

Module

8

Maria Teresa Siniscalco
and Nadia Auriat

Questionnaire
design



Quantitative research methods in educational planning

These modules were prepared by IIEP staff and consultants to be used in training workshops presented for the National Research Coordinators who are responsible for the educational policy research programme conducted by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ).

The publication is available from the following two Internet Websites:

<http://www.sacmeq.org> and <http://www.unesco.org/iiep>.



International Institute for Educational Planning/UNESCO

7-9 rue Eugène-Delacroix, 75116 Paris, France

Tel: (33 1) 45 03 77 00

Fax: (33 1) 40 72 83 66

e-mail: information@iiep.unesco.org

IIEP web site: <http://www.unesco.org/iiep>



September 2005 © UNESCO

The designations employed and the presentation of material throughout the publication do not imply the expression of any opinion whatsoever on the part of UNESCO concerning the legal status of any country, territory, city or area or of its authorities, or concerning its frontiers or boundaries.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means: electronic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from UNESCO (International Institute for Educational Planning).

Graphic design: Sabine Lebeau

Typesetting: Sabine Lebeau

Printed in IIEP's printshop

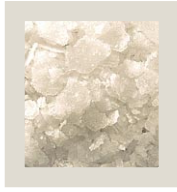


Content

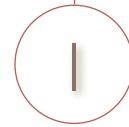
1. Introduction	1
2. Initial planning	3
Why a new questionnaire – and when?	3
Relationships between research problems, research hypotheses, and variable construction	5
Characteristics of research hypotheses	6
Specifying variables and indicators	6
Operationalization of research questions	8
The problem of the cross-national validity of educational concepts, definitions, and data collection instruments	12
1. What is formal education?	12
2. Distinguishing public and private service providers	14
3. What is a school?	16
4. What is a student?	17
5. What is a teacher?	18
3. The design of questions	22
Question structure	22
1. Closed questions	23
2. Open-ended questions	26
3. Contingency questions	28
Guidelines for writing questions	29

Specifying the characteristics of respondents	34
A checklist for reviewing questionnaire items	35
4. Examples of questions	38
Student background	38
1. Gender and age	39
2. Socio-economic background: occupation, education, and possessions	41
Teacher characteristics	47
School location	49
Learning, teaching, and school activities	51
1. Student reading activity	51
2. Teacher activities	52
3. School head activities	53
Attitudes, opinions, and beliefs	55
1. Likert scaling	55
5. Moving from initial draft to final version of the questionnaire	63
Two widely-used patterns of question sequence	63
General guidelines for item placement	65
Covering letters and introductory paragraphs	66
Drafting instructions for answering questions	68
Training of interviewers or questionnaire administrators	70
Pre-testing the questionnaire	72
Basic steps in pre-testing	75
Reliability and validity	76
1. Validity	76
2. Reliability	78

The codebook	79
6. Further reading	84



Introduction

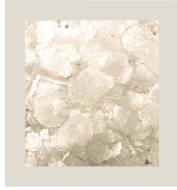


- This module provides guidance for the design of standardized questionnaires that are to be administered in school systems to students, teachers, and school heads. The module is divided into four sections that cover initial planning, the design of questions, examples of question types, and moving from a draft to a final questionnaire.

After reading this module, the reader should be able to design a quality survey questionnaire that is suitable for addressing the research issues at hand. He or she will know how to:

- Decide on the target population for the questionnaire.
- Identify the variables and indicators that will address the research issues and hypotheses on which data are to be collected.
- Develop demographic, knowledge, attitude, and practice questions.
- ‘Close’ open ended quantitative and qualitative questions and design skip, filter, and contingency questions, where appropriate.
- Decrease response bias and maximize response rates.
- Design probe questions and interviewer or respondent instructions on the questionnaire.

- Conduct a pilot test of the questionnaire, and adjust its final design according to the results.
- Prepare a codebook for data entry.



Initial planning

This section reviews the steps required to determine the need for a new questionnaire, and looks at how a general research problem needs to be translated into a number of specific research questions and hypotheses. It examines the problem of valid cross-national instruments and provides helpful hints and recommendations for using comprehensive and precise definitions of key educational concepts.

