



CHOOSING OR DESIGNING EFFECTIVE SURVEY INSTRUMENTS:TIPS TO APPLY

Assoc. Prof. Dr. Norsiah binti Mat

Postgraduate Talks

SBM-COB

Kedah Room

9.00AM-11.00AM

7 November 2017

Outline

- Using the right instrument -- Develop, adopt @ adapt an instrument
- General Rules for Writing Survey Questions
- Instrumentation Criteria for existing instrument
- Develop, adopt @ adapt an instrument
- Choosing **Survey Response Scales**
- **Pre-testing & Pilot Testing Questionnaire**
- Steps in Developing your own Research Instrument
- Questions and Comments

Measurement instrument

- Refers to various methods through which a researcher obtains data from respondents for his research work.
- in order to make sense out of the data collected, each variable must be measured in such a way that its magnitude or quantity can be clearly identified
- **In general, the following data collection methods work for qualitative research:**
 - Document review.
 - In depth interviews.
 - Observation methods.
 - Focus group discussion
- **Quantitative research data collection methods include:**
 - Surveys with closed-ended questions.
 - Clinical trials/experiments.
 - Extracting data from computer and information systems.
 - Observing, counting and recording events that are well defined (for example, counting the number of people who come into a restaurant).

Measurement instrument



General Rules for Writing Survey Questions

- The ideal question accomplishes three goals:
 - It measures the underlying concept it is intended to tap
 - It doesn't measure other concepts
 - It means the same thing to all respondents
- The following rules help to accomplish this:
 - **Avoid technical terms and jargon** -words used in surveys should be easily understood by anyone taking the survey.
 - **Avoid Vague or Imprecise Terms** - usually, it's best to use terms that will have the same specific meaning to all respondents. E.g.,
it's not clear what you get when you ask "How important is it that a candidate shares your values?" You might get a more consistent answer if you asked: "How important is it that a candidate shares your religious values?"

- 
- **Define Things Very Specifically** – E.g., don't ask: "What is your income?" A better question would be specific and might ask: "What was your total household income before taxes in 2016?"
 - **Avoid Complex Sentences** -- Sentences with too many clauses or unusual constructions often confuse respondents. How easy will it be for a typical person to answer: "Do you think the increase in the rate of immigration, controlling for the economy, is higher or lower than the increase in the rate of crime in your area?"
 - **Provide Reference Frames** -- make sure all respondents are answering questions about the same time and place. For example, if you ask: "How often do you feel sad?" it's better to provide a reference frame: "How often have you felt sad during the past week?"

- **Avoid Double-Barreled Questions** --questions should measure one thing. Double barreled questions try to measure two (or more!) things. For example: “Do you think the Prime Minister should lower taxes and spending.”
- **Answer Choices Should Anticipate All Possibilities** --if a respondent could have more than one response to a question, it's best to allow for multiple choices. If the categories you provide don't anticipate all possible choices, it's often a good idea to include an “Other-Specify” category.
- **If You Want a Single Answer, Make Sure Your Answer Choices Are Unique and Include all Possible Responses**
 - If you are measuring something that falls on a continuum, word your categories as a range. For example What punishment should this person receive: No punishment, Punishment not including jail time, Up to five years in prison, From five years to ten years in prison, From ten years to 20 years in prison, More than 20 years but less than life in prison, Life in prison, or the death penalty?”

- **Avoid Using Leading and Loaded Questions** – a leading question is a question that suggests or implies certain answers. E.g., “Do you believe that US should immediately withdraw troops from the failed war in Iraq?” and a loaded question is emotionally charged. E.g. “Do you support or oppose the GST?.”

Develop, adopt @ adapt an instrument

- Poor instrument selection adds noise and error to your research
- "Development of new instrument is a complex and difficult process that requires considerable training. Therefore, we recommend that **you make certain no suitable instrument is available before developing your own**" (Gall, Gall, & Borg, 2003, p. 216).
- The development and validation of a new instrument could be an entire thesis in and of itself.

