Research instrument development and data collection Technique

Asst. Prof. Dr. Hashima E Nasreen

Contents

• Steps in questionnaire design
• How to develop a research instrument
• Concepts of validity
• Types ad measurement of reliability
• How to construct data dictionary
What is a research Instrument

The Paper/document on which data is collected in called research instrument.

1. A form with no questions-only points for recording particular information
2. A check list – used to observed some situation, put tick marks against the particular point
3. A questionnaire - all information is collected by putting questions
   - Interviewer-administered questionnaire
   - Self-administered questionnaire
TYPES OF QUESTIONS

• Open-ended questions (Unstructured)
• Closed questions (Structured)
• Semi-structured questionnaires (containing completely open as well as partially pre-categorized and closed questions)

OPEN-ENDED QUESTIONS

Such questions are useful for obtaining in-depth information on:
• opinions, attitudes and suggestions of informants, or
• sensitive issues.

For example:
‘What is your opinion on the services provided in the ANC?’ (Explain why.)
‘What do you think are the reasons some adolescents in this area start using drugs?’
➢ The answers to these questions are written down as closely as possible in the words of the respondents.
Advantages of completely open-ended questions

• Allow you to probe more deeply into issues of interest being raised.
• Issues not previously thought of when planning the study may be explored, thus providing valuable new insights on the problem.
• Information provided in the respondents’ own words might be useful as examples or illustrations, which add interest to the final report.
• Often, re-reading an answer in a later phase of the analysis offers the possibility for different interpretations in relation to other data collected, which would have been impossible if the answer had been pre-categorised.

Risks of completely open-ended questions

• Skilled interviewers are needed to get the discussion started and focused on relevant issues and to record all information collected. A big risk is incomplete recording of all relevant issues covered in the discussion.
• Analysis is time-consuming and requires experience; otherwise important data may be lost.
Closed / Structure questions

CLOSED QUESTIONS have a list of possible options or answers from which the respondents must choose.

Advantages of closed questions:
• It saves time
• Comparing responses of different groups, or of the same group over time, becomes easier

Risks of closed questions:
• In case of illiterate respondents, the interviewer may be tempted to read the list of possible answers in the given sequence, thereby influencing the choice of response and introducing bias.

For example:

What is your opinion on the following statement ‘Women who have induced abortion should be severely punished.’

1. Strongly agree
2. Agree
3. Not sure/no opinion
4. Disagree
5. Strongly disagree
For example:

- Did you eat any of the following foods yesterday? (Circle yes if at least one item in each set of items is eaten)

<table>
<thead>
<tr>
<th>Peas, beans, lentils</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish or meat</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Eggs</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Milk or cheese</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Insects</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

CHECKLISTS

Observations Ventilated Pit-latrines
(Tick the appropriate boxes)

<table>
<thead>
<tr>
<th>Evidence of use: worn footpath to VIP</th>
<th>Yes (clear)</th>
<th>Little (hardly noticeable)</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toilet smells</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flies in and around the toilet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fouling around the toilet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defecation in the area around the homestead</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Prerequisite for developing a research instrument

Before developing a research instrument:
• Hypothesis, objectives and all variables of interest should be known
• Operational definitions of the variables should be clear and stated
• Decision about scaling or categories of the variables should be made

Qualities of a Good questionnaire

• The questionnaire should be of appropriate length (20-30 minutes acceptable)
• Question should be relevant to the hypothesis
• Question should be in specific, not vague
• Language of the question should be simple and clear
• Layout is easy to read and pleasant to eye
• Sensitive questions must be worded exactly
• Double barreled questions should be avoided
Points to be noted for designing a Questionnaire

• Follow your hypotheses & objectives and to determine the content of your questions
• Make sure your questions are valid. The question must ask for information which the respondent is capable of providing an informed response
• Make the question as specific as possible
• Ask others to review and proofread your survey
• Always pretest
• Make instructions as clear as possible.

