APPENDIX E:

EDITORIAL PANEL REVIEW UPDATE PRO FORMA

This form is for the purpose of submitting for an editorial panel review update where there is already a published test review appearing on the ASSA website, but a publisher / distributor has made substantive changes to the test since the publication of the initial review.

1. Please answer all of Sections 1–5, even if there have been no factual changes.
2. If there is new information, please highlight the items that have changed and note in the table any information that may be relevant. Otherwise, please leave these sections blank.
3. Where there are changes to items, scoring or computer-generated reports please send relevant documents to support this
4. For Section 5, **ONLY if you have made changes**, in the first column (rating in last review):
    a. Give the rating allocated for the item in the last review, and enclose supporting documentation in a copy of your manuals.
    b. Do not indicate what you would like the new rating to be.
    c. For any amendments / 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.
    d. Do not send research articles without stating where they are relevant and where the relevant information can be found.

1 General information

1.1	**Instrument name (local version)**	
1.2	**Short name of the test** *(if applicable)*	
1.3	**Original test name** *(if the local version is an adaptation)*	
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** *(if different to current distributor/publisher)*	
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 Classification

2.1	**Content domains** *(select all that apply)*	□ Ability - General □ Ability - Manual skills/dexterity □ Ability - Mechanical □ Ability Learning/memory □ Ability - Non-verbal/abstract/inductive □ Ability - Numerical □ Ability - Perceptual speed/checking □ Ability - Sensorimotor □ Ability Spatial/visual □ Ability - Verbal □ Attention/concentration □ Beliefs □ Cognitive styles □ Disorder and pathology □ Family function □ Group function □ Interests □ Motivation □ Organisational function, aggregated measures, climate etc □ Personality – Trait □ Personality – Type □ Personality – State □ Quality of life □ Scholastic achievement (educational test) □ School or educational function □ Situational judgment □ Stress/burnout □ Therapy outcome □ Values □ Well-being □ Other (please describe):
2.2	**Intended or main area(s) of use** *(please select those that apply)*	□ Clinical □ Advice, guidance and career choice □ Educational □ Forensic □ General health, life and well-being □ Neurological □ Sports and Leisure □ Work and Occupational □ Other (please describe):
2.3	**Description of the populations for which the test is intended**	

2.4	**Number of scales and brief description of the variable(s) measured by the instrument** Please indicate the number of scales (if more than one) and provide a brief description of each scale if its meaning is not clear from its name.	
2.5	**Response mode** If any special pieces of equipment (other than those indicated in the list of options, e.g. digital recorder) are required, they should be described here. In addition, any special testing conditions should be described.	□ Oral interview □ Paper & pencil □ Manual (physical) operations □ Direct observation □ Computerised □ Other (indicate):
3.6	**Demands on the test taker** Which capabilities and skills are necessary for the test taker to work on the test as intended and to allow for a fair interpretation of the test score? It is usually clear if a total lack of some prerequisite impairs the ability to complete the test (such as being blind and being given a normal paper-and-pencil test) but the requirements listed should be classified as follows: "Irrelevant / not necessary" means that this skill is not necessary at all – such as manual capabilities to answer oral questions verbally. • "Necessary information given" means that the possible amount of limitation is stated. • "Information missing" means that there might be limitations on test users without the specific capability or skill (known from theory or empirical results) but this is not clear from information provided by the test publisher e.g. if the test uses language that is not the test taker's first language.	Manual capabilities (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing Handedness (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing Vision (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing Hearing (*select one)* □ irrelevant / not necessary □ necessary information given □ information missing Command of test language (understanding and speaking) (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing Reading (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing Writing (*select one*) □ irrelevant / not necessary □ necessary information given □ information missing