- 
- Using a pre-developed instrument to measure the key variables in your study has a number of advantages:
 - saves the time and energy necessary to develop a new instrument
 - since the instrument has already been validated, you do not have to spend the time and energy necessary to validate a new instrument.
 - If you use an instrument developed by other researchers, then you know that your study fits well into the worldwide knowledge base on your variables of interest.

Selecting an Existing Instrument

- **Address the following questions:**
 - Do instruments already exist that measure a construct the same or very similar to the one you wish to measure?
 - **Clearly define construct or concept** –some concepts are relatively concrete (e.g., age, sex, number of children)
 - Some concepts such as loyalty, job involvement, OCB and the likes present greater problems in terms of definition and measurement.
 - **Operational definition of the construct** – gives meaning to a concept by specifying the activities or operations necessary to measure it.
 - **E.g., The concept of financial performance – can be measured by financial ratio such as return on equity (ROE), Return on Asset (ROA) and return on investment (ROI)**
 - How well do the constructs in the instruments you have identified match the construct you have conceptually define for your study?
 - Not rely on the title of the measure alone, but need to review the actual items used in a questionnaire
 - Adherence to theory

...continued

- Is the evidence of reliability and validity well established?
 - Has it been validated in a population similar to the one you will be studying

- Did previous studies have a large amount of missing data , either on the measure itself or on items within the measure?

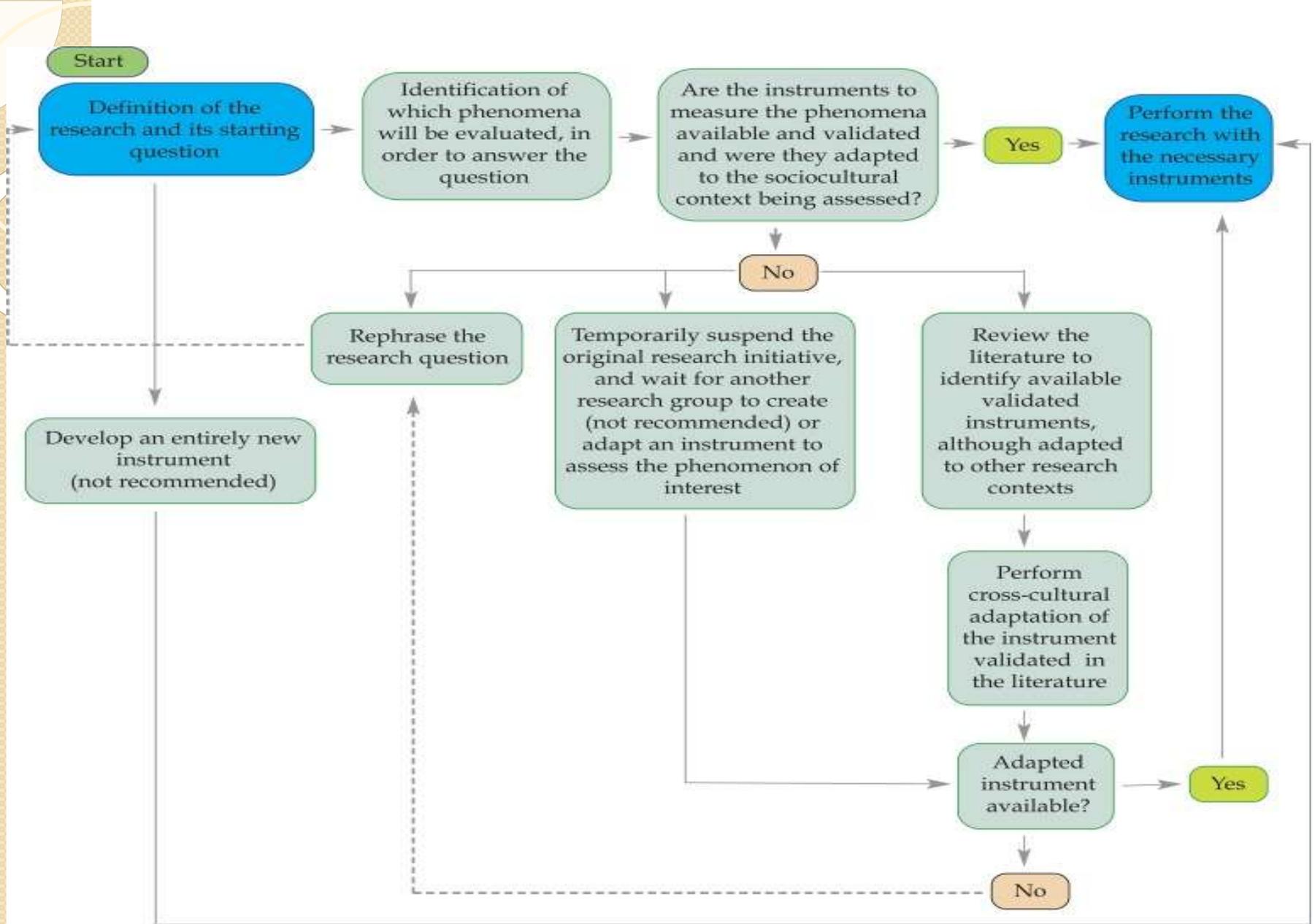
- Is the instrument in the public domain ?
 - If not, obtain permission for use from the author

