

QUESTIONNAIRE DESIGN

A working paper
to be discussed on June 5, 2010

by

Bidhan Acharya

Associate Professor
Central Department of
Population Studies
Tribhuvan University



Enumerators in Dhanusha in Exercise for Filling in the Questionnaire of Social Exclusion Survey of CNAS, TU

For a

*Training –cum-Workshop
in Research Methodology*

Organised by

*Centre for Post-graduate Studies
Nepal Engineering College*

in Collaboration with

**University Grants Commission
Nepal
Sanothimi, Bhaktapur**

Pulchok, Lalitpur, Nepal
2-11 June 2010

D31b	Family members employed in urban centres	1. Yes	2. No	8. DK	9.NS
D31e	Better access to Markets	1. Yes	2. No	8. DK	9.NS
D31f	Better access to Loans	1. Yes	2. No	8. DK	9.NS
D31g	Improve in transportation and communication	1. Yes	2. No	8. DK	9.NS
D31h	Changes in Family size	1. Yes	2. No	8. DK	9.NS
D31i	Others (Specify	1. Yes	2. No	8. DK	9.NS
D32	Have you experienced also some negative changes around your community?	1. Yes	2. No	8. DK	9.NS
D33	How much do you think the society is adversely affected by negative changes, on the following?				
D33a	Loss of wage earner family member	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33b	Family responsibility as the burden for females	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33c	Unemployment	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33d	No market for products due to social exclusion	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33e	Price hike, inflation	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33f	National capital goes to India	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33g	Diminishing of traditional skills	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33h	Traditional skills are threatened	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33i	Increased number of sick and injured	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS

D33j	Lack of money due to conflict	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33k	Destruction of property and resources due to conflict	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33l	Lack of access to education due to conflict	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33m	Low profit due to conflict	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33n	Displacement of people due to conflict	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33o	Only women remained in agriculture	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33p	Environmental change and increased natural disaster	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33q	Loss of property due to increased tendency of theft	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33r	Bankruptcy due to (forced) migration and debt	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33s	Political instability, strike and <u>Barah</u>	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS
D33t	Others (Specify	1. Highly affected	2. Affected	3. Not Affected	8. DK 9.NS

Section-E, Public Health

E1	Was anyone in your family sick during last one year?	1. Yes	2. No	8. DK	9.NS
E2	What did you do first when the person was detected with sickness?				
E2.1	Looked for a health worker / doctor	1. Yes	2. No	8. DK	9.NS
E2.2	Looked for a faith healer	1. Yes	2. No	8. DK	9.NS
E2.3	Offerings at temple or places (deval/rahd) of spirits	1. Yes	2. No	8. DK	9.NS

TABLE OF CONTENTS

1. General Background	2
2. Types of Questionnaire	2
2.1: Content of Questions:	3
2.2: Type of Questions	3
Close ended and Open ended	3
Scaling type of questions	4
Matrix type of Questions	6
Ranking type of questions.....	6
Caution: Rank and Category	7
3: Construct of Questions.....	7
4. Pre-coded questions	7
5. Retrieval of the Events.....	8
6. Language of the Questions.....	8
6.1: Eleven ‘No’s in question designing.	9
1. No question without objective:	9
2. No complex language:	9
3. No ambiguous concepts:	9
4. No leading and embarrassing questions (Wording, Leading and threatening):.....	9
5. No shorter checklist (Response Set):	9
6. No longer questions:	9
7. No 2 in 1:	10
8. No double negative (Double barrelled):	10
9. No calculations:	10
10. No longer and vague reference periods:	10
11. No reference of previous questions:	10
7. Sequence and Skipping	10
7.1: Funnel and inverted funnel sequence.....	11
8. Texts, Clarifications and Instructions	12
9. Layout of Questionnaire	12
10. Questionnaire Administration.....	13
11. Reliability Instruments.....	13
Question of Reliability while Questionnaire Designing	13
12. Conclusion	14
Suggested Readings	14

QUESTIONNAIRE DESIGN

Bidhan Acharya

1. General Background

Most of the social science researches are based on the information provided by the target population chosen for the analyses. How to determine target or sample population is normally the part of discussion in sampling design. This paper discusses on the essentials of the questionnaire construction. Questionnaire design is very crucial and important part of the research, because an inappropriate questionnaire misleads the research, academics and policymaking. Therefore, a set of adequate and appropriate questions in a sequential order is required in a questionnaire.

A questionnaire is defined as a document containing questions and other types of items designed to solicit information appropriate to analysis (Babbie, 1990:377). Questionnaire is equally used in survey research, experiments and other modes of observation. Indeed, people ask different questions in their daily life to satisfy their queries. Journalists, market researchers, observers and interested persons in different events ask a set of questions to others. They form an opinion or conclusion according to the answers that they receive. The use of the questionnaire or a set of questions is a similar process in the research. The research questions are firmly constructed to receive answers related to the chosen variables for analysis.

Social science research is the basic tool to inaugurate, illustrate and predict the human behaviour. The findings of research contribute to policy making and enhancement of the academics. These findings are based on the answers collected through different questions asked to the sample population. Naturally, dubious and inconsistent questions produce ambiguous and inharmonious answers that result in futile analyses and incoherent outcome of the research.

Therefore, the analyses, outcomes, general conclusions, recommended policy and identified area for future research all depend on how perfectly the questionnaire is constructed. Thus, questionnaire construction is the most important part of the research and other information collection activities.

Human behaviour is one of the most unpredictable matters in the world. People with similar social setting, same level of education and income might be different from each other in terms of average number of children they have and use and non-use of family planning methods. There are many seen and unseen issues, which guide the behaviour and differential results are observed. Researchers do need to collect information on these issues that are related to their studies. Furthermore, the researches need to collect information on such issues, which are not even discussed in the groups of the people. Even the husband and wife have not discussed on the topic yet, the researchers do want to know their opinions that is expected to reflect the couple. Moreover, many issues in the life of a person, couple, household or society are abstract. They are thought as abstracts and practised as abstracts too. With the help of some sets of questions, the researchers collect information on these abstract behaviours and quantify collected information according to their own definitions. The challenge to the researchers is that they have to copy human behaviour on paper in quantitative terms and analyse them as to reflect what really is happened in the society. The relevant questions are the only tool to know about human characteristics. Therefore, survey questionnaires are also termed as structured in through quantifiable information is collected and analysed.