Why a new questionnaire – and when?

This module addresses the planning and design of standardized questionnaires. A formal standardized questionnaire is a survey instrument used to collect data from individuals about themselves, or about a social unit such as a household or a school. A questionnaire is said to be **standardized** when each respondent is to be exposed to the **same** questions and the **same** system of coding responses. The aim here is to try to ensure that differences in responses to questions can be interpreted as reflecting differences among respondents, rather than differences in the processes that produced the answers.

Standardized questionnaires are often used in the field of educational planning to collect information about various aspects of school systems. The main way of collecting this information is by asking people questions – either through oral interviews (face to face or telephone), or by self-administered questionnaires, or by using some combination of these two methods.

Although survey research, by definition, implies the use of some form of questionnaire to be administered to a sample of respondents, the questionnaire is simply one instrument that can be employed in the study of a research problem. As such, it may or may not be the most suitable tool for the task at hand.

Hence, before deciding on the need for a new questionnaire, one should consider whether or not some of the required information may already be available from other sources, for example, from statistics compiled by governments or research agencies, or from survey research archives. One should also consider whether a suitable questionnaire already exists that could be wholly or partially used.

The planner should also consider whether other means of data collection are more appropriate. These can be (a) field experiments, where people in 'treatment' and 'control' groups respond to a scenario devised by the investigators, (b) content analysis of newspapers or articles, (c) direct observation (such as counting the number of schools in a district, the number of blackboards in a school or the number of students per teacher in a given area), or (d) non-directive interviews where there are no pre-specified questions and the interviewer has a great deal of freedom in probing areas and specific issues during the course of the interview.

Among the types of information that can be collected by means of a questionnaire are facts, activities, level of knowledge, opinions, expectations and aspirations, membership of various groups, and attitudes and perceptions. In the field of educational planning, the information that is collected can be classified broadly into: (a) inputs to education (such as school resources or various background characteristics of schools, teachers or students), (b) learning and teaching processes, and (c) the outcomes of education (such as pupil achievement, attitudes towards school, and measures of school efficiency such as survival rates etc.).

Relationships between research problems, research hypotheses, and variable construction

The development of a questionnaire commences with the transformation of general educational research and policy concerns into specific research questions for which the data are intended to supply an answer. Some examples of general educational policy and research concerns are: (a) policy-makers want to assess the supply of resources in their primary schools, (b) a curriculum expert wants to determine to what extent teaching methods explain differences in reading literacy among 9-year-old students, and (c) a national evaluation agency wants to investigate student attitudes towards science at the end of secondary school.

In the case of the above three examples, it would be necessary to establish empirical evidence for decisions through the collection of data on facts (school resources), activities (teaching methods), and attitudes (students' views towards science), respectively.

A **research hypothesis** is a tentative answer to a research problem expressed in the form of a clearly stated **relation** between independent ('cause') and dependent ('effect') variables. Hypotheses are built around a more general research problem.

Examples of educational research problems derived from the general issue of 'equity' are:

- **How large are differences in the stability of school staff between schools in urban and rural areas?**
- **Is the provision of equipment and supplies distributed to schools dependent on public and private funding?**

These research problems can be translated into research hypotheses as follows:

- **The stability of school staff is greater in rural schools than in urban schools.**
- **Equipment and supplies are more widely available in schools dependent on private funding than they are in schools dependent on public funding.**

Characteristics of research hypotheses

Educational research hypotheses should have the following characteristics.

- Describe clearly, and provide identification of the most important variables in operational terms.
- Specify expected relationships among independent, dependent, and control variables.
- Present a statement in a form that is testable with available research methods.
- Be value free in the sense that they exclude the personal biases of the researcher.

Specifying variables and indicators

Following the identification of the research problem and the formulation of researchable hypotheses, it is necessary to prepare a tentative list of variables and indicators for measuring the specific research questions and hypotheses of interest.

A **variable** is a characteristic that can assume two or more properties. If a property can change either in quantity or quality, then it can be regarded as a variable. There are five main types of variable.