- Will the instrument be acceptable to subjects?

...continued

○ **Example of Adopted and Adapted Instrument**

Variable	Measurement & Instrument		Source
	Original Version	Adapted Version	
<p>Innovation</p> <p>Operational Definition:</p> <p>Innovation refers to intentional efforts to create, introduce and implement of new ideas in order to provide the beneficial new outcomes to enhance performance of organization through an individual innovative behavior (Janssen, 2000)</p>	<p>Creating new ideas for difficult issues (idea generation)</p>	<p>I create new ideas for difficult issues</p>	<p>Janssen (2000). Cronbach's alpha coefficient was 0.95.</p>
	<p>Searching out new working methods, techniques, or instruments (idea generation)</p>	<p>I search out new working methods, techniques, or instruments</p>	
	<p>Generating original solutions for problems (idea generation)</p>	<p>I generate original solutions for problems</p>	
	<p>Mobilizing support for innovative ideas (idea promotion)</p>	<p>I mobilize support for innovative ideas</p>	
	<p>Acquiring approval for innovative ideas (idea promotion)</p>	<p>I acquire approval for innovative ideas</p>	
	<p>Making important organizational members enthusiastic for innovative ideas (idea promotion)</p>	<p>I make important organizational members enthusiastic for innovative ideas</p>	
	<p>Transforming innovative ideas into useful applications (idea realization)</p>	<p>I transform innovative ideas into useful applications</p>	
	<p>Introducing innovative ideas into the work environment in a systematic way (idea realization)</p>	<p>I introduce innovative ideas into the work environment in a systematic way</p>	
	<p>Evaluating the utility of innovative ideas (idea realization)</p>	<p>I evaluate the utility of innovative ideas</p>	



Decision tree to guide the process of choosing an instrument to collect scientific research data

Choosing Survey Response Scales

- It depends on what type of data you are looking for.
- There are 4 types of scales:
 - Nominal – classify into categories (e.g. Gender : 1=Male; 2=Female or ‘Yes’ or ‘No’)
 - Ordinal – categorize and rank order. E.g. rank preference from best to worse
 - Interval –group individual into different categories, rank their order, measure the distance between two points and explain the characteristics of the respondents.
 - Ratio –has an absolute zero point.
 - --group individual into different categories, rank their order, explain the characteristics of the respondents but also measures the proportion in the differences

	Nominal Scale	Ratio Scale								
Example Question	What is your age? 1 – Under 25 2 – 25-35 5 – Over 35	What is your age? ____								
What the Report Might Look Like	 <table border="1"> <caption>Bar Chart Data</caption> <thead> <tr> <th>Age Group</th> <th>Count</th> </tr> </thead> <tbody> <tr> <td>Under 25</td> <td>23</td> </tr> <tr> <td>25-35</td> <td>36</td> </tr> <tr> <td>Over 35</td> <td>12</td> </tr> </tbody> </table>	Age Group	Count	Under 25	23	25-35	36	Over 35	12	Mean Age: 28 Median Age: 25
Age Group	Count									
Under 25	23									
25-35	36									
Over 35	12									