2. Types of Questionnaire

There are roughly two types of questionnaires, structured and unstructured. A mixture of these both is the quasi-structured questionnaire that is used mostly in social science research.

Types of questionnaire

- **Structured**
- **Unstructured**
- **Quasi-structured**

Structured questionnaires include pre-coded questions with well-defined skipping patterns to follow the sequence of questions. Most of the quantitative data collection operations use structured questionnaires. Advantages of such structured questionnaires are - less discrepancies, easy to administer, consistency in answers and easy for the data management.

Unstructured questionnaires include open ended and vague opinion-type questions. Maybe questions are not in the format of interrogative sentences and moderator or the enumerator has to elaborate the sense of question. Focus group discussions use such questionnaire.

Not all questions are easily pre-coded with almost possible alternatives of answers. Given answer alternatives of some questions in standard questionnaire are left as 'others' (please specify). A common and pragmatic practice is that most of the questions are structured, however, it is comfortable to have some unstructured questions whose answers are not feasible to enumerate completely. Such a type of questionnaire is called as quasi-structured questionnaire.

2.1: Content of Questions:

In fact we intend to obtain answers either the **facts** related to the respondents or the **opinion** of the respondents regarding the subjective or even objective matters.

Answers sought	
Facts	past present
Opinions	past present future

Consider an example - the question like - Has anyone of age 6 to 14 from your home gone outside for working (as a child labour)? The answer we can presume that it will come in the format of either Yes or No.

This is FACTUAL question and the expected answer is dichotomously categorized in Y or N that is simultaneously also the CLOSE ENDED question. Okay, if we had have asked another subsequent (that is called as **contingency**) question; *if it is so, is it very common in this community that a boy or girl of age 6 to 14 go out for work for earning?* The answer alternatives we have arranged in **Likert Scaling** format -- as

1. Very Common 2. Somewhat Common, 3. Neither Common Nor Rare, 4. Somewhat Rare and 5. Very Rare.

The answer runs from extreme positive to extreme negative. It is one of the subjective judgements.

Facts are always related to past and present; and opinion could be for all three times - past, present and future.

Example - fact based question:

C7.	Which of these items does your family own in the household?				Nos.
C20a	Bicycle	1. Yes 2. No 8. DK 9.NS	C20a1	If yes, how many?	
C20b	Motorcycle	1. Yes 2. No 8. DK 9.NS	C20b1		
C20c	Car / Jeep	1. Yes 2. No 8. DK 9.NS	C20c1		
C20d	Tractor/Truck	1. Yes 2. No 8. DK 9.NS	C20d1		

Example - Opinion based question

C8.	How do you rate the status of your family compared to other families in your neighbourhood in terms of landholding?	
	1. Better off, 2. Above than average, 3. Average, 4. Below average, 5. At the bottom stratum, 8. Don't Know, 9. Not stated	

2.2: Type of Questions

Close ended and Open ended

Closed ended and open ended are popular type of questions. Literatures illustrate **close ended** as *structured* and **open ended** as *unstructured* ones. Close ended questions have sufficient alternatives to select or to fit in the information given by the respondent. However, open ended are open for the

answers. The verbatim must be recorded in open ended and has problems for the post-coding in data management. In short, close ended questions are difficult to construct and open ended are difficult to analyse.

Example: Structured / close ended question

C9.	Is your annual cereal/grain production sufficient to feed your family for the whole year?	
	1. Yes → C9_1 2. No → C8 8. DK → C9_1 9. NS → C9_1	
C10.	If not sufficient, for how many months can your cereal/grain production feed your family?	
	1 2 3 4 5 6 7 8 9 10 11	

C9 is FACT based and C10 is OPINION based, but both are structured. If there is answer 1 in C9 the question c10 would not be asked. Such questions carried by the answer to previous questions are called as *Contingency Questions*.

And if we add another question as -

C11.	In case of food shortage, how do you manage for survival?
------	--

It is an open ended question. Answers might be several. Maybe one respondent have more than one answers too. Consider some possible alternatives for answer.

Purchase/borrow food	Barter in neighbourhood	Eat out in the place of labour	Eat out in hotel
Alternative food / bring from the Jungle	Seek humanitarian support	Seek friends' and relatives' support	
Seek government support	Others (Specify.....)		

If you know if not cent percent, but majority of the possible alternatives you can manage your question accordingly so that the data management would be easier later. Collection of information only is not important, but analysis of that information is very important. Therefore, Pre-TEST of questionnaire is important. And, if I know the above alternatives I may manage the question as:

C12.	In case of food shortage, how do you manage for survival? <i>Please select one answer from each row.</i>					
C12a	Purchase/borrow food	1. Yes	2. No	8. DK	9.NS	
C12b	Barter in neighbourhood	1. Yes	2. No	8. DK	9.NS	
C12c	Eat out in the place of labour	1. Yes	2. No	8. DK	9.NS	
C12d	Eat out in hotel	1. Yes	2. No	8. DK	9.NS	
C12e	Alternative food / bring from the Jungle	1. Yes	2. No	8. DK	9.NS	
C12f	Seek humanitarian support	1. Yes	2. No	8. DK	9.NS	
C12g	Seek friends' and relatives' support	1. Yes	2. No	8. DK	9.NS	
C12h	Seek government support	1. Yes	2. No	8. DK	9.NS	
C12i	Others (Specify.....)	1. Yes	2. No	8. DK	9.NS	

The term 'structured' and 'unstructured' is used sometimes for answers and sometimes for the question itself. If I make a checklist of points to be asked to respondents like: Food → Sufficiency → Management in Insufficiency, it becomes - unstructured question, because the question wording is not fixed and I may ask the same question to various respondents differently that may lead differential answer too. If we consider C12 as one question, the total set is also considered as *Multiple Answer* type of question. The information on household facilities as radio, TV, electricity, vehicle etc. is also the multiple answers set question, in which a respondent can select more than one answer.

