

Adapting the Measure of Development and Early Learning (MODEL) Literacy tasks to new languages and contexts

MELQO Adaptation: Introduction

The need to translate and adapt tests to new languages and cultures has increased significantly in recent years.¹ For example, the Early Grade Reading Assessment,² an individually administered assessment for primary grade students, has been used in more than 65 countries and translated into more than 100 languages. Despite the growing need, there is often little guidance on how to adapt tests to new cultures and languages while maintaining the integrity of the instrument's purpose and ensuring that the results are valid and reliable.

This brief seeks to help remedy that by outlining guidelines for adapting the early childhood assessment tool, Measuring Early Learning Quality and Outcomes (MELQO), specifically for the literacy tasks in the MODEL (Measure of Development and Early Learning) direct child assessment module. MODEL is designed to reflect universal early literacy skills for children ages four to six.

Understanding the distinction between translated and adapted tests is important. A *translated test* is one in which only the language changes between the source language and translated target language versions of the test, while the content or targeted constructs stay the same.³ An *adaptation* may also include some changes in the test items in order to ensure that the meaning and purpose of the test are consistent with the original test, and so that items are appropriate for the culture to which the test is being adapted.⁴

Box 1 sets out recommendations for translating and adapting the MODEL child assessment, based on guidelines from the International Test Commission.⁵ The rest of this brief reviews guidelines for translation and language considerations, including recommended quality control procedures, and provides specific recommendations for each of the literacy tasks in the MODEL MELQO.

Language considerations in adaptation and translation

This section provides specific considerations to ensure quality linguistic adaptation and translation of the MODEL module. Table 1 provides a summary, and Box 2 provides an example of the translation process of MELQO modules in Ethiopia.

When adapting an instrument, teams should work with an individual who is a linguist with fluency in both the local language into which the tool will be translated and in the source language, as well as a person who is fluent in the target language and who also knows the content of what is being translated. To minimize possible bias toward the test, the linguist should have no previous knowledge of it.

When checking to ensure that an adaptation has captured the intent of the instrument, it is important to include two individuals, one of whom is a linguist in English (the source language for MODEL) and the other who is a linguist in the local language (target language) involved in the translation. The local language linguist should have little knowledge of the test, in order to minimize possible bias toward the test during the translation by, for example, providing helpful hints in the translation that might bias the test. A third individual, an expert on the content as well as a fluent local language speaker, can help determine whether the translation is asking the question as intended. In summary, of the two local language linguists, one should be quite familiar with the test and one should not know the test.

Following the translation, the tool should be reviewed by experts and tested in the field. Experts in early childhood and literacy assessment will examine other aspects of the measure, including the structure, layout, assessor instructions, and whether the expressions used in the measure are cross-culturally appropriate. The visual layout of the task will be reviewed, including font sizes and formats, and how information is arranged on the instrument.

This brief is the first in a series on early childhood measurement. The series builds on the experience of the Measuring Early Learning Quality and Outcomes (MELQO) modules, designed to provide country-level evidence on child development and learning in response to country demands for national data. This series both summarizes findings and lessons learned to date and offers practical guidance for others wishing to conduct research using the MELQO modules. To access other MELQO briefs and for more information on the MELQO initiative, please visit ecdmeasure.org. This brief was written by Marcia Davidson and reflects the work of the MELQO global team, including but not limited to Magdalena Janus, Linda Platas, and Abbie Raikes.

BOX 1. Guidelines for Tool Adaptation

Finding the right experts

Local experts must be involved in the translation and adaptation process. These experts include individuals with knowledge of the languages and cultures involved, as well as knowledge of the content of the test and the general principles of testing. Individual experts may possess different qualifications, but all must have knowledge of the general principles of testing as well as the test content.

Because translators with knowledge of test development principles can be almost impossible to find, it is important to provide translators with training in this area. For MELQO tool adaptation, local stakeholders, including government officials, early childhood and assessment experts, and language experts should be invited to participate in the country adaptation process so that they can contribute to important decisions about how to adapt the instrument while maintaining its purpose, reliability, and validity.

Choosing translation designs

It is best practice to use multiple translation designs whenever feasible (see Box 2 for recommended examples). When possible, two versions of the translated instrument are created, one by an expert with knowledge of the construct(s) being assessed, and the other by a local language expert without knowledge of the goals of the assessment instrument.⁶ If two simultaneous translations are not feasible, one might have an expert in the target language translate or adapt the test (forward translation) to ensure that the construct being measured is the same in both languages, and then check for accuracy by asking a different person who is expert in the source language to translate the new translation back to the source language, using backward translation design.

