Insight2Impact

Advancing Financial Inclusion

Developing valid and reliable survey scales

October 2017

cenfri

FMT FINMARK TRUST

Author

Grant Robertson, Data4FI

About insight2impact

Insight2impact | i2i is a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design.

i2i is funded by the Bill & Melinda Gates Foundation in partnership with The MasterCard Foundation.

For more information:

Visit our website at **www.i2ifacility.org.**
Email Mari-Lise du Preez (i2i's Partnerships Manager) at **mari-lise@i2ifacility.org.**
Call us on **+27 21 913 9510.**

Contents

Introduction

The purpose of this note is to describe a process that can be used to design valid and reliable financial inclusion measuring instruments. It is often the case that important measurement dimensions are excluded from financial inclusion surveys. This could be due to an oversight on the part of the research team, lack of a suitable measure or a lack of resources in survey teams to develop suitable measurement frameworks and subsequent measuring instruments. Whatever the reason, the impact of this is that financial service providers, policymakers and regulators will miss key information that can assist them in making better strategic decisions. In such cases, it is necessary to develop new instruments that are valid and reliable measures of the phenomena of interest. Failure to design measuring instruments scientifically will result in poor-quality data and ultimately poor decision-making. Having well-developed survey modules and scales that reflect relevant measurement frameworks can accelerate the provision of appropriate data to inform financial inclusion strategy.

As surveys are the primary research method used by financial inclusion researchers, the questionnaire remains one of the most important data collection tools. Often, these questionnaires have psychometric problems or are not subjected to any psychometric scrutiny[1]. This note presents a systematic process that can be followed when devising new measuring instruments, be that an entire questionnaire, a module, scale or even a single item.

Before we present this process, it is important to highlight some of the benefits that are derived from systematically building and including new scales and modules in surveys.

1 Hinkin, Tracey, Enz, 1997

The benefits of scale development

The development and dissemination of scientifically developed and standardised measuring instruments, in the form of questionnaires, modules or scales, have the following benefits:

- They introduce new and important dimensions to surveys that should be of universal concern to financial inclusion researchers and policymakers.
- They are cost-effective to deploy once developed.
- They are flexible and can be deployed in several ways, either as a standalone survey or as part of an existing survey.
- They provide a quick way of introducing something innovative and fresh into existing financial inclusion surveys.
- They are largely standardised but can be customised for local contexts.
- They are credible, as they've been validated and imbedded within a measurement framework and/or theory.
- They allow for better cross-country and cross-survey comparison.
- They can benefit from network effects, as other researchers can give their input and feedback so that the instrument can be improved.
- They can produce shorter questionnaires. Questionnaires with too many items can create problems with respondent fatigue or response biases. Shorter questionnaires are cheaper to administer and thus increase survey sustainability.

A lack of focus on proper instrument development that ensures a valid and reliable measure of a theory or measurement framework threatens our understanding of economic development phenomena and reduces our impact on important development outcomes. A key objective of the i2i facility is to create relevant measurement frameworks[2] that are operationalised through survey modules and scales that can be deployed in existing or new surveys. It should be noted that not all measurement framework data can and should be collected through surveys when other data sources are more suitable.

2 http://i2ifacility.org/resources/publications

Definitions of key terms

As in any scientific endeavour, we need to follow a systematic process when building a measuring instrument. Before we present the scale development process, let's first clarify the meaning of some important terms.

Theory

A theory is a scientifically acceptable body of principles offered to explain phenomena. Theories are fact based (not mere guesses) and are presented in such a way that they are testable. As such, a conceptual framework needs to be presented and hypotheses need to be formulated for testing. In social sciences, the phenomena to be explained is usually a behaviour. A theory also usually makes some predictions about future behaviours.

Model

Models are derived from theories in an attempt to help simplify and test them. They are more pragmatic and represent reality, as opposed to being abstract like a theory is.

Measurement framework

A measurement framework combines theory and data to describe a condition necessary to achieve an objective. It consists of an indicator or set of indicators populated by data. The theory explains why the condition is important for the objective and why the indicators are valid proxies for the condition and any changes therein[3].