While dealing with the types of questions we should not forget the rating and scaling format of questions. There are three popular formats of rating¹ and

Scaling type of questions

Scaling is used to quantify the subjective or responses. The marriage between quantitative and qualitative methods to substantiate the findings of researches is in practice in contemporary times.

¹ Thurstone Scaling (Equal Appearing); Guttman Scaling (Cumulative); and Likert Scaling (Summative)

Scaling methods do support for shaping the open ended type and opinion based answer seeking questions. Though there is always room for questioning the validity of answers by the respondents, these methods assist researchers to arrive at some general conclusions regarding those issues that are not quantifiable.

- Scaling is a process of making a kind of composite measure consisting several items that have a logical or empirical structure among them (Babbie, 379).
- The set of categories or range of scores on a variable is called as scale, and the process of assigning scores to objects to yield the measures of a construct is called scaling (Judd, Smith and Kidder, 146).
- Measuring the complex and abstract concepts more accurately is scaling (Kothari, 94).
- Scaling is the arrangement of possible opinions of respondents in a coherent order of behaviour or attitudes, in which a person could judge him/her to be fit in certain standpoint (Bidhan).

Scaling is the branch of measurement that involves the construction of an instrument that associates qualitative constructs with quantitative metric units. Scaling evolved out of efforts in psychology and education to measure "unmeasurable" constructs like *authoritarianism* and *self esteem*. In many ways, scaling remains one of the most arcane (mysterious or hidden) and misunderstood aspects of social research measurement. And, it attempts to do one of the most difficult of research tasks -- *measure abstract concepts*.

Please refer to following for extended knowledge on scaling techniques:

Chapter VII: Scaling in Judd, Smith and Kidder (7)(146-170)

Chapter VIII: Index and Scale Construction in Babbie (1)(147-173)

Chapter VI: Measuring Attitudes in Sudman and Bradburn (13)(148-173)

Chapter V: Measurement and Scaling Techniques in Kothari (Rc 11)(94-115)

<http://www.socialresearchmethods.net/kb/scaling.htm>

Though, rating and scaling is dealt separately in research methods, but I did not observe 'scaling' as a separate topic of discussion in this workshop. And, type of question needs to know scaling too. Therefore, an elementary knowledge of scaling is included here. Three types of scaling as following are popular in social research.

1. Thurstone scaling (Equal Appearing): This is simply grouping and arrangement of answers given by the respondents. In Time Allocation Survey in 1996 in CDPS, we had 1156 types of activities collected by the respondents from some 18000 households. We grouped them into 99 different categories of work in the initial two stages; and finally prepared 9 major divisions of those groups. All activities had equal importance and none of them was to cancel out the other. However, if our intention is to identify the most income generating activity, we can rank them as poorest as 1 to highly income generating as 9. This is an example of Thurstone Scaling.

2. Guttman Scaling (Cumulative): In this method the cumulative nature of current (or expressed) opinion of respondent includes the previous ones too. For example consider the series of answers below.

- I believe that Nepal should allow more Indian immigrants in.
- I would be comfortable if a new Indian immigrant moved next door to me.
- I would be comfortable with new Indian immigrants moving into my community.
- It would be fine with me if new Indian immigrants moved onto my block.
- I would be comfortable if my children have Indian immigrants as their friends.
- I would permit a child of mine to marry an Indian immigrant (adapted from Trochim).

Every subsequent answer alternative has the idea of previous alternative as an inherent attribute of it. Therefore, this is cumulative. If the person is not uncomfortable to marry his/her children with an Indian immigrant, there would be no question of objecting their immigration.

3. Likert Scaling (Summative): It is arrangement of opinions from extreme negative to extreme positive. Its options are like :

1. Fully Agree, 2. Somewhat Agree, 3. Neither Agree nor Disagree, 4. Somewhat Disagree and 5. Fully Disagree.

Researches have found that scaling options arranged in horizontal fashion are more reliable than in vertical fashion in questionnaire if the respondents are well educated and questionnaire are self administered.

An Example from CNAS Questionnaire for Social Inclusion Survey is below.

Do you agree with your children or family members (either boys or girls) if they have inter-caste (or cross cultural) marriage with the following groups?		Strongly Agree	Somewhat Agree	No Comment	Somewhat Disagree	Strongly Disagree	Don't Know	Not Stated		
P6_1	Hill Brahmin	1	2	3	4	5	8	9		
P6_2	Hill Chhetri	1	2	3	4	5	8	9		
P6_3	Other Hill Janjati	1	2	3	4	5	8	9		
P6_4	Terai origin	1	2	3	4	5	8	9		
P6_5	Janjati - Terai Origin	1	2	3	4	5	8	9		
P6_6	Hill Dalits - Sarki, Kami, Damai	1	2	3	4	5	8	9		
P6_7	Terai Dalits	1	2	3	4	5	8	9		
P6_8	Muslims	1	2	3	4	5	8	9		
P6_9	Christians	1	2	3	4	5	8	9		
P6_10	Any religious people	1	2	3	4	5	8	9		
P6_11	Other (specify)	1	2	3	4	5	8	9		

Matrix type of Questions

Matrix question provide multiple answers to more than one persons or elements. The Matrix is of, indeed, not questions but the type of format they are arranged. Sometimes these questions are also called as *Battery Questions*. These questions are in the form of tables where mostly the information is titled without proper sentencing like a question. These can satisfy the need for multiple variable information on many persons. The major disadvantage of such type of questions is the difficulty in data entry and analysis process if it is to be merged with respondent's file². Example of the household roster is best fit to the battery questions. A careful explanation is required in analytical phase.

Example of Matrix format or Battery Questions

B1. SN	B2. Name (Optional)	B3. Sex	B4. Age	B5. Relation to HH head	B6. Ethnicity	B7. Religion	B8. Literacy	B9. Education	B10. Occupation	B11: Monthly Income	B12: Survey Status
01											
02											
03											
04											
05											

Ranking type of questions

Ranking is also the best fitting information of respondent for *lowest to highest categorised variables* of research. These are in fact the qualitative type of judgement but quantified in numbers. For example - If I ask you a question -

How do you score to Current government's performance? Give your score in Full marks 10.