3.7	**Items format** *(select one)* Two types of multiple-choice formats are differentiated. The first type concerns tests in which the respondent has to select the right answer from a number of alternatives as in ability testing (e.g., a figural reasoning test). The second type deals with questionnaires in which there is no clear right answer. This format requires test takers to make choices between sets of two or more items drawn from different scales (e.g., scales in a vocational interest inventory or a personality questionnaire). This format is also called 'multidimensional', because the alternatives belong to different scales or dimensions. In this case it is possible that the statements have to be ranked or the most- and least-like-me options be selected. This format may result in ipsative scales (see question 3.8). In Likert scale ratings the test taker also has to choose from a number of alternatives, but the essential difference with the multiple choice format is that the scales used are unidimensional (e.g., ranging from 'never' to 'always' or from 'very unlikely' to 'very likely') and that the test taker does not have to choose between alternatives from different dimensions. A scale should also be marked as a Likert scale when there are only two alternatives on one dimension (e.g., yes/no or always/never).	□ Multiple choice (ability testing, or right/wrong) Number of alternatives: .... □ Multiple choice (mixed scale alternatives) Number of alternatives: .... □ Likert scale ratings Number of alternatives: .... □ Open □ Other (please describe)
3.8	**Ipsativity** As mentioned in 3.7 multiple choice mixed scale alternatives *may* result in ipsative scores. Distinctive for ipsative scores is that the score on each scale or dimension is constrained by the scores on the other scales or dimensions. In fully ipsative instruments the sum of the scale scores is constant for each person. Other scoring procedures can result in ipsativity (e.g. subtraction of each person's overall mean from each of their scale scores)	□ Yes, multiple choice mixed scale alternatives resulting in partially or fully ipsative scores □ Yes, other item formats with scoring procedures resulting in partially or fully ipsative scores □ No, multiple choice mixed scale alternatives NOT resulting in ipsative scores □ Not relevant

3.9	**Total number of test items and number of items per scale or subtest** If the instrument has several scales or subtests, indicate the total number of items and the number of items for each scale or subtest. Where items load on more than one scale or subtest, this should be documented.	
3.10	**Intended mode of use (conditions under which the instrument was developed and validated)** (*select all that apply*) This item is important as it identifies whether the instrument has been designed with the intention of it being used in unsupervised or uncontrolled administration conditions. Note that usage modes may vary across versions of a tool. Note. The four modes are defined in the *International Guidelines on Computer-Based and Internet Delivered Testing* (International Test Commission, 2005, pp. 5-6).	□ *Open mode*: Where there is no direct human supervision of the assessment session and hence there is no means of authenticating the identity of the test-taker. Internet-based tests without any requirement for registration can be considered an example of this mode of administration. □ *Controlled mode*: No direct human supervision of the assessment session is involved but the test is made available only to known test- takers. Internet tests will require test-takers to obtain a logon username and password. These often are designed to operate on a one- time-only basis. □ *Supervised (proctored) mode:* Where there is a level of direct human supervision over test-taking conditions. In this mode test-taker identity can be authenticated. For Internet testing this would require an administrator to log-in a candidate and confirm that the test had been properly administered and completed. □ *Managed mode:* Where there is a high level of human supervision and control over the test-taking environment. In CBT testing this is normally achieved by the use of dedicated testing centres, where there is a high level of control over access, security, the qualification of test
3.11	**Administration mode(s)** *(select all that apply)* If any special pieces of equipment (other than those indicated in the list of options, e.g. digital recorder) are required, they should be described here. In addition, any special testing conditions should be described. 'Standard testing conditions' are assumed to be available for proctored/ supervised assessment. These would include a quiet, well-lit and well-ventilated room with adequate desk-space and seating for the necessary administrator(s) and candidate(s).	□ Interactive individual administration □ Supervised group administration □ Computerised locally-installed application – supervised/proctored □ Computerised web-based application – supervised/proctored □ Computerised locally-installed application – unsupervised/self-assessment □ Computerised web-based application – unsupervised/self-assessment □ Other (indicate):

3.12	**Time required for administering the instrument** *(please specify for each administration mode)* The response to this item can be broken down into a number of components. In most cases, it will only be possible to provide general estimates of these rather than precise figures. The aim is to give the potential user a good idea of the time investment associated with using this instrument. Do NOT include the time needed to become familiar with the instrument itself. Assume the user is experienced and qualified. • Preparation time (the time it takes the administrator to prepare and set out the materials for an assessment session; access and login time for an online administration). • Administration time per session: this includes the time taken to complete all the items and an estimate of the time required to give instructions, work through example items and deal with any debriefing comments at the end of the session. • Scoring: the time taken to obtain the raw scores. In many cases this may be automated. • Analysis: the time taken to carry out further work on the raw scores to derive other measures and to produce a reasonably comprehensive interpretation (assuming you are familiar with the instrument). Again, this may be automated. • Feedback: the time required to prepare and provide feedback to a test taker and other stakeholders. It is recognised that time for the last two components could vary enormously - depending on the context in which the instrument is being used. However, some indication or comments will be helpful.	Preparation: Administration: Scoring: Analysis: Feedback:

3.13	**Indicate whether different forms of the instrument are available and which form(s) is (are) subject of this review** Report whether or not there are alternative versions (genuine or pseudo-parallel forms, short versions, computerised versions, etc.) of the instrument available and describe the applicability of each form for different groups of people. In some cases, different forms of an instrument are meant to be equivalent to each other - i.e. alternative forms. In other cases, various forms may exist for quite different groups (e.g. a children's form and an adult's form). Where more than one form exists, indicate whether these are equivalent/alternate forms, or whether they are designed to serve different functions - e.g. short and long version; ipsative and normative version. Also describe whether or not parts of the whole test can be used instead of the whole instrument. If computerised versions do exist, describe briefly the software and hardware requirements. Note that standalone computer-based tests (CBT) and online packages, if available, should be indicated.	