3 http://access.i2ifacility.org/Publications/i2i%20MFW%20Note%201%20-%20Intro%20to%20measurement%20frameworks_Digital.pdf

Domain

A domain is an overarching label for a set of related constructs. For example, mental tests are designed to measure different cognitive domains, such as verbal fluency, memory, arithmetic skills and spatial visualisation. Financial inclusion, for example, has domains such as financial health, access, uptake and usage. Domains can be represented by a single or multiple constructs and scales.

Construct or latent variable

A construct is an unobservable entity. It can only be measured indirectly through questions or statements. Financial health is a construct in the same way that gravity and temperature are. It cannot be measured directly but rather through items that are believed to represent the construct.

Measuring instrument

A measuring instrument is a device used to collect information. In this note, a measuring instrument refers to a questionnaire, module, scale or test.

Module

A survey can be made up of several modules, such as a demographic, hobbies and interest, financial health, financial products module and so forth. Each module could be a collection of separate but related questions (such as age and gender in a demographic module), a single scale (such as quality of life) or a number of scales that each measures various dimensions of culture (such as power, individualism, masculinity, uncertainty and avoidance). So, a module is a part of a survey that measures a specific dimension of interest. A module should ideally only measure a single domain. It should be able to slot into an existing survey without a need to redesign the entire survey. The module can contain a single scale or multiple scales, or it can just be a set of individual questions that measure something of interest.

Scale

A scale is a composite measure of several items (questions/statements) that have a logical or empirical structure among them. This allows us to measure the direction and intensity of a construct. A single module might have several scales if the dimension you are interested in is multi-faceted. A scale is made up of many items. Examples of scales include financial health, value for money, trust, loyalty, service quality and quality of life. A key characteristic of a scale is that it is best measured by a number of items that reflect various aspects of the construct or dimension of interest.

Item

An item is a single question or statement in a questionnaire and is usually the smallest building block of a questionnaire, module, scale or test.

Response scales

In some instances, we are interested in phenomena that have no underlying latent structure, such as frequency or recency of financial product use. These types of items are usually presented in a way that the response options define the metric of interest. In such cases, careful selection of the response options in terms of the type of scale, number of options, anchor points, midpoints, the range of each option, etc. is important, as it will inform the resulting metrics. A response scale thus has one item (question or statement) that is linked to a set of response options. It is thus not a scale, but can be a module or question in a survey.

We will now proceed to discuss a systematic approach to developing a scale.

The diagram is an illustration of how response scales, items, scales and modules build on one another to create a questionnaire.

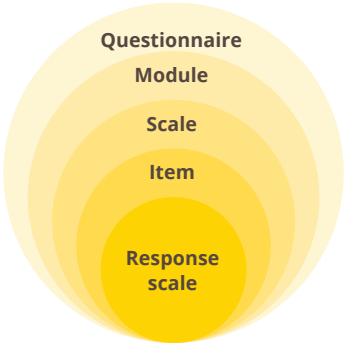


Figure 1
Design hierarchy for a questionnaire

> Examples of scales include financial health, value for money, trust, loyalty, service quality and quality of life. A key characteristic of a scale is that it is best measured by a number of items that reflect various aspects of the construct or dimension of interest.

Scale development process

This section deals specifically with the process of developing a scale, as opposed to developing single-item measures such as frequency of use, which is an indicator of how often a financial device is used. Frequency isn't something that we need to measure using multiple items. Modules can be made up entirely of these types of indictors and do not require a full-scale development cycle. If the goal is to develop a scale (such as financial health, wellbeing or literacy), then the 13-step process described below is applicable. The steps are not discrete or linear; they overlap, and you can return to previous steps as new learnings emerge.