² Please discuss on this issue in much greater detail in Data Management session with Dr. Kabita Bade Shrestha

Now your answer can be ranked from 0 to 10. In most of the ranking the scores in between lower and upper limits can be left un-labelled.

Caution: Rank and Category

Many researches do categorise the interval scale variables prior to data collection. Such efforts generally rank the variables.

For example – If I ask you to select one appropriate answer that best fits your income and give alternative as 1. Rs. 5,000 to 9,999 ; 2. 10,000 to 14,999; 3. 15,000 to 19,999; 4. 20,000 to 24,999 and 5. 25,000 and above – the income is ranked. However, consider a precaution that such ranking are very harmful if done with interval scale variables as age, income, number of times of visit (eg ANC/PNC) or number of years of schooling. In above example, the income level as 19550 =3, and 10340=2; and also 15000=3, and 14950=2. If you need to run statistical tools as Correlation and Regression you should NOT group the answer alternatives in the level of questionnaire design. That will completely mislead the phenomena. Grouping should be done only in data analysis phase. Advantage of ranking is researcher can interpret as the problem is set and designed to analyse, and demerit is that it overlaps one answer to another, messes-up the answers and there exists always possibility of colliding of two ranks together.

No Categories in the phase of questionnaire design for variables in Interval Scale

3: Construct of Questions

The design of questionnaire should be consistent to the expected analysis of the data. The objectives and hypotheses (in quantitative) and/or the research questions (in qualitative) guide the methodology and plan for the analysis of data. Therefore, the objectives and/or hypotheses of research determine how a questionnaire should be organised and what items (constructs/correlates/variables - in the form of questions) should be included. Therefore, the conceptualisation process of research is important. Let me tell you that the observation (information) or intuition is important to generate a research. Then you have a big question to your mind - why it happens to some situations and why not to others? Researchers are the rational people and they want to study the differences and reasons as well as the consequences. Researches have noble ideas of recommending policy options for mitigating gaps in most of the social research. Thus the observation (problem) follows a path via objective, literature survey; identification of variables; conceptual framework; hypotheses / research questions; methods, analysis, conclusion and lastly to recommendation. It's a journey from dissatisfaction of researcher to satisfaction by recommending the solutions or finding of new relations between or among variables.

Before constructing a questionnaire

- Have clear *Objectives* and *Hypotheses*
- Have a listing of selected variables and items
- Have section breaks of topics

An initial collection of coherent items or variables related to the objectives and hypotheses is always beneficial. Therefore, a listing of variables before questionnaire design is suggested. Such a listing helps identify the important dimensions of the domain of interest. The arrangement of variables in a sequential manner from origin to destination is conceptual framework. The questionnaire design must include all variables appear in conceptual framework and variables identified.

Also, the questionnaire designers are required to be familiar with the vocabulary of respondents. Understanding of the vocabulary of target population makes questions simple and easily answerable. Furthermore, a questionnaire might contain items related to more than one general topic. The topics of questionnaire are if specifically categorised into sections that makes questionnaire more practical and easy to administration.

4. Pre-coded questions

Questionnaire for quantitative studies should be ready to *standardised interviewing*. Standardised interviewing is

- Use pre-coded answer alternatives in quantitative researches
- Pre-coded questionnaire minimizes the cost of data management

the administration of questions by professional interviewers who are not experts in the subject matter. All questions, answers, skipping patterns, information and directions to the interviewer should be clearly scripted in the defined order. Probe questions should have clear direction on when to probe with which syntax.

Obviously, close-ended questions with defined answers are encouraged. Close-ended questions are accommodating in data management. However, close-ended questions sometimes create problems especially in behavioural research, if questionnaire is not adequately pre-tested in the homogeneous group of sample population.

Open-ended questions may receive various responses. These responses require categorisation and post-coding before they are entered into microcomputer. Post-coding is tedious and seeks high level personnel that are familiar with entire research. Therefore, open ended questions and post-coding is not favoured in contemporary quantitative researches.

Some qualitative research methods prefer open-ended questions. Open-ended questions provide more freedom to the interviewers as well as respondents. What exactly the respondent is feeling and what exact language (verbatim) is the response is sometimes very important in anthropological in-depth research. However, the further analyses of qualitative information need high level human resource with much time and resources.

5. Retrieval of the Events

The respondents have to retrieve the answers to the posed questions from their memory. This retrieval system has three dimensions. First is the *top-down retrieval* in which the questions are designed from general to more specific information. For example a research regarding the knowledge on the modes of transmission of HIV/AIDS is conducted. Though the objective of the research is to know whether the people at risk know about the different ways of transmission, the questions should not begin with "Do you know the modes of transmissions of HIV/AIDS ". It should rather start from the 'heard' about HIV.

Retrieval of the events by respondents

- Top-down
- Sequential and
- Parallel

Second, the *sequential retrieval* refers to logical sequencing of the events within the theme. In this method the reverse chronological order might serve better. Theoretically, the most recent event is assumed as to serve as reminder of the previous event. Questions should be asked on the recent event first, than immediate past and previous events sequentially. For example in the above research - "When did you have a recent meet with a prostitute" might serve as a reminder of "Had you met her before this too".

General rules of question crafting

- Clear objective
- Simple language
- Clear concepts
- Without bias
- Adequate answer options
- Shorter questions
- Single question at a time
- Affirmative sentences
- Mathematics not imposed
- Short/clear reference periods
- Avoid question reference

The reverse chronological ordering is important especially for the questions regarding birth history information in demographic researches and job history information in employment researches.

Finally, *parallel retrieval* reflects the environmental structure that governs the life-course of the persons. Many events happen at the same period in the life of persons. If a period is referred, some other parallel events at that period are easy to retrieve for the respondents. Considering the exemplary research the response of *Had you visited abroad, Where, When did you visit Bangkok last time, Did you visit to commercial sex worker there*, types of question should be sequentially arranged.

