International Adaptation plan Guidelines

For all new products to be developed, you have been kindly asked to complete an International Adaptation plan.

Our international R&D team will review your plan. This means that your plan can be:

- ✓ approved with/without comments, no additional review is needed. Your Licensing team will feedback to you upon approval.
- ✓ not approved, additional information is required. Your Licensing team will contact you to request additional information or adjustments in your plan. You will be asked to re-submits your plan once adjusted. This process continues until adaptation plan has been formally approved and filed.
- ✓ Our R&D team will check the outcome of your adaptation work against the adaptation plan submitted. In this, we assume that you have completed your work in line with the research plan and accepted the quality criteria.

Criteria/minimum standards per development phase

Please note that not all criteria/minimum standards listed below are relevant for, or applicable to, all kinds of tests. You should always consult the original test version manual to understand the requirements for any specific test. Model research plans can in some cases be made available to you upon request. If needed, please contact the License team.

As our valued partner you should be familiar with, and work in accordance with, the recommendations for test adaptation set up by the International Test Commission (2017), The ITC Guidelines for Translating and Adapting Tests (Second edition), as well as the quality criteria established by the European Federation of Psychologist´s Associations (EFPA) 2013.

While there is a standard approach to test adaptation aimed at ensuring high quality in the translated version of the test, we acknowledge that it can occur that a smaller effort is warranted for a certain test in a certain market (e.g., a translation only or a norm validation). If you are planning for a translation only or a norm validation, a clear rational should be given for this.

In planning and describing the adaption, we assume that you will adhere to any local test quality criteria set up by a relevant professional organization/review body that might exist in the territory.

Translation, Adaptation, Item development

Translation of items can be conducted by professional translator or by a SME (Subject Matter Expert). No matter the actual translation process, a thorough language and cultural review in cooperation with SMEs is crucial to capture intended psychological constructs in new language and culture. Translations should also be conducted with an effort to avoid wording in the target language that might be perceived as discriminatory and/or biased. In our new ways of working we no longer require a back-translation.

We kindly ask you to stay as close as possible to the original content, but new items can be developed due to difficulties to translate/culturally adapt certain items in original test. All new items must have a clear rationale. As we trust upon your expertise on this matter, we will not review translated content or new items. Of course, should you need any advice, you can at any time consult us by contacting the Licensing team.

Pilot study

The purpose of a pilot study is to ensure clarity of items and intelligibility of instructions to examinees of all ages, gender and educational levels, as well as administration guidelines for examiners. The goal of the pilot is thus not to collect quantitative data, but to make sure the test is ready for data collection in the next development stage.

The pilot study should be carried out on a small number (minimum 10) of examinees. When choosing suitable examinees, take into consideration age/gender/examinees educational level/parent´s educational level. In many cases, choosing young examinees or examinees with low educational background will be sufficient to ensure understanding of instructions. Equally important is to make sure that instructions to examiners are clear and unambiguous.

A pilot study is recommended for any adapted test (including translations only) for which the instructions/items are not well “tried and tested” in the adapted language. However, it is not normally required for tests with minimal or no changes.

Try-out

The purpose of try-out study is to ensure construct coverage for translated/adapted content, scales cohesiveness/reliability, item order to ensure adequate difficulty progression (if applicable), mitigate floor and ceiling effects.

The goal of try-out is to ensure that the translated/adapted test has sufficient psychometric qualities to form the basis for the standardization version (i.e. the version that is used for norm data collection, reliability and validity studies)

Based on results from a try-out study, selection of items (translated original items, additional new items that have been developed based on language/culture considerations), item order (if applicable) can be

established, reliability estimates can inform final item choices, DIF analysis (min requirement N=100 of each group) can detect biased items.

Including clinical data (i.e., Intellectually disabled and Gifted cases) in try-out sample will serve to investigate floor and ceiling effects, when needed.

Try-out studies are not needed for all types of tests, but for individually administered test which require items with a difficulty progression over ages, try-out is mandatory.

For symptom/problem rating scales, a try-out study is recommended to ensure reliability of translated/adapted/new items. Another option that can be considered is including (a smaller number of) additional items in standardization version.

Try-out should be conducted on a stratified sample, based on local census. Age, gender, examinee/parent educational level are mandatory stratification variables. Collected try-out sample should not deviate from stratification plan more than max 5 % per age group.

Minimum sample size per age group in individually administered tests, with increasing difficulty levels, is 20.

Inclusion and exclusion criteria should be defined (recommendation is to consult the original version test manual for criteria to use).

The try-out study with the above defined goals and the above defined criteria (adjusted to the test in question) is relevant for tests (and subtests) with a large degree of language and culture adaptation (for example verbal subtests in Wechsler scales) and where decisions around item inclusion and item order need to be defined before standardization.

Norming

The purpose of norming is to create norms representative of the intended population.

The norm sample must be stratified in accordance with relevant variables from national census, those variables would – depending on country - normally be region, age, gender, examinee/parent educational level/socioeconomic status/occupation and ethnicity.

The collected sample must not deviate from stratification plan more than 5 % per age group.

Inclusion of Intellectually disabled and Gifted cases is needed in ability tests.

Minimum sample size depends on type of test/scale but must not – for cognitive/ability/speech and language tests – be less than 60 per age group when the age range is divided into minimum 10 age groups and must include 2 Intellectually disabled and Gifted cases per age group (or corresponding suitable cases to check for floor and ceiling effects in speech and language scales).

For symptom/problem rating scales there might be – depending on instrument – a need for norming that is built on prevalence data for certain psychiatric conditions and/or clinician ratings. etc. There will also be cases where no norms as such are used, and results are interpreted according to different cut-off scores.

· Use description of original development work to inform adaptation plan and make sure different adaptation steps are clearly described.

Description of stratification details, number of cases in norm group must be clearly defined. See example tables in International Licensing Adaptation Plan template.

Norm validation

In some cases, a norm validation may be acceptable for a region (e.g., non-verbal tests where there is already evidence on low culture and language effects on the test). A norm validation typically requires at least N=200 s with a demographically representative sample spread along the age span of the test.

Reliability

· Adaptation plans should have information on planned reliability studies.

· Evidence of reliability should rest upon minimum internal consistency.

· Other types of reliability (test-retest, inter-rater) are recommended as appropriate given the nature of the test.

· For test-retest, sample size should not lower than N = 30

· Internal consistency should be over .70. Exceptions can be founded given the number of items and ceiling/floor effects. This should however only concern disparate reliability estimates and not affect a subscale/scale in a systematic way.

· If no reliability studies are planned, a rational should be provided.

Validity

· Adaptation plans must have information about planned validity studies.

· Validity studies (in addition to evidence for validity based on internal structure) are recommended for evidence of clinical utility; evidence based on relationships with other variables studies (studies in special groups and with other tests) can provide evidence of the instruments usefulness in clinical practice and give support for the inferences made from test scores.

· Minimum sample size in any type of validity study is 20.

· If no validity studies are planned, a rational should be provided.

Psychometric Evaluation

· Adaptation plan should have a summary of the analyses that will be completed to evaluate the normative data, reliability and validity studies.

* https://pearsoneducationinc-my.sharepoint.com/:w:/r/personal/katarina_forssen_pearson_com/_layouts/15/Doc.aspx?sourcedoc=%7B0D5E67E1-F2E6-4DD1-B7AC-FA4AE3CF5428%7D&file=International%20Licensing%20Adaption%20Plan%20Template%202021.docx&action=default&mobileredirect=true&cid=afb6c120-85b9-4f38-a21c-19254c6c40de