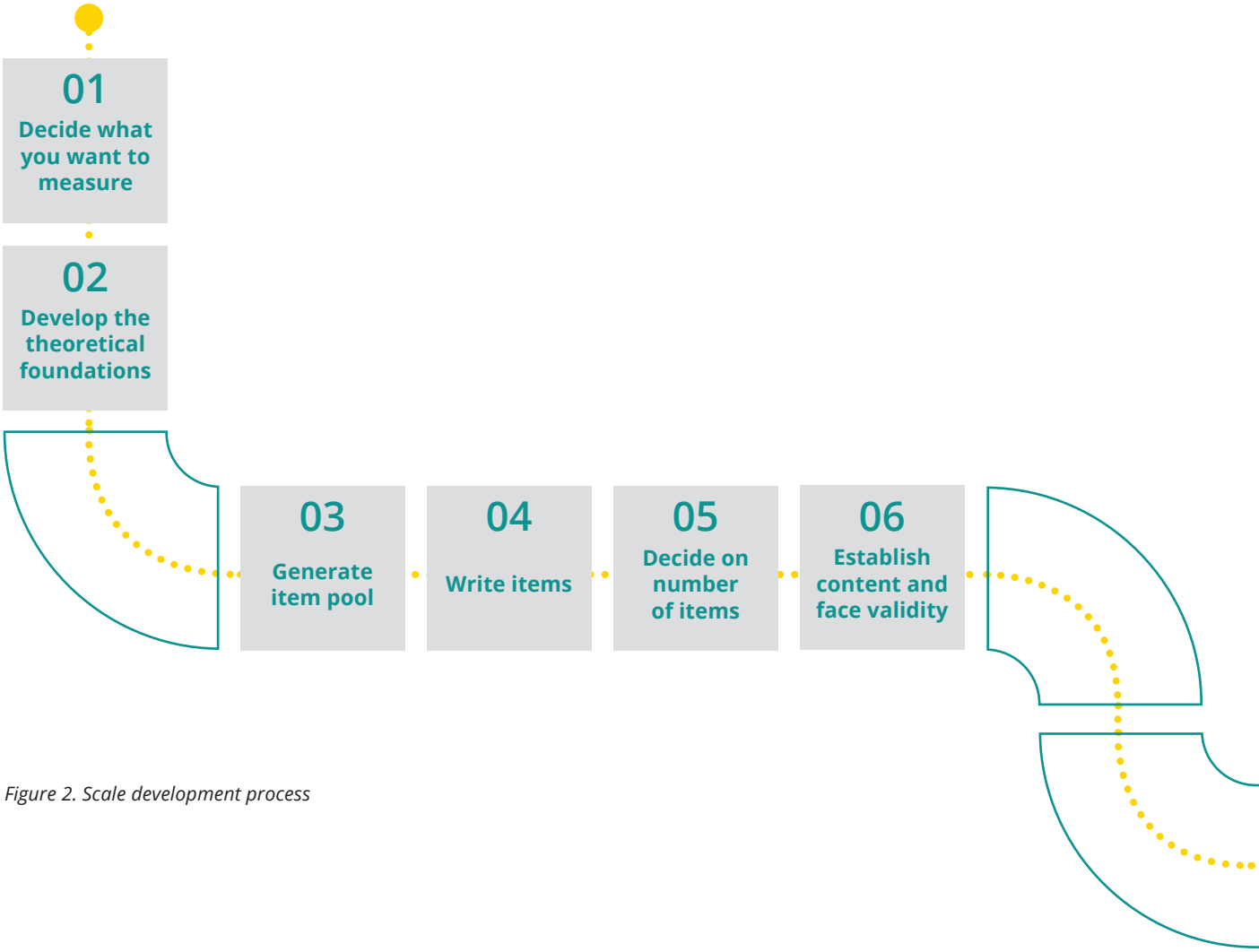


Figure 2. Scale development process

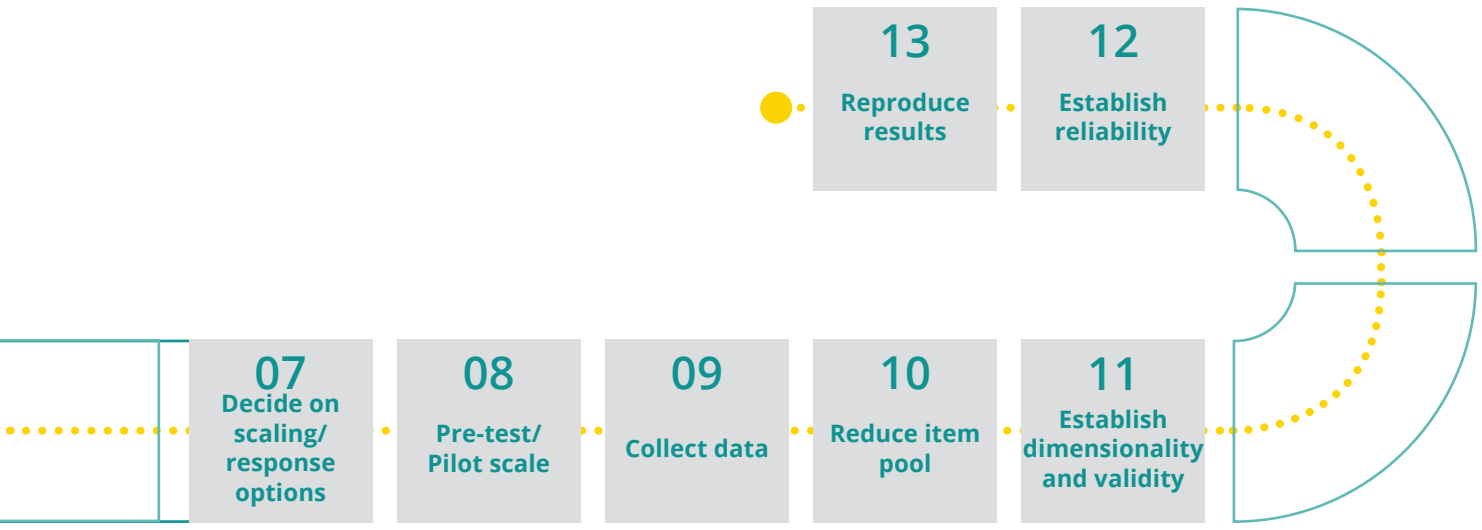
1 Decide what you want to measure

The first step in developing a new scale is to decide what you want to measure. The decision as to what to measure usually arises from a realisation that there is some gap in the current measurement status quo. For example, i2i's needs and usage measurement framework was developed upon the realisation that the focus of most financial inclusion metrics was on access, uptake and products, as opposed to financial needs and usage. Similarly, the development of financial health scales evolved through the realisation that there were limitations in the measurement of financial literacy and capability.

Key considerations during this phase include:

- Be clear on what you want to know, what you know and what you don't know.
- Determine whether the dimension of interest should be measured through a survey, or whether the data can be obtained from another source, such as transactional or regulatory data.
- Ensure that the domain is new and hasn't been developed and tested before.
- Determine whether there is an evidence-based hypothesis of why this measure will be of value in moving the financial inclusion dial.
- Be clear on why existing measures are inadequate.

The output of this phase should be a clear concept note detailing the importance of measuring the new construct, defining it clearly, showing how it is different from existing measures and how this relates to an improvement in the measurement of financial inclusion.



2 Develop the theoretical foundations

The second stage of scale development is the development of a well-articulated theoretical foundation that delineates the content domain for the new measure. It is usually impossible to measure the entire domain of interest, and as such the researcher should ensure that the sample of items selected adequately represents the domain of interest[4]. The researchers should define the construct and specify what is included and excluded.