6. Language of the Questions

The language used in the questionnaire is the most important part of the questionnaire construction. The respondents should feel comfortable with the language used in the questionnaire. Questionnaire language should be non-threatening to the respondents. The investigator(s) should be familiar with the

vocabulary of respondents. Not only in qualitative but also in quantitative researches it is suggested that before pre-testing a team of investigators visit to the sample area and understand their subjects.

Some rules of questionnaire crafting are presented as following. The question designers should be aware of 'what should not be done' while designing questions.

6.1: Eleven 'No's in question designing.

1. No question without objective:

Each question should have objective. Question without objective should be avoided. Example: Proposed research is to assess the knowledge of respondents on sexually transmitted diseases. If the proposed analytical framework has no consideration of the educational (by discipline) background of respondent it is futile to ask: "Which subject did you study at university before you joined the recent job?"

2. No complex language:

The language of the questionnaire should not be complicated to understand. The vocabulary of the respondents should be used in the questionnaire. Simple language is preferred. Use of rhetorical and elite language creates problems while the questionnaire is administered. Example: *Did you realise the complexities of life in a different way by the behaviour of your spouse when you were tested positive with HIV/AIDS?* Instead, the questions like *Do your spouse knows about your HIV positive? (If Yes, Do you find a change in his/her behave? If Yes, What kind of?*

3. No ambiguous concepts:

Ambiguous concepts should not be incorporated in the questions. Example: *What is your opinion about some medical researches that pledge for the high prevalence of transmission of HIV among elite group of Nepal after restoration of multiparty system?* This question has three major elements as medical research, HIV transmission and restoration of multiparty democracy. Elite group and high prevalence are other minor elements. Respondent would not be able to correctly form his/her opinion.

4. No leading and embarrassing questions (Wording, Leading and threatening):

Leading and embarrassing questions should be rephrased. People feel offensive to answer these questions. Such questions also lead towards the biased answers; therefore, these are to be avoided. Example: *Don't you agree that the persons with HIV positive have also rights to marry? Or suppose, you are suffering from HIV positive, should not you have the rights to marry?* Such types of questions insist respondent to provide answers that match the positive or negative tone of the question itself (Yes).

5. No shorter checklist (Response Set):

The answer alternatives should be adequate. To avoid the biases in the stress and pronunciation of interviewer *show cards* for the literate respondents would be desirable. Example: *What was your age when you got married? Options: a. Less than 20 b. 20 or more.* Instead of these two options leave such questions open-ended, because the answer is definite that comes in one or two integers. Or, if the question is like 'What was your occupation when you were detected with HIV positive?' Do not leave the answers with too few options like *a. agriculture, b. service.* Instead, a long list of possible occupations serves well.

The questions should not be designed in such a way that more than one option is merged together at the stage of data collection. Such merge of answer options causes for losing real value of data. Many times, such previously categorised data set does not permit you to run some statistical analyses. Consider the example of educational attainment. Normally it is not suggested to categorise educational attainment at the stage of questionnaire design as 1. Illiterate, 2. Primary, 3. High School, 4. SLC/Intermediate, Plus two, and 5. Bachelor and above. However, the categories must made later by observing the frequencies while at data analysis stage.

6. No longer questions:

Every question should be short and preferably in only one sentence that match with the cognitive capacity of the respondent. However, some longer statements could also be given to provide cues to the

respondents. For example, the methods used intentionally by a couple to avert the birth of child in the future are known as family planning methods. Such methods are natural, mechanical, surgical and chemical in nature. *Had you or your spouse ever used one of such methods with an intention to avert the birth of a child?* The question and preceding statement are relatively longer but the statement clarifies the answer sought by the forthcoming question. Such cue statements and questions are permitted in a questionnaire. Nevertheless, the questions like *"When you first moved from your house to the place of destination and you had an additional baby born, and what was the difference between your income of place of origin and of the place of destination"* type of questions should be avoided.

7. No 2 in 1:

Merging of two questions into one should be completely avoided. Such merging often confuses the respondent and according to the cognitive capacity some respondents serve answers to the latter and some to the former. No all respondents provide answers to the both parts. Example: *When did you visit to your spouse and how many nights did you spend there?* There are clearly 2 questions and they are to be segregated.

8. No double negative (Double barrelled):

Double negatives must be avoided in the language of question. Double negative gives positive meaning but sounds like negation to the statement. It also creates confusion to the interviewers and respondents. Example: *"Do not you want to move from this place not to expose yourself?"* Instead, *"Do you want to move from this place to hide yourself?"* would be better.

9. No calculations:

As far as possible, avoid all calculation seeking questions. Respondents do hesitate to calculate and there is always possibility of receiving wrong answers. Respondents who can not calculate also give wrong answers to hide their ignorance and who can, they also have a tendency of wrong calculation to exhibit their confidence in calculations. Example: *What per cent of your income is spent for the health care per month?* Instead, use "What is your monthly income?" as a preceding question of *"How much did you spend in your health care last month?"* and calculation should be performed in data processing and analyses phase. Do not give strains to the respondents.

10. No longer and vague reference periods:

Reference periods should be clear and preferably shorter. Longer reference period causes recall lapse errors. These errors mislead the research. For example, *after the year of greater earthquake or in these ten years how many times did you visit to health post for antenatal check-ups?* Instead "How many times did you visit to health post for check-up during the period of your last pregnancy (or three months)?"

11. No reference of previous questions:

It is extremely not suggested to ask the questions like *"As I asked in Question number 12 above about*". If reference or cue of previous questions is required recall (or look at) the answer of respondents by stating full question and answer to continue the further interview. Instead *"You said that you had used a method of family planning, could you please mention that method"* would be better. But the preferred way is to maintain a very good skipping pattern to have a sequence of questionnaire administration.

7. Sequence and Skipping

The questions should be started with very simple and general questions to specific ones. To establish a better rapport with the respondents, place non-controversial and common types of questions in the beginning. A beginning with household and non-behavioural information is pragmatic.

