

# How ESTEEM Measures Were Developed



## Motivation

The World Bank's Africa Region Gender Innovation Lab (GIL) and Innovations for Poverty Action (IPA) have partnered to identify which socio-emotional skills are most critical to economic empowerment in Sub-Saharan Africa, and how these vary with gender. In order to examine this question using household surveys, the team examined a wide array of existing resources, and realized the need to adapt measures when possible and validate them for this context.

The Effective Socio-emotional skills To Gain Economic EMpowerment (ESTEEM) tool was developed based on focus groups, cognitive interviews, and iterative psychometric analysis across large samples in Tanzania, Nigeria, and Cote d'Ivoire.

# Criteria

The SES measures needed an instrument accessible to individuals with varying levels of literacy and computer proficiency to ensure its widespread availability and utility. Further, the tool had to be made appropriate for administration by enumerators using offline tablets during home surveys and be contextually relevant with limited game ability and bias. Each measure had to be assessed within five minutes, and a self-report scale and at least one behavioral measure had to be developed for each of the fourteen skills.

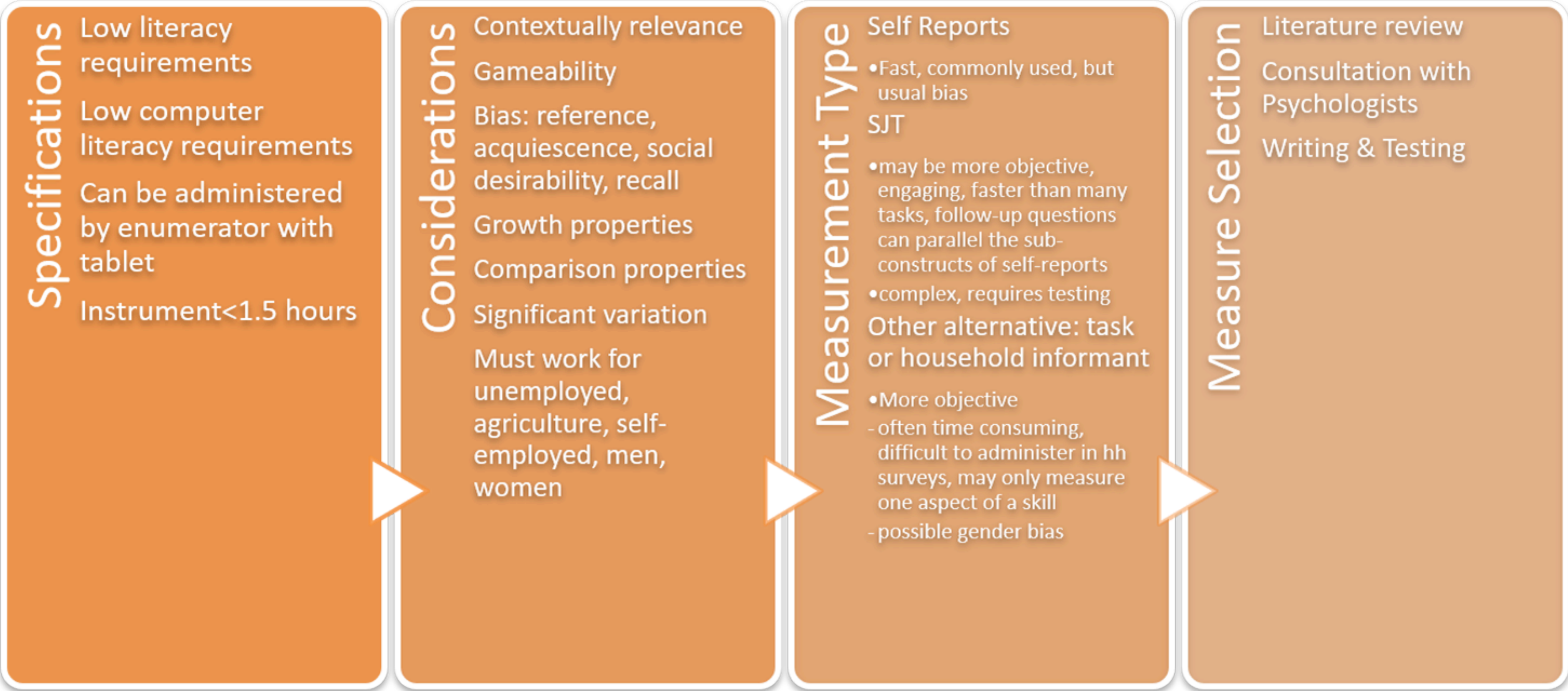
Behavioral measures tested include tasks, household informant interviews, and situational judgment tests. The self-reports were based on existing measures wherever possible, but required extensive adaptation and several rounds of translation to capture each skill.