Even if there is no available theory, you should try to specify a tentative theoretical model that will serve as a guide to developing the scale. This should, at minimum, include a well-formulated definition of the construct as well as a set of hypotheses.

Key considerations during this phase include:

- Identifying all the main dimensions/constructs of interest
- Specifying exactly what will and will not be measured
- Naming the constructs, such as financial health and financial capability
- Providing precise definitions of the constructs you've selected to measure
- Proposing a conceptual or theoretical model to explain the phenomena of interest
- Specifying the underlying relationships between the constructs
- Specifying hypotheses that can be tested
- Specifying the populations and settings in which the construct will be relevant
- Specifying the types of metrics that will be produced once the data has been collected

4 Hinkin, 1998

3 Generate an item pool

Once the domain has been fully researched and articulated, there are two approaches to identifying items: the deductive approach and the inductive approach. The deductive approach assumes that it is sufficient to generate items after a thorough review of the literature and theoretical fundamentals has been considered. This should only be attempted by researchers who have a good understanding of the phenomena under investigation. The inductive approach is used when there is no or little theory to work from. Sometimes researchers will just start generating items, or they will gather information from the target population to help identify items for the scale[5]. Qualitative techniques (such as group or personal interviews) are conducted, and the data content is analysed to assist in identifying relevant items.