Translation designs include both forward and backward translation, and each design has its own limitations. For example, in backward translation, the goal is to make certain that the translated version looks close to the original test, and there is often little attention to reviewing the target language version. Forward translation involves a single translator or (preferably) group of translators who adapt a test from the source to the target language. The primary limitation of forward translation is that the translators must rely on a very high level of inference about the equivalence of the two versions of the test. Another limitation is that because the translators may be better educated than the monolingual examinees, they may miss some of the problems that might confront the examinees.

The intended outcome is that the language of the translated and adapted test will feel natural and acceptable, with the focus on functionality and not on literal equivalence.⁷ Ideally, a double-translation approach can be employed, using a reconciliation procedure in which a third, independent expert or panel reviews two separate forward translations of the test and reconciles any differences in a new single version of the test.

Ensuring equivalence in meanings across populations and cultures

It is important to gather evidence that the test instructions and content have the same meaning across languages and cultures. For example, if a child is to respond to different facial expressions, it is important that the expressions have universal meanings across all the cultures where the test will be administered.

Several strategies can help ensure equivalence, including having a local expert review the translated test for cultural bias and cultural familiarity, reviewing the language and culture by local expert, administering the test to bilingual respondents to provide guidance on the equivalence of the two versions of the test, and trying out the adapted version on a small scale and collecting the feedback from the test responses as well as the assessors who administered the test.

Item format, item scoring, modes of test administration, and other administrative methods also need to be appropriate for all populations tested. Items should be administered in the same way, which requires enumerators to be trained in consistent administration based on culture and language. In addition, enumerators should be trained in how to respond to neutral praise in the local language (sample phrases, such as “uh huh” or “good” should be adapted and agreed upon for local context.)

Pilot-testing for reliability and validity

Pilot data on the adapted test should be collected to check for reliability and validity. Ideally, the pilot testing should include 100 subjects, but that is not always possible. When planning to pilot a newly adapted test version, it is important to ensure that the population from which the sample of children is randomly selected all speak the language of the newly adapted test. In many cultures, there are a number of local languages as well as dialects, so it is always a good idea to check whether a child speaks and understands the language prior to beginning any test administration. Ideally, the trained assessors will attend the adaptation workshop and assist with the process of adapting the measure. The assessor instructions must be standardized across languages. The pilot test should provide important information on whether the test instructions are clear to the child being tested and it is important to carefully observe both the assessor and the child during the piloting.⁸

Evaluating the newly adapted instrument

This involves assessing compatibility between the translated version(s) and the original instrument in four areas: (i) semantic equivalence, to make sure all words and phrases maintain the same meaning and that grammar is correct; (ii) idiomatic equivalence, referring to maintaining the cultural meaning of each item; (iii) experiential equivalence, which addresses whether the items are applicable to the new cultural context; and (iv) conceptual equivalence to determine whether any given term or expression in the task assesses the same aspect in different cultures.⁹

Table 1. Quality control procedures for MODEL literacy tasks translation

	Considerations for MODEL
Step 1: Translate the MODEL tool	Three linguists are needed for translation: <ul style="list-style-type: none">• An expert in English and MODEL literacy tasks• An expert in the local language with no knowledge of MODEL literacy tasks, and• An expert in the local language with knowledge of the MODEL literacy tasks.
Step 2: Review the translated tool for appropriateness	Following translation, MODEL tools should be <ul style="list-style-type: none">• reviewed by early childhood and literacy experts for appropriateness, and• tested in the field with the target population to ensure that procedures and instructions are clear.
Step 3: Evaluate the translated tool for compatibility	There need to be evaluation procedures that can be used as an additional quality control check to identify words or phrases in the translation that remain unclear or to find inconsistencies or conceptual errors in the translation. ¹⁰ Box 2 provides recommended examples of translation design for quality control

BOX 2. Examples of translation designs for quality control

Back-translation

Back-translation is an additional quality control check, but it should be done by individuals who were not involved in the original translation. The purpose of back-translation is to identify words or phrases in the translation that remain unclear or to find inconsistencies or conceptual errors in the translation. It is important that any translated item be *conceptually* consistent with the original item; it need not be literally identical to it. However, back-translation does not sufficiently address whether the translated test works in the same way that the original test works, so another source of quality control is necessary.