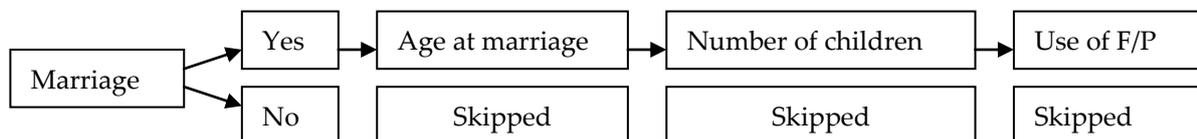
Most of the life course events are age specific. Chronological follow of events either in ascending or descending order sets a sequence. Identify the origin of event, causes, reactions and consequences respectively. Some times which one of the event should be considered first creates problem. For example in urban areas the age at menstruation might be asked to a woman before the question on age at

marriage and age at cohabitation; because menstruation normally occurs before marriage and cohabitation in hill and mountains. Same event in Terai may have different dimension. If a girl is married before menstruation, but she cohabited with her husband after menstruation (*gauna*), in this case, the age at marriage is less important than the age at cohabitation for fertility analysis. The questionnaire formulators should be familiar with such situations and manage their questions accordingly.

Flow chart example to identify skipping patterns

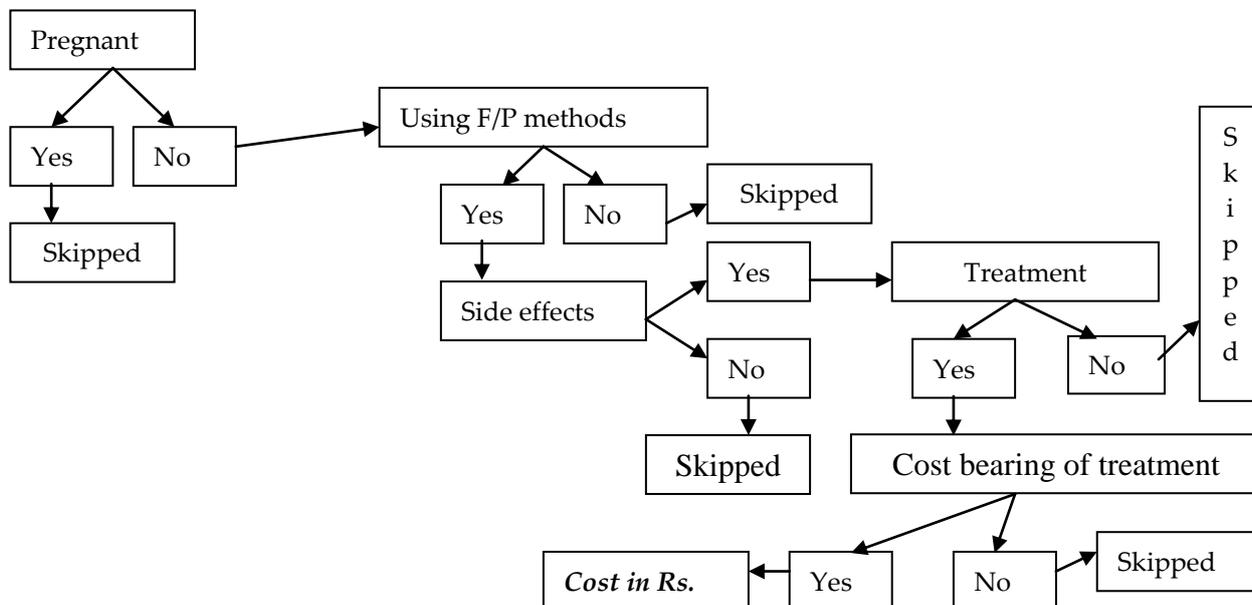
Simple example

Question objective: General information on marriage, CEB and Contraception



Slightly complex example:

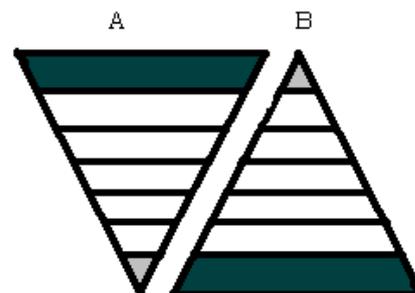
Question objective: know the cost of treatment for side effects of contraception among married women.



If an option is answered in previous question what should not be asked in subsequent questions is skipping. A flow chart can help to identify the skipping patterns in a questionnaire. If a head of the household tells you that he owns no latrine in his household it is futile to ask about the type of latrine. That question should be skipped for him/her.

7.1: Funnel and inverted funnel sequence

Funnel sequence are used to obtain specified information starting from the general one and inverted funned start from the recent the micro event back to the generalised situations. Suppose, among you some of the participants are interested to know how much money is spent by the persons in late night dance restaurants and bars in the Kamalpokhari area of Kathmandu. The situation 'A' starts from generalised themes like, where do you live, why do you come to Kathmandu, where do you work, how much do you earn, and now how much do you spend for the dance bars? That is funnel type question from very general to specific one. The counter-model uses rather opposite method that asks how much have you spent tonight and goes back in sequence to where from



have you come to Kathmandu. If the respondent is supposed to remember gradually the information the upside down funnel method is appropriate, and if the respondent has to gradually evolve the information and has to reach to some climax normal funnel as in A is important.

8. Texts, Clarifications and Instructions

A questionnaire also contains items other than questions. These are information regarding selected sampling unit, statements to clarify the objective of questions and directions to the supervisor/interviewer either to follow the skipping patterns or do calculations or probe the contents.

It is not practical to have large data files in computer. Therefore, different sections of questionnaire are usually entered as separate data files. The MEBDC (Migration, Employment, Birth, Death and Contraception) Survey conducted by the Central Department of Population Studies, Tribhuvan University; DHS 2006, 2001 by New Era are examples of larger surveys in Nepal. Each section and sub section of the questionnaire consists of identification numbers as district, cluster, household and individual ID on the top of the section. This information serve as "merging/matching variables" when there will be a need to extract information from another section³.

9. Layout of Questionnaire

The layout of questionnaire depends on the mode of administration. Developed countries also use self-administered questionnaires. A respondent is expected to fill in the questionnaire without assistance of others in such self-administered questionnaires. Therefore, self-administered questionnaire should be attractive and printed in a very clear manner of directions and skipping patterns. Professional interviewers handle telephone survey⁴ questionnaire; therefore ordinary lay out can also serve well in such cases.