The item pool can be generated in a variety of ways, including:

- A literature review of related research and questionnaires
- Expert interviews
- Some form of qualitative research, such as observations, focus groups, in-depth interviews and ethnography

Have a look at our Questionnaire Design Tool, which can be used to see how questions related to the same construct can be asked differently[6].

> “Even if there is no available theory, you should try to specify a tentative theoretical model that will serve as a guide to developing the scale. This should, at minimum, include a well-formulated definition of the construct as well as a set of hypotheses.”

5 Hinkin, 1998
6 http://access.i2ifacility.org/Community/Questionnaire_design_tool

4 Write the items

Once the item content has been chosen, the actual items need to be written. Good question-writing is an art and a science. Please refer to our implementation guide for guidance on how to write good questionnaire items[7]. In summary, the items should be short, precise, clear, positively phrased and non-ambiguous. It is also important not to mix behavioural questions with attitude or outcomes questions.

The items should be short, without sacrificing clarity, and written at an appropriate reading level for the target audience, which will differ depending on the context.

Every effort should be made to ensure that each item reflects the underlying construct you are measuring. This can be done by continually referring to the theoretical underpinnings of the framework. A good starting point is to look at existing scales and items that have been tested previously.

5 Decide on the number of items

The number of items is important for several reasons. It is always best to ask fewer questions, as it usually produces better-quality data. However, when developing scales, it is usually good practice to eventually have at least three to five items for each construct being measured. Statistically this allows for better estimation of validity and reliability. The variable-to-sample-size ratio is an important consideration for some statistical techniques like factor analysis or principal components analysis. When in the development phase, it is likely that you will have three or four times the number you ultimately want, so you might start off with 12 items for a scale in the hope of ultimately having three or four good indicators.

In theory, the items selected should be done randomly from a pool of items that represent the universe of items relating to the construct of interest.

7 http://access.i2ifacility.org/Community/Implementation_Guide/Questionnaire-design

6 Assess content and face validity – review the item pool

The items should then be assessed for content validity and face validity. Content validity deals with whether the researcher has a sufficient and appropriate sample of items to adequately represent the construct of interest. This process precedes any data collection, as it will help to save a significant amount of time and money by testing only items that have been properly assessed for face and content validity.

Content validity addresses the issue of whether all facets of the construct of interest are being measured. For example, if we are measuring financial product usage, an assessment of content validity would highlight any important dimension that might be missing. This process usually starts off with thorough literature reviews of the construct of interest. Content validity is usually assessed by subject matter experts. Content validity is not measured but rather is "assured" by the expert input.

Face validity is the most basic form of validity testing and the most common. The researcher or a content domain expert simply looks at the contents of the instrument and decides whether it looks like it will measure what it is supposed to measure. Another method would be to ask a sample of potential respondents to score the instrument in terms of its perceived suitability to measure the content area of interest. An alternative approach would involve giving each item to the target population with different definitions, after which they are tasked to indicate which definition they think best describes the item[8]. Face validity establishes what the instrument looks like it is measuring, not what it is actually measuring.

> "Every effort should be made to ensure that each item reflects the underlying construct you are measuring. This can be done by continually referring to the theoretical underpinnings of the framework."

8 Hinkin, 1998

7 Decide on the response options

Item response options are another important consideration. It is important that the scale produce enough variation among the respondents. In demand-side surveys, Likert scales are popular and use either five or seven points. In some markets, this can be reduced to fewer if the conditions warrant it. Other possibilities are semantic differentials and binary options. Please refer to our implementation guide for more on response/answer options.