Evaluation of the adapted tool by the target population includes procedures to examine the clarity of the instructions, their appropriateness, and whether the items are well written. If a respondent (child) is confused by a question or task during the piloting, the item should be explained and the issue noted so that it can be revised for clarity.

Recommended guidelines for adapting the MELQO MODEL literacy tasks

The following general guidance for the adaptation of literacy items provided in the MELQO Overview¹³ includes procedures relating primarily to the structure and orthography of the language to which the MELQO will be adapted. These recommendations should be carefully adhered to in any adaptation process. In addition, it is important to verify that the children who will be tested actually speak the language

Simultaneous translation and comparison

Figure 1 summarizes another methodology for adapting and translating instruments derived from J. C. Borsa, B. F. Damásio, and D. R. Bandeira.¹¹ This reflects a process in which two translations are developed simultaneously by knowledgeable translators, one with knowledge of the construct and the second one without knowledge of the translation goals.¹² The first translation is likely to provide a translation that is most similar to the goals of the instrument, and the second translation is likely to better reflect the language used by the target population.

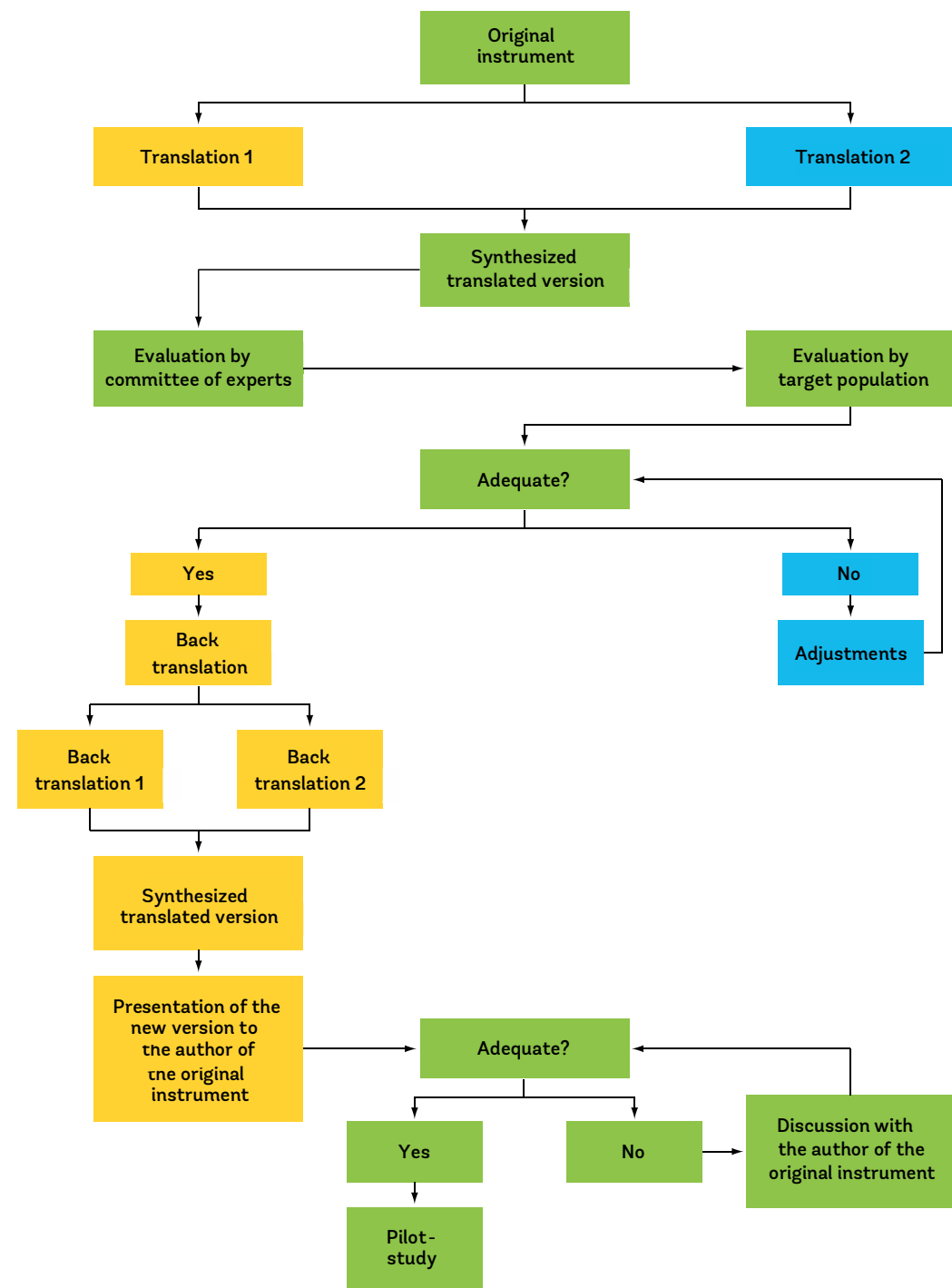
to which the MELQO will be adapted, and to seek out expert recommendations for letter sounds, letter names, and attention to maintaining story length after the adaptation.