- **Dependent variables**
Variables that the researcher is trying to explain (for example, student achievement).
- **Independent or explanatory variables**
Variables that cause, or explain, a change in the dependent variable.
- **Control variables**
Variables that are used to test for a spurious relationship between dependent and independent variables. That is, to test whether an observed relationship between dependent and independent variables may be explained by the presence of another variable.
- **Continuous variables**
Variables that take all values within a particular range.
- **Discrete variables**
Variables that take a number of specific values.

An **indicator** is an empirical, observable, measure of a concept. When an indicator is composed of a combination of variables involving only simple calculations (such as addition, subtraction, division, multiplication, or a combination of these) it is called a 'simple indicator'. When more complex analytical methods, such as factor analysis or regression are used to develop an indicator, the result is referred to as a 'complex indicator'. Examples of simple indicators are: number of school library books per pupil; or teacher/pupil ratio. An example of a complex indicator is a factor score

entitled 'emphasis on phonics' in the teaching of reading formed from three variables: learning letter-sound relationships; word attack skills; and assessment of phonic skills.

Operationalization of research questions

When operationalizing a specific research question to identify an appropriate indicator it is necessary to specify the indicator according to the following components.

- The statistics that will be reported (for example, means or percentages).
- The level of analysis at which the statistics will be calculated (for example, at the student, teacher, or school level).
- The target population and, if any, the sub-populations considered (for example, all primary school students, with the data presented by region, and urban/rural location of the school).
- The specific measure to be used (for example, the number of school library books per student).
- The variables needed in order to calculate a measure on the indicator to be obtained (for example, total school enrolment and number of books in the school library).

Two different indicators of teacher stability were operationalized in data collections conducted by UNESCO and the OECD during mid 1990's. The UNESCO study examined the conditions of primary schools in the least developed countries (Schleicher et al., 1995, pp. 56-59) and the OECD study was focussed on the development of a

broad range of indicators (OECD, 1996, pp. 150-152). These studies offer interesting examples of different approaches to indicator construction. For example, staff stability was defined on the basis of the number of years teachers had been at the school, but the indicator was constructed differently in the two surveys.

In the UNESCO study it was hypothesized that most of the participating countries would be characterized by a high level of staff instability due to population growth and resulting increases in school enrolment rates. Teachers were considered to be 'stable' if they had been at the school for at least three or more years. The level of staff stability for nations was represented by the percentage of teachers in each country who 'had been at the same school for three or more years'. The following variables were needed for this calculation: the overall number of full-time equivalent teachers in the school; the number of teachers having joined the school by year; and, the year of construction of the school building – which functioned as a validity check variable.

In contrast, the indicator of staff stability used by the OECD for developed countries measured the percentage of primary school students enrolled in schools where more than 75 percent of teachers had been employed at the same school for at least five years. In order to build this indicator the following variables were needed: total enrolment per school, the number of teachers per school, and the number of years each teacher had been employed at the school.

Three aspects distinguish these two indicators of school staff stability. First, in the OECD indicator the percentage of stable teachers was weighted by the number of students enrolled. This approach was taken because the goal of the analysis was to provide an indication of how many students were affected by the stability of the teaching force – rather than concentrating on teachers as the unit of analysis. In contrast, the UNESCO study aimed at giving a picture of the teaching body as a whole and therefore employed

teachers as the unit of analysis. Second, stability was defined as teachers being at the school for a minimum of five years in the OECD indicator, and for at least three years in the UNESCO indicator. The reason for the difference between five and three years was that the first study was dealing with a group of the world's most developed countries and the second study concerned developing countries. Third, the OECD indicator defined 'stable' schools as those where more than a certain percentage of teachers (75 percent) were 'stable'. That is, the OECD defined an indicator decision point to distinguish between stable and unstable schools. On the other hand, the UNESCO study, aimed at giving a descriptive picture of the conditions of schooling and therefore did not need to adopt an indicator decision point.

The following table presents the components of the above-mentioned indicators on teacher stability, and highlights the main differences between them.