Key considerations for developing response scales:

- The response scale should be developed simultaneously during the item writing phase.
- Consider the number of points, as this impacts on the reliability and sensitivity of picking up changes in the dimensions you are interested in.
- More points are desirable if the construct under investigation allows for easy discrimination between levels or strength of the phenomena.
- Consider the setting of appropriate start and end anchors for the scale.
- Scales points should not overlap.
- Consider whether the scale should be balanced or unbalanced, i.e. an equal number of negative and positive points and whether there is a mid-point (even or uneven):
  - An even number of points is good when you want to force people to make a decision that commits them to one side of the scale or another.
  - An odd number of points is good when you're interested in when people will take a middle-of-the-road stance, when they will say "don't know", etc.
- Consider the implications of having *nominal, ordinal, interval and ratio scales* for statistical methods used.
- The response scales should be intuitive and relevant to the cultural and literacy levels of the population of interest.

8 Pre-test/pilot the scale

Once the items have been selected and written and a response scale has been attached, they are ready to be pre-tested. At this stage, *cognitive interviews* could be used to assess the quality of the items, or a standard pre-test of the instrument on a small sample of the population could be undertaken. This provides an additional layer of filtering that can weed out poor items. See our implementation guide for more on pre-testing and piloting.

After pre-testing and piloting of the instrument, it can be administered to the entire sample of interest using standard survey practices.

9 Collect data

Once the items have been refined through the pre-testing phase, they are ready to be administered to a sample of the population of interest. At this stage, it is critical that you have a high-quality sampling plan with an adequate sample size to be able to conduct the necessary statistical analysis. If the items are to be administered in languages other than the development language, a back-translating process needs to ensue. The objective of this phase is to collect the data that will allow us to test the psychometric properties of the scale. This means that, in addition to the new scale, existing scales known to be related and unrelated to the new scale need to be included so that the various types of validity can be assessed.

> “It is critical that you have a high-quality sampling plan with an adequate sample size to be able to conduct the necessary statistical analysis.”

10 Reduce item pool

Once the data has been collected and cleaned, it is ready to be analysed. The focus during this phase is usually to select the best items from the tested pool. Usually an item reduction process is followed, using a variety of techniques, which allows us to eliminate items that do not contribute additional information. Doing so also increases the internal consistency or reliability of the scale, usually measured by Cronbach's Alpha. If item reduction isn't required (such as when you are relying on a single item that isn't actually a latent construct, such as frequency), the focus is on whether the measure adds additional insight into the phenomena of interest or discriminates meaningfully in some way.

Some common statistical techniques used include:

- **Principal components analysis (PCA):** PCA is a procedure for identifying a smaller number of uncorrelated variables, called "principal components", from a large set of data. The goal of PCA is to explain the maximum amount of variance with the lowest number of principal components. You can use PCA to reduce the number of variables and avoid multicollinearity, or when you have too many predictors relative to the number of observations[9].
- **Factor analysis:** Factor analysis (including confirmatory/structural equation models) is a method for explaining the structure of data by explaining the correlations between variables. Factor analysis summarises data into a few dimensions by condensing many variables into a smaller set of latent variables or factors[10]. Factor analysis is usually used to test a theory of latent constructs, whereas PCA is usually used as a data reduction technique. Components aren't latent; they are linear composites of the correlated variables.
- **Cronbach's alpha:** Cronbach's alpha is a measure of internal consistency, that is, how closely related a set of items are as a group. It is a measure of scale reliability. A "high" value for alpha does not imply that the measure is unidimensional – that is assessed through factor analysis[11].
- **Correlations:** One could also inspect the correlation matrix among the variables and remove items that correlate highly with other items.

9 http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/multivariate/principal-components-and-factor-analysis/what-is-pca
10 http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/multivariate/principal-components-and-factor-analysis/what-is-factor-analysis/
11 https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean

11 Establish dimensionality and validity of the scale

Surveys can be reproduced by the original researchers or validated by independent researchers as part of the replication process. However, before reproducibility and replication of the entire survey are attempted, a significant amount of time should go into other methods of validation, namely assessing the survey instrument for validity and reliability. Validity and reliability are discussed in more detail below.

12 Validity and reliability

Issues of validity and reliability are often overlooked by practising survey researchers.

This is usually because:

- The work is bespoke and ad hoc in nature and will not be repeated.
- It is time-consuming.
- It can be expensive.
- There is a lack of sufficient resources or skills.