TABLE 12.2 SELECTING STATISTICAL TESTS

Univariate	Bi/multivariate		
	NOMINAL	ORDINAL	INTERVAL/RATIO (Assumption of normality – if not normal use ordinal tests)
<p>NOMINAL</p> <p>2-point scale: gender – 1 = female, 2 = male</p> <p>3-point scale: religion – 1 = Catholic, 2 = Protestant and 3 = Jewish</p> <p>Central tendency: <i>Mode</i></p> <p>Dispersion: <i>N/A</i></p> <p>Compare sample to population: <i>Chi squared</i></p>	<p>Compare 2 or more Groups: <i>Chi Squared</i></p> <p>Compare within same group over times: 2 pts: <i>McNemar test</i> 3+ pts: <i>Cochran's Q</i></p> <p>Relationship with other variables: Yes/no: <i>Chi squared</i> Strength 2 pts: <i>Phi</i> Strength 3+ pts: <i>Lambda</i></p>	<p>Compare 2 groups: <i>Mann-Whitney</i> 3 or more groups: <i>Kruskal-Wallis</i></p> <p>Compare within same group over 2 times: 2 pts: <i>Wilcoxon signed rank test</i> 3+ pts: <i>Cochran's Q</i> 3 or more times (2+ pts): <i>Freidman's test</i></p> <p>Relationship with other variables: Yes/no: <i>Chi squared</i> Strength 2+ pts: <i>Lambda</i></p>	<p>Compare 2 or more groups: <i>ANOVA followed by t-test</i></p> <p>Compare within same group over times (2+ pts): <i>ANOVA followed by t-test</i></p> <p>Relationship with other variables: Yes/no: <i>Pearson's product moment correlation</i> Strength: <i>f-test</i></p> <p>With 2 or more independent and 1 dependent variable: <i>MANOVA</i></p> <p>2+ dependent variable and 3+ groups: <i>Multiple regression or Path analysis</i></p>
<p>ORDINAL</p> <p>TV viewing – order of preference, 1 = sitcoms, 2 = dramas, 3 = movies, 4 = news, 5 = reality TV OR Likert scale 1 = strongly disagree, 2 = disagree, 3 = neutral, etc.</p> <p>Central tendency: <i>Median</i></p> <p>Dispersion: <i>Interquartile range</i></p> <p>Compare sample to population: <i>Kolmogrov Smirnov</i></p>		<p>Relationship with other variables: Small sample <10: <i>Kendall's Tau</i> Larger Sample: <i>Spearman's Rho</i> With 1 variable as dependent: <i>Somer's d</i></p> <p>Small sample <10: <i>Kendall's Tau</i></p> <p>Larger sample: <i>Spearman's Rho</i></p> <p>With 1 variable as dependent: <i>Somer's d</i></p>	<p>Relationship with other variables:</p> <p><i>Jaspen's coefficient of multiserlal correlation</i></p> <p>With the interval/ratio variable as dependent: <i>ANOVA</i></p>
<p>INTERVAL/RATIO</p> <p>Interval – IQ score Ratio – real numbers, age, height, weight</p> <p>Central tendency: <i>Mean</i></p> <p>Dispersion: <i>Standard deviation</i></p>			<p>Relationship with other variables – no dependent/independent distinction: <i>Pearson's product moment correlation</i></p> <p>With 1 independent and 1 dependent variable: <i>Pearson's linear correlation</i></p>

PRE-TESTING & PILOT TESTING QUESTIONNAIRE

- **Pre-testing**

- the purpose of pre-testing a questionnaire is to ensure whether
 - a) *the wording of the questions is correct,*
 - b) *the sequence of questions is correct,*
 - c) *the respondents have clearly understood all the questions,*
 - d) *additional questions are needed or some questions should be eliminated, and,*
 - e) *the instructions are clear and adequate (Kumar et al., 2013).*
- All developed scales, or items, be **it adopted or adapted**, should be pre-tested to confirm whether the questions work accurately in a new setting with the new respondents (Kumar et al., 2013).