Table I Analysis of the teacher stability indicators' components

Components of the final indicator	Teacher stability	
	UNESCO data collection	OECD data collection
Statistics	Percentages of teachers	Percentages of teachers
Unit of analysis	Teacher level	Student level
Target population (and sub-populations)	Primary school teachers, with reference to subgroups of schools defined by type (public/private) and location (urban/rural)	Primary school teachers
Operationalization of the indicator	Three years at the school	Five years at the school
Variables needed	a) overall number of teachers b) number of teachers by number of years of permanence at the same school	a) overall number of teachers b) number of teachers at the same school for at least 5 years
Indicator decision points	Not specified	Schools with at least 75% of stable teachers

The problem of the cross-national validity of educational concepts, definitions, and data collection instruments

The specification of variables and indicators presupposes a general and common agreement on the exact meaning and scope of the terms and concepts employed. However, given the diversity that characterizes different education systems, not to mention the disparities that can sometimes be found among regions, and even among schools within the same system, there is a need for a clear and comprehensive definition of these kinds of terms.

In the following paragraphs some key educational concepts and terms are examined to exemplify the kind of definitional problems that arise when dealing with education issues. Some solutions that can be used to address problems in this area have also been described. The definitions and classifications presented draw mainly on UNESCO, OECD, and EUROSTAT work.

I. What is formal education?

a. Problem/issues to be resolved

A number of questions on the scope of education need to be addressed before meaningful data can be collected on key aspects of education systems. For example, when does formal education start and should a data collection on education statistics include the pre-primary level? How should activities in the field of vocational education and training be accounted for? Is special education provided within or outside regular schools, and should it be covered by the data collection? Should adult education be included in the statistics?

b. Helpful hints

Some directions helping to answer these questions can be drawn from the following comprehensive definition of education proposed within the International Standard Classification of Education (ISCED).

Education is ‘organized and sustained communication designed to bring about learning’ (UNESCO, 1976). ‘Communication’ in this context refers to the relation between two or more persons involving the transmission of information. ‘Organized’ means planned in a sequence including established aims and/or curricula and involving an educational agency that organizes the learning situation and/or teachers who are employed (including unpaid volunteers) to consciously organize the communication. ‘Sustained’ means that the learning experience has the characteristics of duration and continuity. ‘Learning’ indicates any change in behaviour, information, knowledge, understanding, attitudes, skills, or capabilities that can be retained and cannot be ascribed to physical growth or to the development of inherited behaviour patterns.

According to this definition, pre-primary school should be included within the specification of education because, not only does it serve the purpose of giving the child daily care while the parents are at work, it also contributes towards the child’s social and intellectual development. One solution to keeping track of differences among pre-primary programmes is to distinguish between ‘all pre-primary programmes’ and ‘pre-primary programmes with special staff qualifications requirements’. The first area covers all forms of organized and sustained activity taking place in schools or other institutional-settings (as opposed to services provided in households or family settings). The second refers to programmes where at least one adult has a qualification characterized by training covering psychological and pedagogical subject matter.

In the area of vocational training one solution is to exclude vocational and technical training that is carried out in enterprises, with the exception of combined school and work programmes that are explicitly deemed to be part of an education system.

The above definition of education also suggests that special education programmes, whether provided by regular schools or by special institutions, are to be included in the data collection as long as the main aim of the programme is the educational development of the individual.

'Adult' or 'non-regular' education programmes should be included in the statistics only if they involve studies with a subject matter content similar to regular education studies or whose qualifications are similar to those of regular education programmes.

Each of the above points provides some idea of the kind of definitional and classificatory work necessary to overcome national and/or regional differences in definitions and thereby to construct data collection instruments which guarantee the comparability of the data that are collected.

2. Distinguishing public and private service providers

a. Problem/issues to be resolved

Most national and cross-national data collections gather information that will enable schools to be classified according to the education service provider. In many data collections for school systems, this classification is often referred to as 'school type'. This is a complex task because of the need to take variety into account. In some countries virtually all education activities and institutions are public. In other countries private agencies play a substantial role. However, the label 'private' covers

a number of different educational configurations. In some countries 'private schools' are entirely or mostly funded by the central authority, but they are run privately. In other countries 'private schools' are entirely or mostly funded and managed privately.