Financial inclusion demand-side surveys tend to be run at regular intervals and typically measure the same core constructs and indicators over time. This makes them ideal candidates for an in-depth assessment of validity and reliability. As they are run regularly, they are self-reproducing, which is an inherent validation method.

A. VALIDITY

Put very simply, validity assesses whether the instrument measures what it says it measures. There are different types of validity, and all of them should be assessed. Instruments are seldom 100% valid, and some forms of validity are more important or less important, depending on what you are trying to achieve.

Broadly, we are interested in three categories of validity:

- **External validity,** which is the extent to which the results of a study can be generalised from a sample to a population. Establishing external validity for an instrument, then, follows directly from sampling. External validity can be further split into population and ecological validity.
  - Population validity is how well the results can be extrapolated to the population of interest
  - Ecological validity is an assessment of how the survey environment might influence the results. So, it speaks to the extent to which the results can be generalised to "real" world settings – the results aren't artefacts of the survey conditions.
- **Internal validity** refers to the extent to which we can ascribe the outcomes of the research to the independent variables, so how confident you are about your conclusions as to the cause-and-effect relationships.
- **Instrument validity** refers to the appropriateness of the content of an instrument. So, does it measure accurately what you want to know? There are many ways in which instrument validity can be assessed.

> "Instruments are seldom 100% valid, and some forms of validity are more important or less important, depending on what you are trying to achieve."

12 Nunnally, 1978

This note will briefly introduce four types of instrument validity that need to be considered when testing scales.

1. **Content validity**
   As defined on Page 13

2. **Face validity**
   As defined on Page 13

3. **Construct validity**
   A construct is a representation of something that does not exist as an observable dimension of behaviour; and the more abstract the construct, the more difficult it is to measure[12]. Construct validity deals with the complex issues of how consistent a construct (service, financial health, poverty, etc.) behaves relative to the theory from which it was developed. Construct validity looks at the pattern of interrelationships among variables. Unlike content validity, construct validity can be tested statistically. Construct validity can be assessed by looking at the extent to which the constructs correlate (convergent) or do not correlate with each other (divergent). A number of statistical techniques can be used to do this, but they will not be discussed in this note.
   - ***Convergent validity***
     This is the extent to which our construct correlates positively with other items known to measure similar things and have a high portion of shared variance.

   - ***Divergent validity***
     This is the extent to which our construct correlates negatively with different constructs and the extent to which they are distinct from each other.

4. **Criterion-related validity**
   Criterion validity is the evaluation of the extent to which a measure relates to a particular outcome. This can pertain either to measurements taken at the same time or to some measure taken in the future.

   - ***Concurrent validity***
     Concurrent validity is when a new measure correlates positively to an outcome measured at the same time. For example, a scale that measures financial health should correlate with a measure of poverty taken at the same time.

   - ***Predictive validity***
     Predictive validity is when a new measure correlates positively to an outcome measured at some time in the future. A financial health scale should correlate positively with a reduction in national indebtedness.

B. ESTABLISH RELIABILITY

Reliability of a scale is the ability of the scale to produce the same results when administered on different occasions. For example, if you weigh yourself on a scale and step back on it 10 minutes later, you should get the same reading.

There are a number of ways to assess reliability, each of which estimates reliability in a different way. They are:

1. **Test-retest reliability**
   Test-retest reliability methods are used to assess the consistency of a measure from one period to another. It is expected that there will be a high positive correlation between the results taken at different times.

2. **Parallel-forms reliability**
   Parallel-forms reliability is used to assess the consistency of the results of two scales constructed in the same way from the same content domain. For example, if you have 20 items measuring financial health you can create two scales with 10 items each and administer them to the same respondents. You would expect the two scales to correlate highly if they are reliable.

3. **Internal consistency**
   Internal consistency methods are used to assess the consistency of results across items within a scale.

   - ***Cronbach Alpha***
     For internal consistency, Cronbach's Alpha is usually reported, which is the average of all possible split-half estimates. Inter-item correlations, item total correlations and split-half methods are also used to assess internal consistency. Generally, the more items you have the higher Cronbach's Alpha will be[13].