## Specifications

- User-friendly
- Contextually relevant
- Limited game-ability & bias
- < 5 minutes per measure
- Self-report scale & behavioural measures available for each of 14 skills



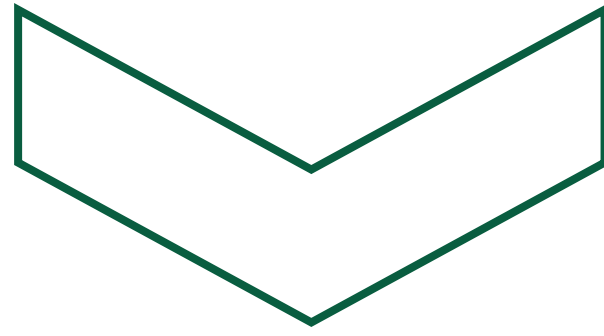
**Figure 1. Measure Development Considerations**



## 4 Key Steps to Content Validity

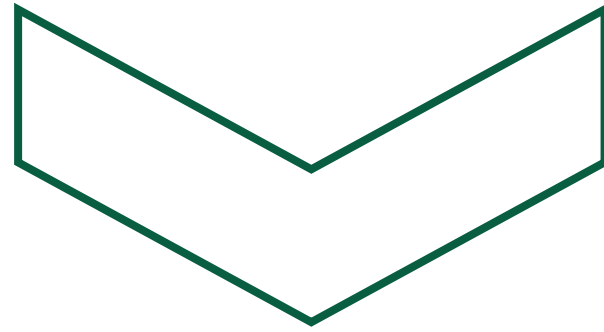
- 1** Focus group discussions and interviews with entrepreneurs were conducted to gather information about the fundamental skills required for success in each sector, and how skills manifested as behaviors. This process also shed light on the differences in key skills between men and women and highlighted fundamental skills for household discussions.
- 2** Cognitive interviews were conducted to verify the understanding of participants. Interviews included questions like “What did you understand with this question,” or “Is there a correct answer?” They were open-ended, critical thinking-based, and played a key role in enumerator training and feedback, timing, order, and engagement.
- 3** The information underwent intensive translation and back-translation processes to ensure accuracy. This process included translation ratings, receiving linguistic feedback from psychologists, and a final translation evaluation. Key words that can create translation mishaps were identified and clarified.
- 4** The content was reviewed by subject matter experts including psychologists, curricula experts, educational assessment specialists, international organizations, and field coordinators.

## Figure 2. 4 Key Steps to Content Validity



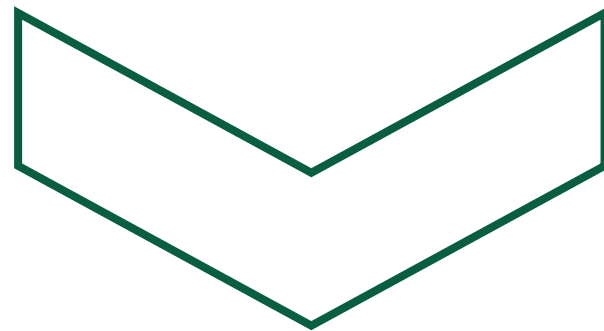
### Focus Group Discussions/Interviews with Entrepreneurs

- ‘Which skills are key for success in each sector, each gender, and household discussions?’
- ‘Which behaviours are associated with each skill?’



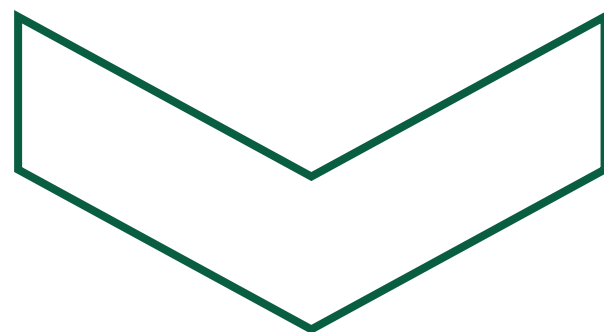
### Cognitive Interviews

- “What did you understand from this question?”
- “Is there a correct answer?”
- Key for enumerator training, feedback, timing, order, engagement



### Intensive Translation Process

- Translation & back-translation
- Translation Ratings, Linguistic Feedback by Psychologists, Final Translation
- Glossary items