When information concerning 'school type' is collected it is not sufficient to ask the respondent (for example, the school head) to classify the school either as public or private. When developing questions in this area, whether the questionnaire is to be addressed to a central authority or to school heads, it is necessary to specify what it is intended by 'private' vs. 'public', or by 'government' vs. 'independent'.

b. Helpful hints

An approach often adopted is to distinguish between the following three categories of schools: schools controlled by public authorities; schools controlled by private authorities but depending on substantial government funds; and schools controlled and funded by private authorities.

Alternatively, it is helpful to distinguish between the 'control' and 'funding' of schools. The terms 'public' and 'private' can be used to indicate control. That is, whether it is a public or a private agency which has the ultimate power to make decisions concerning the institution (in particular the power to determine the general programme of the school and to appoint the officers who manage the school). The terms 'government' and 'independent' can be used to indicate the source of funding. For example, a government school could be defined as one that receives more than 50 per cent of the funds to support its basic educational services from government agencies and/or whose teaching personnel are paid by a government agency; whereas an independent school could be defined as one that receives less than 50 per cent of its overall funding from government agencies.

3. What is a school?

a. Problem/issues to be resolved

A school is often difficult to define in a manner that is consistent for a cross-national data collection. In some cases a school consists of several buildings, managed by the same head-teacher. In other cases, the same building hosts different schools in different shifts at different times of the day. In some cases a school has a well-defined structure, consisting of separate classrooms – with each classroom being endowed with one teacher table and chair, one desk and chair for each student, and a chalkboard in each classroom. In other cases the school is in the open air, perhaps under a tree, where teachers and students sit on the ground, and the students use their knees as writing places. When collecting comparative information on schools, these different scenarios have to be taken into account.

b. Helpful hints

Suppose, for example, that ‘school crowdedness’ – expressed as square metres of classroom space per pupil – is being measured. The result obtained by dividing the number of square metres by the total enrolment will be correct (and comparable across schools) only in a situation where all schools have one shift. But if some schools operate more than one shift, then the results will be misleading.

One solution in this case would be to ask whether the school has shifts, and how many students there are per shift. The crowdedness measure could then be calculated by taking into account the overall number of students for schools with no shifts, but only the students in the largest shift for schools with more than one shift.

4. What is a student?

a. Problem/issues to be resolved

Suppose that student enrolment figures are being investigated. How will the corresponding statistics be calculated and reported? When the focus of the analysis is on rates of participation, what should be done with repeaters, and how should they be distinguished from students enrolling regularly for the first time in a grade or year of study? All these issues need to be taken into account when designing questions on student enrolment figures for an education system.

b. Helpful hints

A distinction should be made between the number of students and the number of registrations. The number of students enrolled refers to the number of individuals who are enrolled within a specific reference period, while the number of registrations refers to the count of enrolments within a specific reference period for a particular programme of study. The two measures are the same if each individual is only enrolled in one programme during the reference period, but the measures differ if some students are enrolled in multiple programs. Each measure is important: the number of students is used to assess participation rates (compared to population numbers) and to establish descriptive profiles of the student body. The number of registrations is used to assess total education activities for different areas of study.

One solution for calculating student enrolment figures would be to choose a given date in the education programme of interest and then to present the number of students enrolled on that date. Another solution would be to report the average number of students enrolled during the academic year. Yet a third possibility would be to report the total number of students

enrolled during the academic year, with the possibility of double counting multiple entrants and re-entrants.

With respect to identifying repeaters, one commonly applied solution is that students are classified as repeaters if they are enrolling in the same grade or year of study for a second or further time.

5. What is a teacher?

a. Problem/issues to be resolved

How can teachers be defined in order to distinguish them from other educational personnel? One approach would be to base the definition on qualifications. However, this could result in an overestimation of the number of teachers because a number of personnel employed in schools may have a teacher qualification but do not actually teach. Another approach would be to define teachers on the basis of their activities within schools, but this alone would not be sufficient to distinguish professionals from those who may act as teachers occasionally or on a voluntary basis. A further issue is the reduction of head-counts to full-time equivalents (if part-time employment applies). How can part-time teachers be converted into full-time equivalents? No questionnaire concerning teacher characteristics can be designed before these issues have been clarified.

b. Helpful hints

The following definition of a teacher provides a useful framework for overcoming ambiguities:

A **teacher** is a person whose professional activity involves the transmission of knowledge, attitudes, and skills to students enrolled in an educational programme.