Sudman and Bradburn (1982:230) have a number of suggestions to make a questionnaire handy and attractive. The trained interviewers administer most of the questionnaires in survey in developing countries. For the easiness of transportation and handling each questionnaire should be printed in quality paper and firmly stitched.

The name of survey conducting organisation and title of the survey should appear on the first page (and if possible, in short in each page) of questionnaire. The name of interviewer, date and time of interview are also to be included in this page.

Section breaks are important to segregate the broader topics. From beginning to the end without section breaks confuses the enumerator and respondent. It is always better to use different colour papers for different sections of questionnaire. It is easy to check consistency even in the sections. Questions should not be printed in crowd. There must be sufficient space for comfort. Font type should be large and clear enough.

Characters are less easy in data entry and processing. So, each question and answer options should be numbered. Sub-questions are to be printed differently than main questions to distinguish them. A single question should never be parted into two different pages. If not so, the interviewers and respondents might think that the question is completed at the end of earlier page.

The answer format should be vertical for individual questions, but if the answers are in scales the answer options should be presented in horizontal fashion. Answer options should be pre-coded, but if there is possibility of answer options other than coded, there must be the option as 'others' with sufficient space to write.

³ Refer to Data Management session of this seminar

⁴ Computer Assisted Telephone Interviewing (CATI) is popular in some commercial surveys, in developed countries. In these surveys the interviewer directly inputs the data in computer as obtained answers by respondents on phone.

The length of questionnaire is always a debatable topic. However, the length of a questionnaire depends upon the research topic and information required. Nevertheless, the questionnaire that is administered in less than half an hour is normally considered as comfortable for both interviewers and respondents.

10. Questionnaire Administration

Prepared questionnaire should be pre-tested to check the vocabulary, consistency and ambiguity. Pre-test provides basis for amendment of questionnaire. No matter how better is the questionnaire, if the administration is performed in no normal environment the results are certainly affected. In haste, within a little time and in the situation of mental tension of either party (interviewer or respondent) the information may be biased.

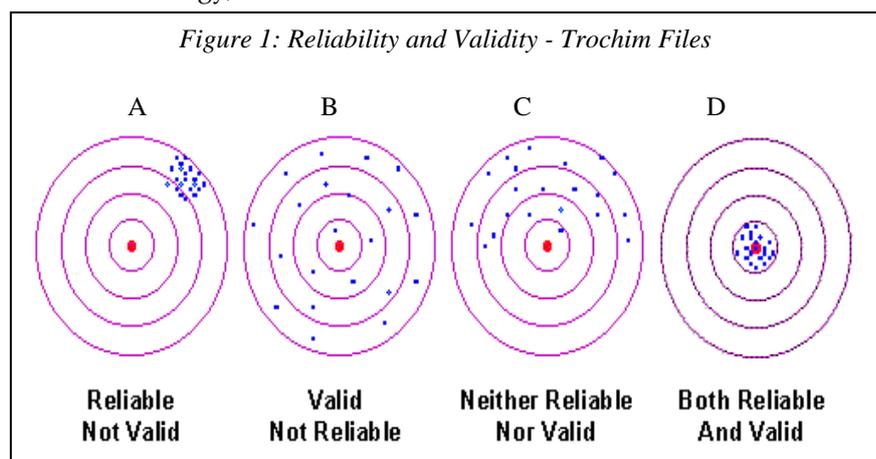
Some interviewers give stress in some of the questions and read lousily the other questions that are not the topic of interest for them. That eventually harms the quality of data.

11. Reliability Instruments

Reliability⁵ is the simply the occurrence of events. If we have asked the question *that how the people manage the waste produced at home*, and the group of illiterate respondents have replied that almost half of time they leave the waste at the streets in open, and if it is the true case some months or years the 'question as an instrument' with its answer alternatives is reliable. Not all instruments in questionnaire are always reliable. And it is not necessary that only reliable tools are used in a research, it depends upon the objective of research. In time series longitudinal surveys behavioural change type questions are asked even if they are found not reliable over the time.

Golafshani defined the reliability as the extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable (Golafshani, 2003:1).

Trochim (2006) has depicted reliability and validity in illustrations. Figure 'A' is representing Reliable Not Valid; the occurrence of event is there for sure. Consider above example and suppose that out of the 10 clusters we selected, people in one cluster deposited their waste on streets unattended. The one tenth do the same is reliable,



but for generalising the conclusions, it's not considered as a valid one. Figure 'B' is scattered, that means the same 10 percent in all clusters do the same; therefore, it has no reliability but the information is valid. Figure 'C' represents that the behaviour is distributed arbitrarily and in half of the clusters only with no reliability and no validity. Figure 'D' showed both Reliability of occurrence and validity of association.

Question of Reliability while Questionnaire Designing

There are both types of arguments that the phase of questionnaire design should be or should not be worried about the reliability of instruments. The reliability could be measured only after the data collection and questionnaire design has nothing to do with it is one kind of version. However, the pre-

⁵ Validity very often comes with reliability; 'Validity' is the association between / among variables. In this case if in all clusters the illiterate people use to throw their wastages in all clusters the information is valid. There is the association between Literacy and Waste Management.

test is important and some observation of reliability of the measures must be done prior to finalising the questionnaire is another pragmatic observation. The ultimate objective of questionnaire is to make a research success in policy formulation. If in a 'pre-test' cent percent of respondents were found with depositing the waste in container, maybe this question is also futile.

12. Conclusion

A bad questionnaire, a bad administration of good questionnaire, a bad coverage of samples with good questionnaire and good administration all cause for less valid data. Questionnaire construction is one of the initial activities that are critical for a pragmatic social science research. If the questionnaire is far from the reality, it will collect a gamut of garbage information. A misled information produces misleading conclusion and futile recommendations. Thus the efficiency of research entirely depends upon the excellence of questionnaire. Take adequate time, if necessary have a series of pre-tests and finalise the questionnaire. A perfect questionnaire is considered as half job is completed.

This is not references cited, but just a list of some publications that contain important information on survey research methods and questionnaire design.