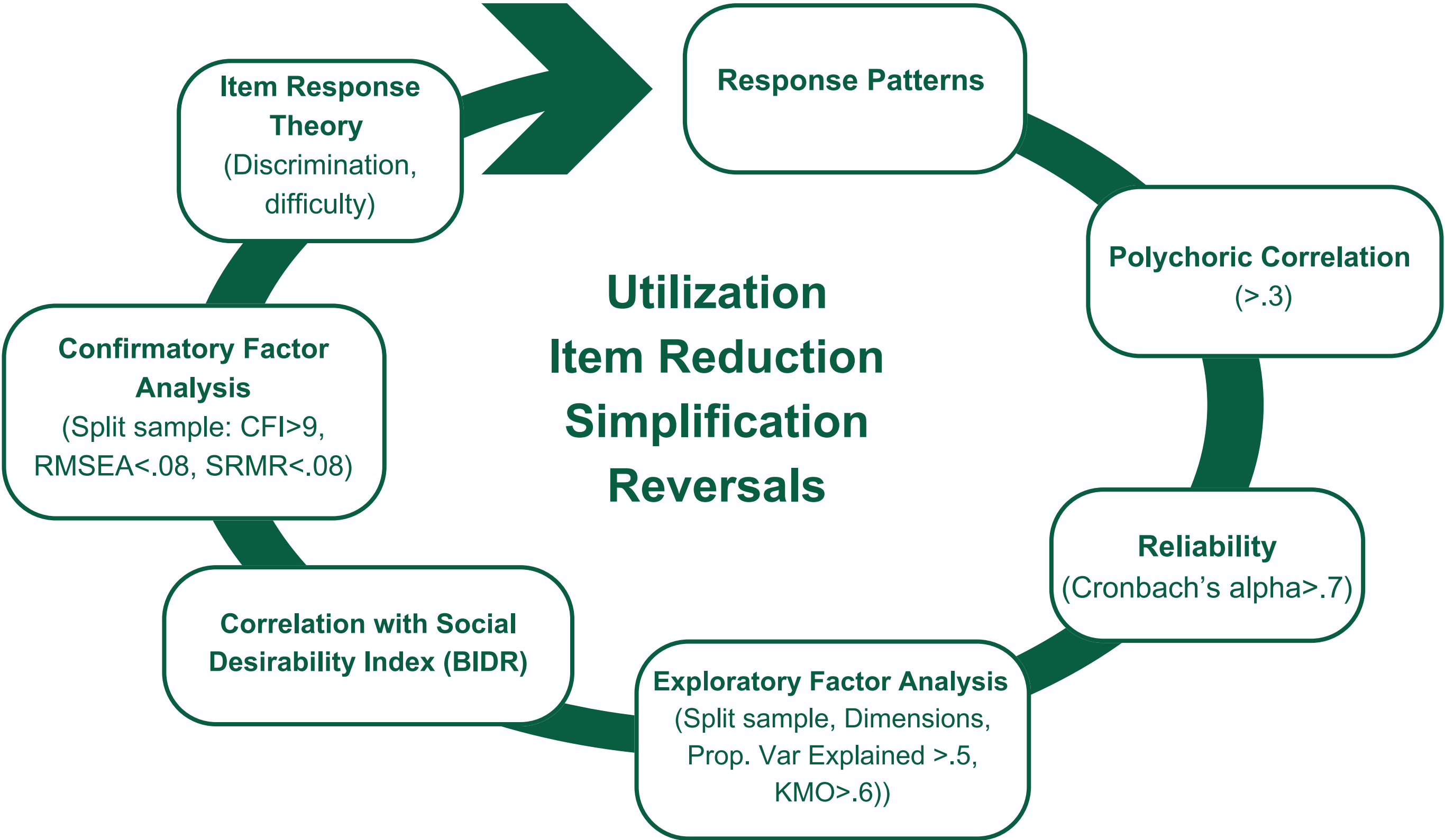


### Review from Subject Matter Experts

- Psychologists
- Curricula Experts
- Education Assessment Experts
- International Organizations
- Field Coordinators

# Figure 3. Psychometric Analysis Conducted in Each Setting

The team adopted an iterative approach for improving the measures. Across various rounds of piloting, baselines, and follow-ups, the team ran cognitive interviews and examined the psychometric properties of the measures: response patterns, correlation, reliability, exploratory factor analysis, correlation with a social desirability index, confirmatory factor analysis, and item response theory. This allowed us to test for understanding and remove items that were not aligned with the measure of a skill. Despite several attempts, reversed items and regular items could not be used in the same scale.





# Recent Progress and Next Steps



This psychometric analysis has been conducted in Tanzania, Nigeria, and Cote d'Ivoire, and self-reported measures have also been tested in Congo and Rwanda. The team continues to conduct a range of **survey experiments** investigating the length of scales, the gender of individuals in situational judgement tests, the gender of the household informants, and the effectiveness of a unipolar scale as opposed to a bipolar scale.