The above definition is based on the concepts of (a) activity (excluding those without active teaching duties), (b) profession (excluding people who work occasionally or on a voluntary capacity in educational institutions), and (c) educational programme (excluding people such as some school principals who provide services other than formal instruction to students).

Note that according to this definition, principals, vice-principals, and other administrators without teaching responsibilities as well as teachers without active teaching responsibilities for students in educational institutions are not classified as teachers.

For the reporting of head-counts, individuals who are linked to multiple educational programmes, such as teachers who divide their work between public and private institutions, or between levels of education, or between different functions (for example, teaching and administration) should be pro-rated between those levels, types of institutions and functions. Suppose, for example, that there are 100 full-time teachers that (on the average) devote 80 per cent of their statutory working time to teaching and 20 per cent to the function of headmaster. In this case 80 full-time teachers should be reported under the category 'teacher' and 20 full-time teachers should be reported under the category 'other professional personnel'. If countries cannot pro-rate educational personnel, the classification could be based on the activity to which they devote the majority of their working time.

With respect to part-time conversion, the distinction between full-time and part-time teachers, as well as the calculation of full-time equivalents, is based on the concept of statutory working time. One solution to this conversion problem is to classify as 'full-time' those teachers employed for more than 90 percent of their statutory working time, and as 'part-time' those teachers employed for less than 90 percent of the statutory working time. The classification of individuals linked to multiple educational programmes as full-time

or part-time teachers will depend on the total number of working hours over all levels, types of institutions, and functions.

In some cases the solutions found will be used to define the target population for which data will be collected – as shown in the examples given in the paragraph on formal education. In other cases the definitional work will contribute directly to the design of questionnaire items – as shown in the examples given in the previous section on service provider. In yet other cases definitions and explanations will be used to prepare accompanying notes that provide instructions on how to answer specific questions – as shown in the examples given for the conversion of part-time teachers and students into full-time equivalents).

EXERCISES

1. Prepare three research hypotheses concerning factors that influence student achievement, and then identify some appropriate independent, dependent, and control variables.
2. Specify the variables needed to construct each indicator in the following list.
 - a. Time spent daily for homework by students.
 - b. Teacher academic education.
 - c. Availability of school library.
 - d. Yearly instructional time in Grade I.
 - e. Average class size in school.
 - f. Teacher/pupil ratio.
3. An educational planner has defined an “Indicator of the Level of Qualifications for Primary School Teachers” as the percentage of primary school pupils in schools with at least 50 percent of teachers having completed secondary school. For this indicator specify (a) the statistical units used, (b) the unit of analysis, (c) the target population, and (d) the variables required to construct the indicator.
4. Suppose you have prepared a questionnaire to be answered by all teachers in a sample of schools. How would you specify what is a teacher in the notes accompanying the questionnaire in order to prevent the questionnaire being filled in by other educational personnel?



3

The design of questions

Once the indicators and variables of interest have been identified and their components have been defined, one may begin designing the corresponding questionnaire items. It is important to note that the number of questions in a questionnaire does not coincide necessarily with the number of variables. Sometimes more than one question needs to be asked to operationalize one variable.

This section is concerned with types of questions and response formats. It examines and discusses the advantages and disadvantages of three key types of question structure: open, closed, and contingency. It then gives writing tips for structuring questions and the response categories that accompany them. The section ends with advice on how to avoid response bias and pitfalls in question writing.

Question structure

Two important aspects of questionnaire design are the structure of the questions and the decisions on the types of response formats for each question. Broadly speaking, survey questions can be classified into three structures: **closed**, **open-ended**, and **contingency questions**.

I. Closed questions

Closed (or multiple choice) questions ask the respondent to choose, among a possible set of answers, the response that most closely represents his/her viewpoint. The respondent is usually asked to tick or circle the chosen answer. Questions of this kind may offer simple alternatives such as 'Yes' or 'No'. They may also require that the respondent chooses among several answer categories, or that he/she uses a frequency scale, an importance scale, or an agreement scale.

