

**M4C**

Measurement for Change

*Slide Deck 4:*

# Fit for Purpose Selecting Appropriate Tools



*In this slide deck we ask important questions about the suitability of Specific Tools to Specific contexts*

# Let us start with a story

- ❖ A colleague from a sister organization was coming to visit for the first time. This colleague came from a country with a mild climate, that varied little from season to season. The visit was planned for the coldest time of year in the place to be visited.
- ❖ The message went out, “*Make sure you bring a coat*”. But, without the experience of the biting winds and freezing temperatures, the coat that was brought was too thin, and unable to keep the visitor warm.
- ❖ A coat is a universal garment, but to be fit for purpose, it needs to be made to fit the local requirements.
- ❖ Think on.....

# Who is a test designed for?

- ❖ The content of most instruments has been derived from a single culture.
- ❖ The standard administration procedure is also made to fit that context.
- ❖ Thus, each tool, in its original form, will most certainly be limited in the range of behaviours that are seen as important, common, or even recognizable to that culture.

# What makes a tool fit for purpose ?

If the test copyright does not allow you to make any changes, either to

- ❖ Content (pictures or more appropriate vocabulary), or to the
- ❖ Administration format (e.g. oral instead of self-fill)

**THEN IT IS PROBABLY NOT WISE TO USE IT.**

# But .....

If you can adjust the test to the requirement of the local culture this will almost certainly improve the ability to explore with reliability and validity the *within population variability* that you are trying to measure.

# What is Within Population Variance ?

This is a way of evaluating the ability of a tool to pick up meaningful differences.

- ❖ If you are able to make any changes, however subtle, even to the way the tool is administered, then you are

## DEFINITELY

- ❖ Improving the ability of the test to pick up differences between sub-groups in your population.
- ❖ This will impact on the effect sizes that you can measure.

The questions that follow are those most frequently explored in test selection.

The ideas outlined, and answers provided, are based on decades of experience in the field, and the evidence base that this has built up.

**DO NOT MAKE ASSUMPTIONS  
OR**

**YOU WILL NOT BE ABLE TO INTERPRET YOUR RESULTS**

# 1. Does the test selected already exist in the language I will use?

- ❖ Good place to start, but this does not mean you should use it, just that it will save you time if you decide to.
- ❖ Precedence is not enough justification to select a test.



## 2. What is the reliability of the tool in the local context?

- ❖ The test should have a track record of being reliable and responsive to the key developmental question you are interested in
- ❖ A test has also to be reliable and valid within the context in which you intend to apply it.

### 3. How well has the tool been translated, and can further improvements be made?

- ❖ Those with local expertise must read through the translation with a critical eye.
- ❖ The reviewer needs to understand the purpose of the test to ensure key concepts are appropriately expressed.
- ❖ Adjustments may need to be made for particular dialects.

## 4. Is the language of the test actually practical?

- ❖ Sometimes translations (as opposed to adaptations) cannot be made, as there is no comparable vocabulary.
- ❖ This is particularly a problem for mental health forms, where English terms often **DO NOT** translate accurately into other languages.

## 5. What is the current standard process of administration?

Factors that match administration procedures to the local context have the biggest impact on test reliability.

- ❖ Is the tool appropriate to the educational/experience level of respondents, and to their culture of sharing information?
- ❖ Consider, for example, that ALL parental response formats seem to be sensitive to the educational level of the respondent

## 6. Where is the test to be delivered?

Is it to be delivered at the household level or at a community site?

This might impact

- ❖ How many participants can be included
- ❖ How much data can be collected per day to maintain a high quality (given walking distances etc.)
- ❖ The details that can be observed
- ❖ Who will be included, and who missed in the data collection process

## 7. What is the experience/training of the team available to administer the instrument ?

- ❖ Have you allocated sufficient time for the team to learn to administer the test properly ?
- ❖ And for them to work on ongoing inter-rater standardization (checking between the team) ?
- ❖ All tools are prone to variability between assessors, irrespective of the experience of the team, so it is important to work with the team to overcome this
- ❖ And it is important to pay keen attention to randomization issues in the allocation of respondents across the assessment team.

*And finally, but probably most importantly –*

## 8. The meaning of the test itself

- ❖ Does the team actually understand what the tool is about?
- ❖ No tool is really accurate or perfect, so you, and your assessment team, need to understand what it is capable of telling you,
- ❖ and what it does not tell you.

# General Rules when using parental response questionnaires

- ❖ Questions that require a simple yes-no response are the most influenced by response set (parents generally choosing to they feel is the *best* response, thus reducing the variability of scores).
- ❖ Questions requiring ratings benefit from a brief training on what the ratings mean, and a visual prompt sheet that helps the responder select the rating.
- ❖ Examine, and maybe control for, educational level of the respondent in your results.



- ❖ There is a community of experience out there
- ❖ If in doubt, share the question, and there will be a suggestion that can help .....

***Assessment is an  
Exploration***

***Enjoy the Adventure !!!!!***

# Contributors:

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_

*Thank You*