The quality of your research depends on the quality of your questionnaire

Quantitative Instruments

(Wording Questions)

• Do not use leading questions that suggest a particular answer or the researcher’s viewpoint
  e.g., Why does working in the public sector make people lazy and careless?)
  • Do not use loaded questions that bias people towards particular answers
  e.g., Do you think it is important to have a strong police force in this time of crisis?)
  • Beware of double-barreled questions. Rephrase them into two questions whenever possible
  e.g., Should the government reduce its financial help and its technical assistance to other countries?
  • Avoid double negatives (e.g., not prohibit), they can be confusing. If absolutely necessary, emphasize the “not” by underlining it
OVERVIEW OF DATA COLLECTION

TECHNIQUES

Data-collection techniques allow us to **Systematically** collect information about our objects of study (people, objects, phenomena) and about the settings in which they occur.

In the collection of data we have to be systematic. If data are collected haphazardly, it will be difficult to answer our research questions in a conclusive way.

**Various data collection techniques**

- Using available information
- Observing
- Interviewing (face-to-face)
- Self administering questionnaires
- Focus group discussions
- Measurement
Using available information

• Usually there is a large amount of data that has already been collected by others, although it may not necessarily have been analyzed or published.

  For example, analysis of the information routinely collected by health facilities can be very useful for identifying problems in certain interventions or in flows of drugs supply, or for identifying increases in the incidence of certain diseases.

• Analysis of health information system data, census data, unpublished reports and publications in archives and libraries or in offices at the various levels of health and health-related services, may be a study in itself.

OBSERVATION

– It is a technique that involves systematically selecting, watching and recording behaviour and characteristics of living beings, objects or phenomena.

Observation of human behaviour is a much-used data collection technique. It can be undertaken in different ways.

• Participant observation: The observer takes part in the situation he or she observes.
  (For example, a doctor hospitalized with a broken hip, who now observes hospital procedures from within).

• Non-participant observation: The observer watches the situation, openly or concealed, but does not participate.
Interview

An INTERVIEW is a data-collection technique that involves oral questioning of respondents, either individually or as a group.

Administering written questionnaires

• Self-administered questionnaire is a data collection tool in which written questions are presented that are to be answered by the respondents in written form.

A written questionnaire can be administered in different ways, such as by:

• Sending questionnaires by mail with clear instructions on how to answer the questions and asking for mailed responses;

• Gathering all or part of the respondents in one place at one time giving oral or written instructions, and letting the respondents fill out the questionnaires; or

• Hand-delivering questionnaires to respondent and collecting them later.

Parts of a questionnaires

1. A title: Clear and brief
2. An Introductory remark
3. Instructions
4. An identification number
5. The questions
6. The ending
Designing Questionnaires

Orient the respondent

• First, briefly describe the purpose of the research study, explain how data gathered by the survey will be used, and by whom it will be used.
• Is the survey confidential or anonymous?
• requesting the informant’s consent to be interviewed and assuring confidentiality of the data obtained.

Designing Questionnaires

First Questions

• The first several questions should be relevant to the study itself so that the respondent quickly understands what the survey is about and becomes engaged.
• The first questions should be straightforward with relatively few categories of response.
Designing Questionnaires

Middle Questions
• Respondents should be eased into sensitive topics by asking them what they think is important or what they do prefer.
• Do not first ask respondents to agree or disagree with a position or sensitive issue.

Designing Questionnaires

Sequence your questions logically!
• Questions should be grouped in sections
• Each section should have a logical sequence
• Avoid making the respondent jump around mentally
• Help respondents shift gears by introducing a new section. For example, “Now, we would like to get your opinion on some related areas".
Pre-Testing Questionnaires

• Can they clearly understand what is being asked?
• Does the flow of the questions make sense?
• Will other people have difficulty?
• Which questions in particular might pose problems?

Pre-Testing Questionnaires

• Are there too many "neutral", "don’t know" or "don’t remember" responses?
• Do you need additional questions relevant to the research?
• Do you need to provide more space for written responses?
• Did respondents respond appropriately to open-ended questions?
VALIDITY

The concept of validity

• Validity is the ability of an instrument to measure what it is intend to measure
• Degree to which the researcher has measured what he has set out to measure (Smith, 1991)
• Extent to which an empirical measure adequately reflects the real meaning of the concept under consideration (Babbie, 1989)
Why validity?

- Validity is done mainly to answer the following questions:
  - Is the research investigation providing answers to the research questions for which it was undertaken?
  - If so, is it providing these answer using appropriate methods and procedures?