How often do your parents ask you about your homework?

(Please, circle one answer only)

- | | |
|---------------------------|---|
| Never | 1 |
| 1 or 2 times a week | 2 |
| 3 or 4 times a week | 3 |
| Nearly every day | 4 |

The main advantage of closed questions are:

- the respondent is restricted to a finite (and therefore more manageable) set of responses,
- they are easy and quick to answer,
- they have response categories that are easy to code, and
- they permit the inclusion of more variables in a research study because the format enables the respondent to answer more questions in the same time required to answer fewer open-ended questions.

The main disadvantages with closed questions are:

- they can introduce bias, either by forcing the respondent to choose between given alternatives or by offering alternatives that otherwise would not have come to mind,
- they do not allow for creativity or for the respondent to develop ideas,
- they do not permit the respondent to qualify the chosen response or express a more complex or subtle meaning,
- they can introduce bias, where there is a tendency for the respondent to tick systematically either the first or last category, to select what may be considered as the most socially desirable response alternative, or to answer all items in a list in the same way, and
- they require skill to write because response categories need to be appropriate, and mutually exclusive.

The response format for closed questions can range from a simple yes/no response, to an approve/disapprove alternative, to asking the respondent to choose one alternative from 3 or more response options.

The possibility of format effects or response bias for this type of question can be reduced by changing the sequence of response categories and values. For example, if responses to an item range from 1 to 5, going from negative to positive, then a number of items in the questionnaire can be designed to have 1 as the most positive alternative and 5 as the most negative. This is a particularly important technique for the construction of attitude scales.

Some closed questions may have a dichotomous response format, which means only two mutually exclusive responses are provided.

What is your sex?

(Please tick one box only)

- Male
- Female

For the above example a dichotomous response format is appropriate. However, this type of format should not be overused in a survey because it elicits much less information than multiple choice formats. For example, if seeking information on degree of interest in public affairs, the question “Do you read a daily newspaper?” yields a yes/no response. This could be reworded to: “How many times per week do you read a daily newspaper?”, to which multiple choice responses could be:

1. Seven times a week
2. Five to six times a week
3. Three to four times a week
4. One to two times per week
5. Less than once per week
6. Never

Such a multiple category response format would provide more specific and more useful information than the dichotomous one.

2. Open-ended questions

Open-ended or free-response questions are not followed by any choices and the respondent must answer by supplying a response, usually by entering a number, a word, or a short text. Answers are recorded in full, either by the interviewer or, in the case of a self-administered survey, the respondent records his or her own entire response.

What are your favourite TV programmes?
(Please specify their titles)

.....

What do you like most about school?

.....

The main advantages of open-ended questions are:

- they allow respondents to express their ideas spontaneously in their own language,
- they are less likely to suggest or guide the answer than closed questions because they are free from the format effects associated with closed questions, and
- they can add new information when there is very little existing information available about a topic.

The main disadvantages of open-ended questions are:

- they may be difficult to answer and even more difficult to analyze,

- they require effort and time on behalf of the respondent,
- they require the development of a system of coded categories with which to classify the responses,
- they require the respondent to have some degree of writing ability, and
- respondent handwriting can be illegible.

There is always the possibility with open-ended questions that responses may come in very different forms, and these may lead to answers that cannot be systematically coded for analysis. For example, if asked “When did you leave school?”, the respondent may answer in a variety of ways: “Seven years ago”. “When I got my first job”. “When my brother started going to high school”. “When my parents moved into this house”.

If the survey is administered by an interviewer, appropriate probing helps clarify such answers. In the case of a self-administered survey, guidance by writing specific instructions on how to answer the question can often minimize the number of responses that have very different dimensions.

Care should be taken in writing open-ended questions so as to avoid formats that elicit a dichotomous yes/no or agree/disagree response. In addition, the wording of questions should seek to reduce the possibility of eliciting responses that are aligned along very different dimensions and therefore cannot be systematically coded. For example, asking “What do you think about your school?” can elicit responses such as ‘nothing’ or ‘school is useless’. However, asking “What recommendations would you have for improving your school?” would be more likely to elicit informative answers.