...continued

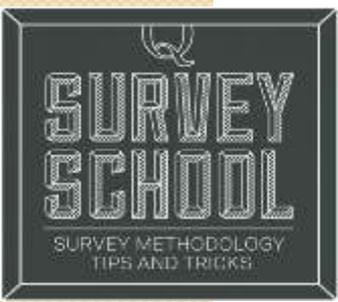
- A variety of pretest methods are available (e.g., informal pretests with small samples of respondents, cognitive interviews, expert panels), researchers frequently fail to pretest their questionnaires sufficiently.
- At least 50 people should be asked to participate (Kumar, Talib., & Ramayah, 2013); a sample of 30 participants is reasonable for pre-testing questionnaires (Perneger, Courvoisier, Hudelson, & Gayet-Ageron, 2015); Willis (2005) recommended a sample size between 5 and 15 individuals for the large-scale surveys.
- There is no clear-cut rule for the pre-test sample size – depends on length and complexity of the questionnaire
- A pre-test requires no statistical analysis

- 
- Pre-testing should be conducted using actual respondents—broadly representative of the target population to be chosen for main data collection (Cooper & Schindler, 2011; Kumar et al., 2013).
 - Upon successful processing of the first round, a second round of pre-test with the revised version of the questionnaire is highly recommended.
 - Pre-testing is often understood as a pilot study, both serve distinctive purposes.

5 Strategies for testing survey

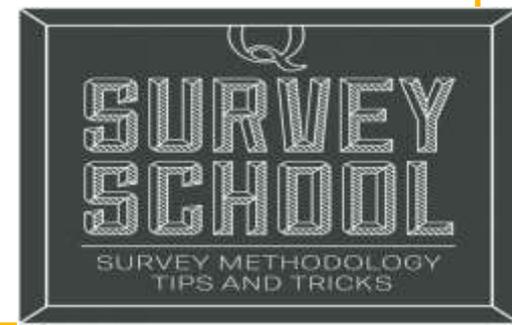
1. Respondent debriefing

- During the debriefing method, the researcher would carefully observe the respondent when he/she fills out the questionnaire.
- Once completed, the researcher would ask him/her to reveal any problems with the questions (For this approach, you will need to add several evaluation questions to the end of your survey for the respondents to answer.
- These can be open-ended or closed-ended questions and usually focus on assessing respondent comprehension and interpretation of survey questions. It should also include overall evaluations of the survey content, time, satisfaction and difficulty.



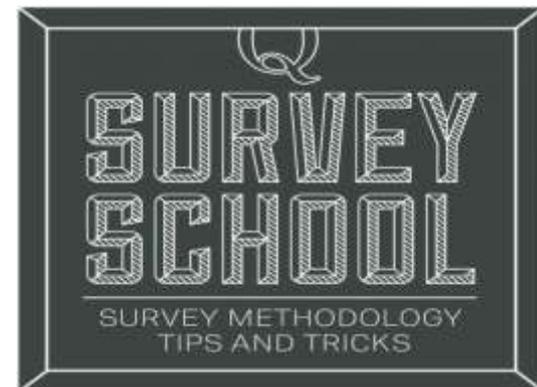
2. Cognitive interviewing

- “Cognitive interviews” are a good way to really understand what is going on the minds of your respondents when they are answering your questions.
- These are typically performed face-to-face with a small sample of 5–15 respondents.
- As the respondents answer each survey question, they are asked to “think aloud,” which can include paraphrasing, providing retrospective thinking or providing judgments of their confidence in what each question means.