Suggested Readings

- Babbie, Earl, 1990. *Survey Research Methods*, Second Edition. California: Wadsworth Publishing Company.
- Campbell, J. Gabriel, Ramesh Shrestha and Linda Stone, 1979. *The Use and Misuse of Social Science Research in Nepal*. Kathmandu: Centre for Nepal and Asian Studies, Tribhuvan University.
- Judd, Charles M., Eliot R. Smith and Louise H. Kidder, 1991. *Research Methods in Social Relations*, Sixth edition. Florida: Harcourt Brace Jovanovich, Inc.
- Lucas, David (Ed.), 1985. *Asking Demographic Questions*, Demography Teaching Notes V. Canberra: Australian National University.
- Sudman, Seymour and Norman M. Bradburn, 1982. *Asking Questions: A Practical Guide to Questionnaire Design*. (San Fransisco: Jossey-Bass Publishers.
- Weller, Susan C. and A. Kimball Romney, 1988. *Systematic Data Collection* (Newbury Park: Sage Publications.
- Golafsani, Nahid, 2003. "Understanding Reliability and Validity in Qualitative Research" in *The Qualitative Report*, Vol. 8 No. 4. Website - <http://ace.upm.edu.my/~lateef/Handouts%20-%20dce%205920/golafshani%20-%20reliability%20and%20validity%20in%20qual%20research.pdf>
- Trochim, William M K, 2006. *Social Research Methods, Knowledge Base*. Website - <http://www.socialresearchmethods.net/kb/index.php>
<http://www.analytictech.com/mb313/principl.htm>
<http://www.statpac.com/surveys/questionnaire-design.htm>
<http://www.socialresearchmethods.net/tutorial/Colosi/lcolosi2.htm>

Pokhara University Affiliated
Nepal Engineering College
Centre for Post-graduate Studies
Pulchok, Lalitpur Nepal
Research Methods Workshop
Questionnaire for Participants' Survey

The questions and answers in this questionnaire will be treated confidentially. They will not be used other than research purposes.

Mark a circle (ex: Ⓔ) to numbers fitting the answer .

A. GENERAL

Q_1. Name (Optional)	c d
Q_2. Completed Age	
Q_3. Sex: 1. Male, 2. Female 3. Transsexual	
Q_4. What is your Ethnicity?: 1. Khas (Hill) (Bahun, Chhetri, Sanyasi); 2. Janjati /Aadibasi (Hill) 3. High Caste Tarai (Tarai Brahman, Bhumihar), Rajput 4. Other Tarai Caste (Yadav, Ahir, and Baniya); 5. Hill Dalits; 6. Tarai Dalits; 7. Tarai Janjati 8. Muslims 9. Others (Specify.....)	
Q_5. Religion : 1. Hindu, 2. Bouddha; 3. Kiranti; 4. Islam; 5. Christians; 6. Jain; 7. Shikh 8. Others (Specify.....)	
Q_6. Place of Birth: 1. Urban 2. Semi-urban 3. Rural	
Q_7. Development Region of Birth Place. 1. EDR, 2. CDR; 3. WDR; 4. MWDR 5. FWDR	
Q_8. Ecological Belt of Birth Place 1. Mountain 2. Hills 3. Tarai	
Q_9. What is your marital Status? 1. Unmarried (Single/ Never Married) →Q11 2. Married and living with spouse 3. Married but not started to live with spouse 4. Separated 5. Divorced 6. Widowed 7. Other's (Specify)	
Q_10. How many persons live together in the house you are living now?	

B. ECONOMIC

Q_11. What is your Occupation?	
Q_12. What is your monthly income? Rs	
Q_13. What amenities are available at your home? a. Landline telephone 1 Yes, 2 No, 8 DK, 9 NS b. Cell phone with you 1 Yes, 2 No, 8 DK, 9 NS c. Motorcycle 1 Yes, 2 No, 8 DK, 9 NS d. Computer 1 Yes, 2 No, 8 DK, 9 NS e. Refrigerator 1 Yes, 2 No, 8 DK, 9 NS f. Four + wheeler 1 Yes, 2 No, 8 DK, 9 NS g. Washing Machine 1 Yes, 2 No, 8 DK, 9 NS	

C. PARTICIPATORY

Q_14. How many of your family members are associated with in any of the political organisations? Men Women	
Q_15. How many of your family members are associated with in any of the social organisations? Men Women	

Q_16. How many of your family members are associated with in any of the NGOs? Men Women	
Q_17. How many of your family members are associated with in any of the following agencies? Agency M F a. School Management b. Forest Users' Group c. Road Users' Group d. Mothers' Group e. Small Credit Group f. Vegetable Growers' Group g. Others	

D. WASTE MANAGEMENT

Q_18. How the people in your neighbourhood manage their daily wastage? 1. Dispose in their own yard/field → Q21 2. Leave on the street → Q22 3. Deposit in Container nearby → Q20 4. Collectors collect and dispose →Q19 5. Bio-degradable are disposed within compound and others left on the streets 6. Others (Specify	
Q_19. At what time the collectors come? AM/PM →Q22	
Q_20. How often does the municipality change the container? →Q22 1. Daily 2. In two/three days 3. Once in a week 4. Fortnightly 5. Once in a month 6. Never 8 DK 9 NS	
Q_21. Do you produce some sort of compost fertiliser out of the waste disposed within homestead premises? 1. Y 2 N 8 DK 9 NS	
Q_22. Do you know that the metallic and plastics wastes could be recycled and reused? 1. Y 2 N 8 DK 9 NS	
Q_23. Do you agree to sell non-biodegradable items in a community centre? 1. Y 2 N 8 DK 9 NS	
Q_24. Are you interested to join a cooperative for recycle and reuse of non-biodegradable item in the community? 1. Y 2 N 8 DK 9 NS	
Statement: There is a good market for the cramps (kabadi) of non-biodegradable items. A factory of Plastics is in the process of being established.	
Q_25. How much are you interested to invest in that factory of recyclable and reusable materials? 1. Fully Interested 2. Somewhat Interested 3. Indifferent 4. Somewhat not Interested 5. Not Interested at All	

Thank you for your cooperation!