   - ***Split half reliability***
     Split half reliability is another type of internal reliability measure. A scale can be split into a number of parts and each part compared to the other. If the parts correlate highly, the scale is thought to be internally consistent.

   It should be noted that reliability is necessary but not sufficient to show validity. Once you have a scale that has shown to be valid and reliable, the final step in the process is to ensure that your research is reproducible.

13 https://www.socialresearchmethods.net/kb/reltypes.php

13 Reproduce results

One of the foundational components of the scientific method is the idea of reproducibility. For research to be considered valid, it must be replicated. If a research study is replicated using the same procedures and the results found to be similar to the original research, this provides support or validation for the original findings[14]. As such, it gives us confidence to use the research findings to make informed decisions.

Reproducibility vs repeatability

Reproducibility should be conceptually distinguished from repeatability.

- Repeatability is where the same research team repeat their research, using the same protocols to test and verify their results.
- Reproducibility is tested by a replication study, which must be entirely independent and generate identical/similar findings to that of the original research.

If the original sample size was large enough, it would be possible to split the sample and conduct the same process on the split sample to verify the psychometric properties (validity, reliability) of the scale.

Replication study

A replication study involves repeating a study, using the same methods but with different subjects and researchers. The researchers will apply the existing theory and methods to the same or new situations to determine generalisability to different populations and sub-populations.

The main determinants of this type of study include:

- To test that results are reliable and valid
- To determine the role of extraneous variables
- To apply the previous results to new situations
- To inspire new research, combining previous findings from related studies[15]

Only once results have been reproduced and replicated can we claim that our measures are valid and reliable.

14 Rand & Wilensky, 2006
15 https://explorable.com

Conclusion

Good research begins with good measurement. This can be achieved if a systematic and scientific process is followed. A perfectly designed and executed survey will not produce good data if the underlying measurements are not scientifically sound. Good measurement will deepen our understanding of financial inclusion and will ultimately result in better policy and strategy decisions.

Bibliography

Cronbach, L.J. (1990). Essentials of psychological testing. 5th edition. Harper Collins.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires [electronic version]. Retrieved [insert date], from Cornell University, School of Hotel Administration.

Hinkin, T.R., Tracey, J.B., Enz, C.A (1997) Scale Construction: Developing Reliable and Valid Measurement Instruments. Cornell University School of Hotel Administration, The Scholarly Commons.

Hinkin, T.R. 1995. A review of scale development practices in the study of organizations. Journal of Management 21(5):967-988

Huysamen, G.K. (1987). Psychological measurement. Academica.

http://access.i2ifacility.org/Publications/i2i%20MFW%20Note%201%20-%20Intro%20to%20measurement%20frameworks_Digital.pdf.

https://explorable.com

http://access.i2ifacility.org/Community/Questionnaire_design_tool/

http://www.ibtimes.com/scrutinizing-scientific-method-researchers-massive-open-access-study-fail-replicate-2071483

https://temme.wiwi.uni-wuppertal.de/fileadmin/_migrated/content_uploads/fuchs_diamantopoulos_2009.pdf

http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/multivariate/principal-components-and-factor-analysis/what-is-pca/

https://stats.idre.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/

https://www.socialresearchmethods.net/kb/reltypes.php

Murphy, K.R. & Davidshofer, C.O. (1991). Psychological testing: principles and applications. 2nd edition. Prentice Hall.

Nunnally, J. C. (1976). Psychometric theory (2nd ed.). New York: McGraw-Hill

Rand, W., Wilensky, U. (2006). Verification and validation through replication: a case study using Axelrod and Hammond's ethnocentrism model. North American Association for Computational Social and Organization Sciences (NAACSOS), South Bend.

How to find us:

Get involved. Contact us.

Grant Robertson
E: grant@i2ifacility.org

Bobby Berkowitz
E: bobby@i2ifacility.org

Damola Olowade
E: damola@i2ifacility.org

Shirley Jeoffreys
E: shirley@i2ifacility.org

+27 11 315 9197
i2ifacility.org

@i2ifacility
/insight2impact
/insight2impact
/i2ifacility

Established by

cenfri

FMT FINMARK TRUST

Sponsored by

BILL & MELINDA GATES foundation

mastercard foundation