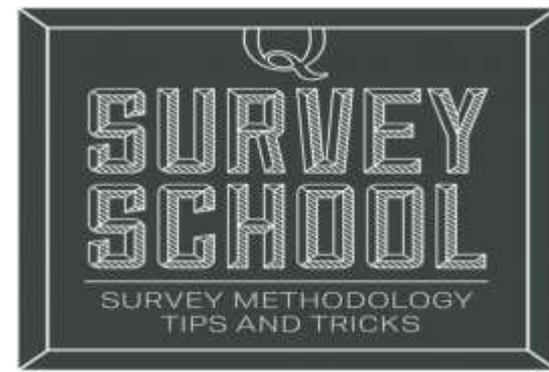
3. Expert evaluation

- Your survey can be dramatically improved by feedback from two types of experts:
 - 1) topic experts that have deep knowledge and expertise about the subject matter of your survey,
 - 2) survey methodologists that have expertise in how to collect the most accurate data for your research question.
- These expert evaluations can help shape the content and form of your survey and result in better data quality and more valuable insights.



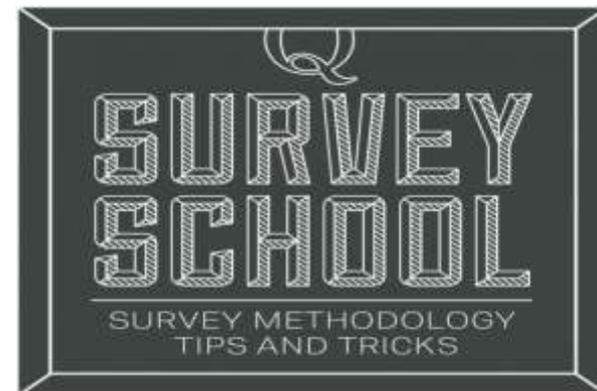
4. Focus groups

- In the preliminary phases of questionnaire development, it can be very helpful to ask a focus group discuss your survey.
- These discussions, which are usually semi-structured discussions between 7–15 people led by a moderator, are particularly helpful for clarifying basic concepts in the survey and evaluating perceptions of respondent burden or topic sensitivity.



5. Experiments

- Splitting a pre-test sample of respondents into groups
- testing different variations of your survey design and content for understanding the results you will get when you field your main survey.
- These experiments are particularly useful for understanding how changes in question wording, questionnaire design, visual layout, question order, and many other methodological factors may influence the data you collect.





“I don’t understand this question”

“The option I want isn’t available”

“I can’t find the next section”

“This is getting boring. Why is it so long?”