3 Measurement and scoring

3.1	**Scoring procedure for the test** *(select all that apply)* Bureau services are services provided by the supplier - or some agent of the supplier - for scoring and interpretation. In general, these are optional services. If scoring and/or interpretation can be carried out ONLY through a bureau service, then this should be stated in the review.	□ Computer scoring with direct entry of responses by test taker □ Computer scoring by Optical Mark Reader entry of responses from the paper response form □ Computer scoring with manual entry of responses from the paper response form □ Simple manual scoring key – clerical skills only required □ Complex manual scoring – requiring training in the scoring of the instrument □ Bureau-service – e.g. scoring by the company selling the instrument □ Other (please describe):
3.2	**Scores** Brief description of the scoring system to obtain global and partial scores, correction for guessing, qualitative interpretation aids, etc).	
3.3	**Scales used** (*select all that apply)*	*Percentile Based Scores* □ Centiles □ 5-grade classification: 10:20:40:20:10 centile splits □ Deciles □ Other (please describe): *Standard Scores* □ Z-scores □ IQ deviation quotients etc (e.g. mean 100, SD=15 for Wechsler or 16 for Stanford-Binet) □ College Entrance Examination Board (e.g. SAT mean=500, SD=100) □ Stens □ Stanines, C-scores □ T-scores □ Other (please describe): □ Critical scores, expectancy tables or other specific decision-oriented indices □ Raw score use only □ Other (please describe):
3.4	**Score transformation for standard scores**	□ Normalised – standard scores obtained by use of normalisation look-up table □ Not-normalised – standard scores obtained by linear transformation □ Not applicable

4 Supply conditions

This defines what is provided, to whom, under what conditions and at what costs. It defines the conditions imposed by the supplier on who may or may not obtain the instrument materials. If one of the options does not fit the supply conditions, provide a description of the relevant conditions.

4.1	**Documentation provided by the distributor as part of the test package** *(select all that apply)*	□ User Manual □ Technical (psychometric) manual □ Supplementary technical information and up-dates (e.g. local norms, local validation studies etc.) □ Books and articles of related interest
4.2	**Methods of publication** *(select all that apply)* For example, technical manuals may be kept up-to-date and available for downloading from the Internet, while user manuals are provided in paper form or on a CD/DVD.	□ Paper □ CD or DVD □ Internet download □ Other (specify):
4.3	**Test-related qualifications required by the supplier of the test** *(select all that apply)* This item concerns the user qualifications required by the supplier. For this item, where the publisher has provided user qualification information, this should be noted against the categories given. Where the qualification requirements are not clear this should be stated under 'Other' *not* under 'None'. 'None' means that there is an explicit statement regarding the lack of need for qualification.	□ None □ Test specific accreditation □ Accreditation in general achievement testing: measures of maximum performance in attainment □ Accreditation in general ability and aptitude testing: measures of maximum performance in relation to potential for attainment □ Accreditation in general personality and assessment: measures of typical behaviour, attitudes and preferences □ Other (specify):
6.7	**Professional qualifications required for use of the instrument** *(select all that ap- ply)* This item concerns the user qualifications required by the supplier. For this section, user qualification information should be noted against the categories given. Where the qualification requirements are not clear this should be stated under 'Other' *not* under 'None'. 'None' means that there is an explicit statement regarding the lack of need for qualification. For details of the EFPA user standards, consult the latest version of these on the EFPA website.	□ None □ Practitioner psychologist with qualification in the relevant area of application □ Practitioner psychologist □ Research psychologist □ Non-psychologist academic researcher □ Practitioner in relevant related professions (therapy, medicine, counselling, education, human resources etc.). Specify: …… □ EFPA Test User Qualification Level 1 or national equivalent □ EFPA Test User Qualification Level 2 or national equivalent □ Specialist qualification equivalent to EFPA Test User Standard Level 3 □ Other (indicate):

5 Evaluation

In the following section: **ONLY if you have made changes**, give the rating allocated for the item in the last review, in the first column. For any amendments/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	CHANGES/NEW INFORMATION WITH DATE AND DOCUMENT AND PAGE REFERENCE
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?		