Preparatory work

Before proceeding to the translation and adaptation of individual literacy items, three preparatory tasks are worth undertaking:

- Determine the type of script (e.g. alphabetic, alpha-syllabic, abjad) that is most appropriate.
- If possible, determine the frequency of letters or syllables (will be needed for the letter-name task).
- If possible, identify a bank of 50 common words likely to be known by preschool children.

Figure 1. Translation design option for cross-cultural adaptation



Source: J. C. Borsa, B. F. Damásio, and D. R. Bandeira, "Cross-cultural Adaptation and Validation of Psychological Instruments: Some Considerations," *Paidéia (Ribeirão Preto)* 22(53) (2012): 423–32.

Box 3. Tool translation experience in Ethiopia

In 2017, the Ethiopian Development Research Institute (EDRI) and the National Educational Assessment and Examinations Agency (NEAEA) conducted an adaptation and language translations for a national MELQO study. To capture the diverse linguistic, ethnic, and socio-economic groups in Ethiopia, the study included six (of 88) regional language groups (Amharic, Tigrigna, Afaan Oromo, Af Somali, Sidaamu Afoo, and Berta); the majority of the population (more than 70%) speaks one of these six languages.

The team translated all the tools that require direct interaction with a local language speaker from English into the six regional languages. Those tools were the Direct Assessment, Parent Caregiver Report, Teacher Survey, Teacher Questionnaire, and Head Teacher Questionnaire. The materials that only the enumerator uses (MELE classroom observation, manuals) were translated from English to Amharic only, as Amharic is the national language spoken by all enumerators.

Following a national adaptation exercise, a translation workshop was held in Ethiopia that included language experts representing the six languages, along with a moderator. The following process was followed:

1. Forward translations were done from English to the six local languages.
2. For each translation, a discussion was held between the two independent translators in the presence of a moderator, who was an English language expert with local language background. This step helped to assess and correct any deviations in the translation from the original English to the translated local language.
3. A general moderated discussion was held among all translators across language groups, based on the English core versions of the tools, concerning the meaning of items.
4. A verification was made by independent experts to evaluate the translations for compatibility. Six additional language experts (different from those involved earlier) verified the equivalence of the original English version tools with the locally translated versions.

The case of Ethiopia is somewhat unique, given the complexity and diversity of languages in the country, but overall it was challenging to ensure competent expert translation of a new set of instruments. Thus, the moderation and evaluation for compatibility were essential steps for quality control.

Construction and adaptation of literacy items

Task 1: Literacy interest. Check the face images to make sure that their intended meaning is consistent with children's interpretation. This requires field testing with a sample of children. It is recommended that if there is any concern that a child does not understand the task, add one or two sample items in which the assessor asks the child how (s)he feels when they are eating something delicious for happy or ask how the child feels if they lost their favorite toy.

Task 2: Expressive language. For this task, the literal translation of body parts is one step. For this task, the assessors need to understand the specific instances in which they need to query the child further (e.g., mouth/lips for tooth).

Task 3: Expressive vocabulary. This is a semantic fluency task, and it is important when adapting or translating this task that the category of items selected is one in which young children from a range of socioeconomic backgrounds will be able to generate an adequate number of responses. Animals and things one can eat should be categories in which many responses are likely to be generated by young children as long as they are not intimidated by the test context or by the adult assessor. This task was attempted as part of a different assessment with primary grade children in a country in West Africa with poor results, because the children were reluc-

tant to generate responses. This is a significantly different task than asking a direct question, and it is important that the assessor establish a good rapport with the child before initiating it.

Task 4: Letter identification. Use the frequency list to identify the easiest and hardest letters appropriate for beginning readers (in alpha-syllabic languages, children would not study some graphemes until second grade, so the letters should be drawn from those they are expected to learn in first grade). Randomly list the easiest letters in the first column and the most challenging in the second column.