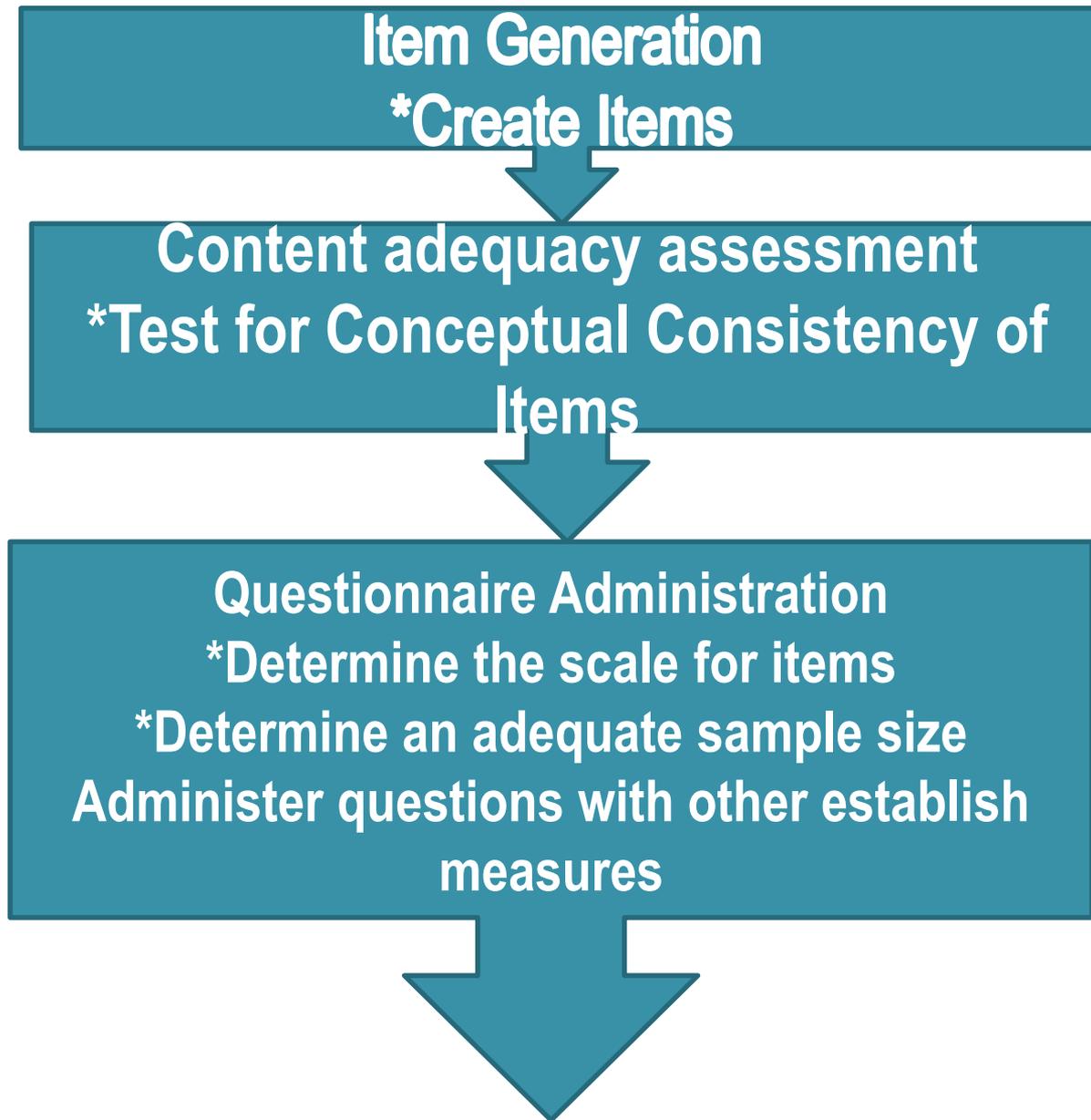
“Why is it asking about that? That makes me uncomfortable”