Types of validity

- Content validity
- Criterion validity
- Construct validity
- Face validity
- Concurrent validity
- Predictive validity
Content validity

- Use logical reasoning and hence easy to apply
  - Justify if each question in relation to the objective of the study
  - Easy if questions relate to tangible matters
- Extent to which a measure instrument covers a representative sample of the domain of the aspects measured
- Whether items and questions cover the full range of the issues or problem being measured

Face validity

- The extent to which a measuring instrument appears valid on its surface
- Each question or item on the research instrument must have a logical link with the objective
Problem with content validity

• Based on subjective logic, no definitive conclusion can be drawn

• Extent to which questions reflect the objectives of the study may differ. If wording changed or questions substituted, magnitude of link changes

Criterion validity

• The extent to which a measuring instrument accurately predicts behaviour or ability in a given area

• The measuring instrument is called ‘criteria’

• Predictive validity: if the test is used to predict future performance
  – Measurement of sugar exposure for caries development
• Concurrent validity: If the test is used to estimate present performance or person’s ability at the present time not attempting to predict future outcomes
• e.g. measurement of DMFT for caries experiences

Construct validity

• Most important type of validity
• Assesses the extent to which a measuring instrument accurately measures a theoretical construct it is designed to measure
• Correlates performance on the instrument with the performance of an established instrument
Construct validity

• Method 1: Factor Analysis
  – To examine empirically the interrelationship among items and to identify clusters of items that share sufficient variation to justify their existence as a factor or construct to be measured by the instrument
  – Then these items are gathered into common factors
  – Common factors are synthesized into fewer factors and then relation between each item and factors is measured
  – Unrelated items are eliminated

• Method 2: diagnostic test
  – By calculating sensitivity and specificity using two instruments (instrument to be tested vs. standard instrument)
Reliability

- Definition: It is the ability of an instrument to create reproducible results
- Each time it is used, similar scores should be obtained
- A questionnaire is said to be reliable if we get same/similar answers repeatedly
- It can be measured by estimating correlation coefficients
Reliability measured in aspects of:

**Stability**
- Done to ensure that same results are obtained when used consecutively for 2 or more times
- Test-retest method is used

**Internal consistency**
- To ensure all subparts of a instrument measure the same characteristic (homogeneity)
- Split-half method/ cronbach's alpha

**Equivalence**
- Used when two observers study a single phenomenon simultaneously
- Inter-rater reliability

Statistical calculation for reliability

- Correlation coefficient
- Can range from – 1 to + 1
- 0 indicates absence of any relationships

<table>
<thead>
<tr>
<th>Correlation coefficient</th>
<th>Strength of relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>+/- 0.7 to 1.0</td>
<td>Strong</td>
</tr>
<tr>
<td>+/- 0.3 to 0.69</td>
<td>Moderate</td>
</tr>
<tr>
<td>+/- 0.0 to 0.29</td>
<td>None to weak</td>
</tr>
</tbody>
</table>
In a nutshell…
Both reliable and valid

Reliable: the shots *reliably* hit the same part of the target.

Valid: the shots are clustered at the centre, where they were aimed.

Data dictionary
What is data dictionary

• A data dictionary, as defined in the IBM Dictionary of Computing, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format."

• In other words, a 'data dictionary' describes the structure and attributes of data 'items' to be used within a software application (usually a database).

• A data dictionary includes the names and descriptions of the tables and the fields contained in each table. It also documents information about the data type, field length and other things such as validation.

Purpose

• The main purpose of the data dictionary is to provide metadata, or information about data.

• Technically, it is a database about a database.

• There is no one set standard in terms of layout or the level of detail to which a data dictionary should be produced.
Data dictionary: an example for layout

<table>
<thead>
<tr>
<th>SI No.</th>
<th>Variable</th>
<th>Definition</th>
<th>Description</th>
<th>Scale of measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Marital status</td>
<td>One’s situation with regard to marriage.</td>
<td>Single 1, Married 2, Separated 3, Divorced 4, Widowed 5</td>
<td>Categorical Nominal</td>
</tr>
</tbody>
</table>

Than you