For languages that read from right to left, reverse the columns; put the easiest letters in the right column and the most challenging ones in the left column, and present to the child in that order.

If a child begins by telling the assessor the sound the letter makes, the assessor should prompt by saying, "Tell me the name of the letter and not the sound it makes."

Task 5: Letter sound identification. This task should only be administered if the team training the assessors in the new language has evidence that the assessors can reliably hear

and reproduce the sound of each letter of the alphabet. In some languages, there is no distinction between the sound and the name of the letter, but when the sound differs from the name, letter sounds may not have been taught.

Task 6: Initial sound discrimination. This is likely to be very difficult for some young children as many cultures do not teach children the sounds they hear in words in preschool or later in the primary grades. Sufficient practice is very important to ensure that the child understands the task. More than one practice item is recommended.

Task 7: Listening comprehension. The story for this task should be carefully evaluated during the adaptation phase to ensure that it is appropriate for young children in the context with the newly adapted tool. Assessor training is also very important for this task as the story should be read at a reasonable pace with appropriate prosody in a manner that is engaging to the child. The translation must not result in a passage that is much more difficult or easier than the original passage, and it must have generally the same meaning across cultures.

The approximate length of the story for the listening comprehension task should be maintained. However, the length of a passage in an agglutinative language may be longer but with the same meaning due to the length of the words, so length can vary to some degree. In English, this story contains 110 words. The number of words may vary by language, but do not shorten or lengthen the content of the story (add or delete story events) just to get the same number of words. Also maintain the transitions (“after a while,” “so”), since those help children remember the sequence.

The story can be adapted in the following ways:

- The two animals can be changed to two animals that are common in the country.
- The story opening can be changed to what is commonly used in the language.
- Task 8: Name writing. Check to see if the child knows how to hold a pencil or piece of chalk and appears to know how to use it prior to beginning this task. A warm-up would be helpful in which the assessor asks the child to draw a line or mark on the page to observe the child’s level of skill in using a writing implement.

Task 8: Name writing. Check to see if the child knows how to hold a pencil or piece of chalk and appears to know how to use it prior to beginning this task. A warm-up would be helpful in which the assessor asks the child to draw a line or mark on the page to observe the child’s level of skill in using a writing implement.

Notes

¹ R. K. Hambleton, P. F. Merenda, and C. D. Spielberger, *Adapting Educational and Psychological Tests for Cross-Cultural Assessment* (New York: Psychology Press, 2005).

² RTI International, *Early Grade Reading Assessment* (Research Triangle Park, NC: RTI International, 2006).

³ M. Bowles and C. W. Stansfield, “Standards-based Assessment in the Native Language: A Practical Guide to the Issues” (2008). Retrieved September 5, 2017 from <https://www.edweek.org/media/maz-guide%20to%20native%20language%20assessment%20v15-blog.pdf>

⁴ C. W. Stansfield, “Oral Translation as a Test Accommodation for ELLs,” *Language Testing* 28(3) (2011): 401–16.

⁵ International Test Commission, *The ITC guidelines for Translating and Adapting Tests* (Second edition) (2017) [www.InTestCom.org].

⁶ J. C. Borsa, B. F. Damásio, and D. R. Bandeira, “Cross-cultural Adaptation and Validation of Psychological Instruments: Some Considerations,” *Paidéia* (Ribeirão Preto) 22(53) (2012): 423–32.

⁷ International Test Commission, *The ITC guidelines for Translating and Adapting Tests*.

⁸ International Test Commission, *The ITC guidelines for Translating and Adapting Tests*.

⁹ International Test Commission, *The ITC guidelines for Translating and Adapting Tests*.

¹⁰ S. Turkan, L. C. De Oliveira, O. Lee, and G. Phelps, “Proposing a Knowledge Base for Teaching Academic Content to English Language Learners: Disciplinary Linguistic Knowledge.” *Teachers College Record* 116 (4)(2014). Retrieved from <https://www.tcrecord.org/Content.asp?ContentID=17361>.

¹¹ Borsa, Damásio, and Bandeira, “Cross-cultural Adaptation and Validation of Psychological Instruments.”

¹² D. E. Beaton, C. Bombardier, F. Guillemin, and M. B. Ferraz, “Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures,” *Spine* 25(24)(2000): 3186–91.

¹³ UNESCO, UNICEF, Brookings Institution, and World Bank, *Measuring Early Learning Quality and Outcomes: Overview* (Washington, DC: 2017).