FILM 3.
USER PACK

The aim of this User Pack is to help you get the most out of our Language Assessment film series.

What is included?

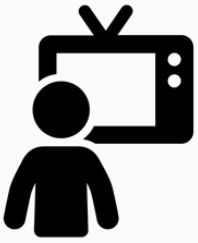
1. Tips for Watching - suggestions for how to get the most out of viewing the film.

2. **Reflection Questions** - questions to help you relate the film content to your context.

3. Talk to Text Cards - cards which provide a text version of the topics discussed in the film.

Film 3:
Assessing Speaking and Listening Skills:
What to check statistically about a test?

Watch the film **here.**



Tips for Watching

- Use a device with good sound quality – the audio is an important part of the film.
- Find a quiet location to watch the film – you can also watch the film as a group with colleagues.
- Use the 3 Rs as you watch: Rewind, Reflect, Repeat – pause and re-watch parts of the film to help you consider what has been said or allow more time to read on-screen text.
- Watch the film 2 or 3 times – this will help you to grasp any new ideas presented.
- Take notes while you watch – you may want to talk with colleagues about the content of the film afterwards.

Film 3:
Assessing Speaking and Listening Skills:
What to check statistically about a test?

Reflection Questions

Take a few minutes to think about what you have learned in the film. If you are viewing the film as a group, you can discuss these questions together.

1. In the film, we learn about **floor effects** (where the test was too difficult and most learners remained at the floor with a score of 0) and **ceiling effects** (where the test was too easy and most learners reached the ceiling, scoring full marks). **Have you ever used a test which was too easy or too difficult? How useful did you find the results from this test?**

2. Think about the different ability levels among students in your context (e.g., in early childhood centers; lower primary school). **How easy/difficult do you think it would be to find tests which are well-suited to the range of ability levels in the group? How would you go about doing this?**

3. The film mentions that sometimes test questions and test stimuli disadvantage particular groups of students. For example, a student in a rural, non-electrified village may not know what a computer is, and would therefore be disadvantaged if they had to name a picture of a computer in a test. **Have you come across any test questions or stimuli which you think disadvantaged learners in your context? In what way were they disadvantaged by the test, and what impact could that have on their education?**

TALK TO TEXT CARDS

Film 3:
Assessing Speaking and Listening Skills:
What to check statistically about a test?

1.VOCABULARY

A child's language system is multidimensional. Several building blocks come together to make a language system. These building blocks are: phonology, vocabulary, grammar and discourse.

Let's start with vocabulary. Vocabulary is about all the words that a child knows about; all the words in the language. This could be action words, like play, skip, wriggle, roll. This could be describing words like small, squat, fat, stubby. This could also be, for example, thinking words. It could be emotion words. It could be words that describe quantity, words that describe magnitude - a lot of different kinds of words.

2. THE RELATIONSHIP BETWEEN THE CHILD'S LANGUAGE ABILITY AND QUESTION DIFFICULTY

The Rasch Model was developed by the Danish statistician, Georg Rasch, in the 1950s. The core concept of the Rasch Model is that every time you have a student interact with a question on a test - or what we often refer to as an item on a test - that involved the trade off relationship between two elements. The first element is the ability of the student. The second element is the difficulty of the item.

As the ability of the student far exceeds the difficulty of the item, the likelihood that the student will get a correct response to the item goes up. Vice versa, as the difficulty of the item goes up relative to the student's ability, then the probability of them getting a correct answer to that question goes down.

One wants to develop a set of items - a set of questions - that cover a range of difficulty levels and that are well-targeted to the range of ability levels inherent in the group of students that you are trying to evaluate.

This can go wrong. Sometimes, when people develop sets of questions that are overall too difficult for many of the students taking the test, then we get what is referred to as a **floor effect**. This means a large proportion of the students, or even any substantial proportion of the students, score zero on the test: they don't get any questions correct.

Through the lens of the Rasch Model, this is problematic because we do not have any information about the ability level of those students. We simply know that the test was too difficult for them.

This also applies at the other end of the scale with the higher ability: if a test is far too easy for a proportion of the students taking the test, then this can result in what is referred to as a **ceiling effect**: we don't have any information about the relative abilities of the students. We simply know that the test was too easy for them.

3. TEST QUESTIONS SHOULD ALL GET TO ONE COMMON UNDERLYING CONSTRUCT

This is the idea that every question on the test assesses the common construct and only assesses a common construct. The idea is that there is an underlying common construct that should determine performance on the test. In the case of something like expressive vocabulary, one would want to design test questions that are all getting at this common underlying construct.

For example, the memory load of the questions would want to be minimised such that it is manageable for students of all abilities with respect to their expressive vocabulary.

Another key consideration in creating questions for a unidimensional test is, for example, the familiarity of any stimuli for the students who are responding to the test questions. One would intend to design these questions such that they draw on stimuli that are reasonably and similarly accessible to all students.

This can be particularly important to consider when designing questions and tests for low-resourced environments. For example, when considering the sorts of stimuli that are included in test questions for something like expressive vocabulary, one wouldn't want to ask the students about items in the household they wouldn't necessarily be familiar with - like a laptop computer.

This is an important factor to keep in mind when designing questions for a test, because it can easily go awry. For example, if one was to design a test looking at expressive vocabulary, one would want to think about other constructs that could potentially influence the responses to the questions.

4. EACH TEST QUESTION SHOULD PROVIDE AN INDEPENDENT PIECE OF INFORMATION

You shouldn't have questions that depend on responses to earlier questions. The idea is that all questions provide an independent piece of information about the ability of the students taking the test. this is important because sometimes people make the mistake of creating questions that depend on a correct response to an earlier question. If that is the case, you doubly disadvantage students who didn't know the answer to the first question. This is the assumption of **local independence**.

5. CONCLUSION

You really want a test which has items that are easy so that even the children who are just starting to become good users of language can get some aspects of the test successfully. You also want items at the other end which are difficult enough so that those who are performing very well, and are advanced for their age, will still encounter some items in a language assessment which will challenge them; which will be just out of their reach.

Talk Together
Supporting Oral Language Development