scores provided for different sample groups		
2.3.4	How old are the studies?		
2.3.5	Reviewers comments, evaluation & recommendation on bias and equivalence		
2.4	NORMS		
2.4.1	NORM-BASED INTERPRETATION		
2.4.1.1	Norms appropriate for local use		
2.4.1.2	Norms appropriate for the intended applications.		
2.4.1.3	Sample size overall		
2.4.1.4	Sample size (continuous/inferential norming)		
2.4.1.5	Procedures for sample selection		
2.4.1.6	Stratification/representativeness of the norm sample		
2.4.1.7	How old are the norm studies?		
2.4.2	DOMAIN-REFERENCED INTERPRETATION		
2.4.2.1	Expert judgement – judges appropriately selected and trained?		
2.4.2.2	Expert judgement – number of judges used adequate?		
2.4.2.3	Expert judgement – critical score for size of inter-rater agreement coefficient		
2.4.2.4	How old are the normative studies?		
2.4.3	CRITERION-REFERENCED INTERPRETATION		
2.4.3.1	Rationale used in developing critical scores		
2.4.3.2	How old are the normative studies?		