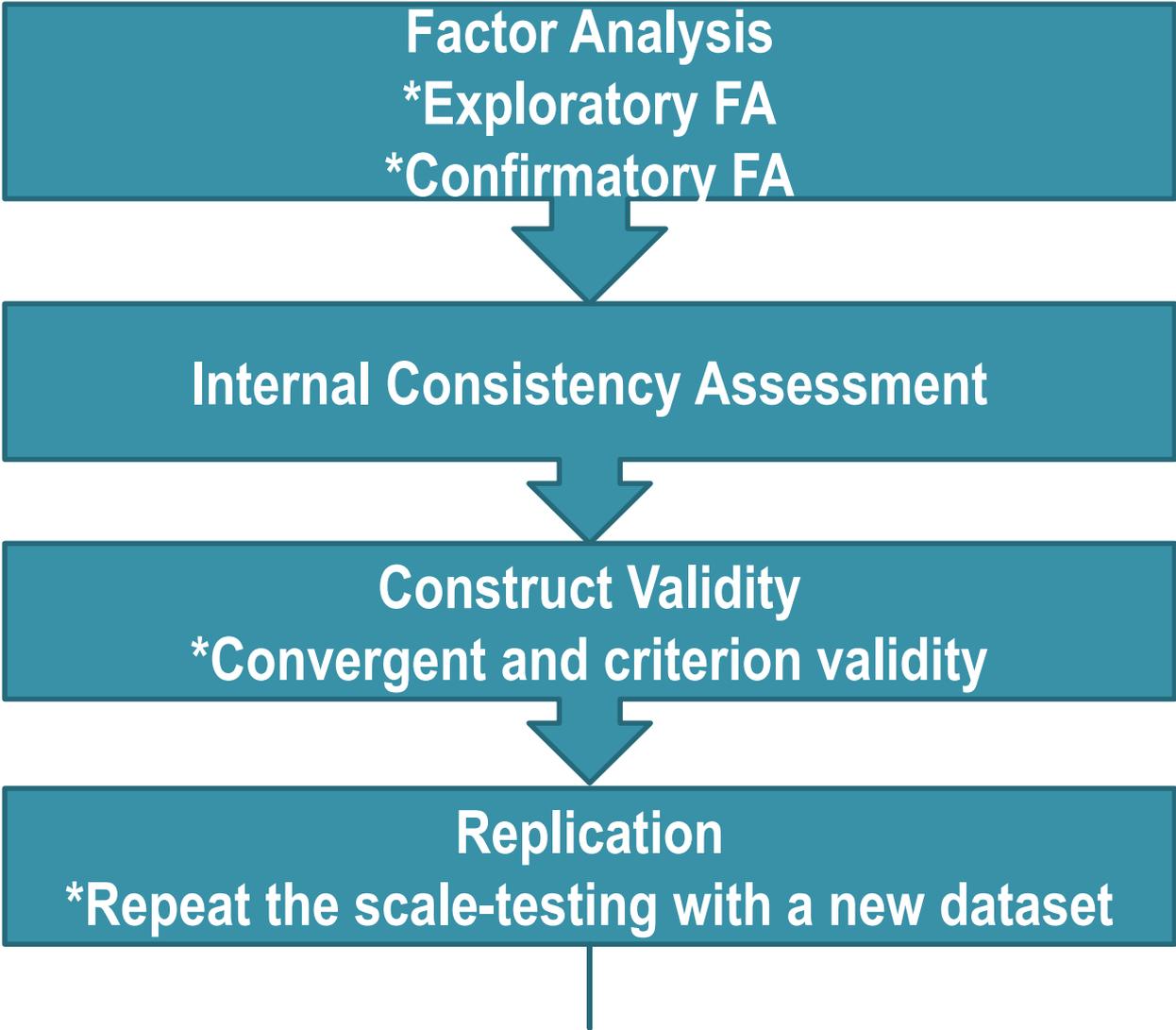
Careful notes should be taken on the problems encountered during pre-testing and possible solutions should be identified

PILOT TESTING/STUDY

- A pilot study is a small-scaled version or trial run—a key step to ensuring a full-fledged study will be carried out successfully
- Objectives for conducting a pilot study:
 - a) testing adequacy of research instruments,
 - b) assessment of the feasibility of a full-scale project,
 - c) assessing whether the research protocol is realistic and workable,
 - d) revealing logistics issues,
 - e) collecting preliminary data,
 - f) ensuring whether the sampling frame and technique are effective,
 - g) determining sample size
- Cooper and Schindler (2011) suggested a sample between 25 and 100 individuals; a range from 10 to 30 individuals are enough for a pilot test (Hill, 1998; Isaac & Michael, 1995); several scholars suggested that the sample size should be 10 percent of the sample projected for the main study (e.g. Connelly, 2008).
- Pilot study requires statistical analysis

Steps in Developing your own Research Instrument





```
graph TD; A["Factor Analysis  
*Exploratory FA  
*Confirmatory FA"] --> B["Internal Consistency Assessment"]; B --> C["Construct Validity  
*Convergent and criterion validity"]; C --> D["Replication  
*Repeat the scale-testing with a new dataset"]; D --> E[" "];
```

Factor Analysis
*Exploratory FA
*Confirmatory FA

Internal Consistency Assessment

Construct Validity
*Convergent and criterion validity

Replication
*Repeat the scale-testing with a new dataset

Guidelines for developing/constructing your own measure

- Refer articles by
 - Onwuegbuzie, A.J., Bustamante, R.M., & Nelson, J.A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, 4(1) 56–78
 - Timothy, R. H., Tracey, J.B. & Enz, C.A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research*, 21(1), 100-120.
doi:10.1177/109634809702100108