A good case for using open-ended questions is when the aim is to have the respondents reply spontaneously, or when the investigator

is pilot testing the first version of the questionnaire, or when the investigator wants to collect evidence on the parameters of an issue with the aim of later formulating a multiple choice or closed version of a question.

Generally, open-ended questions can produce useful information in an interviewer administered survey, provided that the interviewers are alert and trained to probe ambiguous responses. In self-administered surveys, it is useful to provide instructions on the format of the response that is required so as to minimize opportunities for the respondents to answer the question according to very different dimensions.

3. Contingency questions

A contingency question is a special case of a closed-ended question because it applies only to a subgroup of respondents. The relevance of the question for a subgroup is determined by asking a **filter question**. The filter question directs the subgroup to answer a relevant set of specialized questions and instructs other respondents to skip to a later section of the questionnaire.

The advantage of contingency questions is that detailed data may be obtained from a specific subgroup of the population. Some questions may apply only to females and not to males; others may apply only to people in school, and not to those who are employed. At the base of good contingency questions are clear and specific instructions to respondents.

The formats for filter and contingency questions can vary. One option is to write directions next to the response category of the filter question.

Are you enrolled in secondary school?

1. Yes (answer the following question)
2. No (skip to question 5)

Alternatively, the contingency question can be placed at the end of the questionnaire set apart from ordinary questions that are to be answered by everybody:

ANSWER THIS FINAL SET OF QUESTIONS ONLY IF YOU PLAN ON ENTERING AN ADULT EDUCATION COURSE NEXT YEAR.

OTHERWISE, YOU HAVE NOW COMPLETED THE QUESTIONNAIRE.

Guidelines for writing questions

There are no all-purpose rules that, if followed, will automatically result in a well-written questionnaire. There are, however, some basic principles that, when violated, usually result in respondent confusion, misunderstanding, lack of comprehension, or response bias.

a. Keep the vocabulary simple

A first rule concerns the vocabulary used in writing questions and answer categories. The rule is 'keep it as simple as possible'. This implies using simple words, avoiding acronyms, abbreviations, jargon, technical terms, and abstract or general words.

- If a rare or technical term has to be used, then its meaning should be explained. For example, a question concerning the frequency with which teachers teach their students to understand different styles of text should be accompanied by a definition of each kind of text.

Narrative:

texts that tell a story or give the order in which things happen.

Expository:

texts that provide a factual description of things or people or explain how things work or why things happen.

Documents:

tables, charts, diagrams, lists, maps.

- Acronyms and abbreviations should always be spelled out in the questionnaire. Do not assume that respondents will or should know what an acronym represents.
- When a general term is used, concrete examples should be given to clarify its meaning. For example, a question on learning activities included in the International Educational Achievement (IEA) Reading Literacy Teacher Questionnaire included the following items, for which the respondent had to answer on a four-point frequency scale.

How often are your students typically involved in the following activities?

- silent reading in class
- learning new vocabulary systematically (*for example, from lists*)
- learning to use illustrations (*for example, graphs, diagrams, tables*)

The words 'systematically' and 'illustrations' were too general to be understood in the same way by all respondents. Examples were therefore provided to clarify their intended meaning.

Finally, it is recommended to avoid words that may have an ambiguous meaning. In education, the word 'hour' may have different meanings. For example, many education systems refer to a lesson length or period as an hour even though the lesson is only forty-five minutes long. In order to measure the yearly instructional time at a given educational level, it is therefore necessary to know the length (in minutes) of an 'hour' of instruction, the number of minutes of instruction per week, and the number of school weeks per year. If this information is known, then calculations can be made later for instructional time per day, or week, or year.

b. Keep the question short

Closely related to keeping vocabulary simple is avoiding lengthy questions. Generally, it is recommended to hold questions to 25 words or less. If a longer sentence is used then it should be broken up so that there will be several shorter sentences.

c. Avoid double-barrelled questions

These are single questions that ask for two things and therefore require two answers. "Do you have your own table or your own room to do your homework?" "Do you think it is good idea for children to study geography and history in primary school?" In such instances, respondents do not know what to do if they want to say 'Yes' to one part of the question but 'No' to the other.

d. Avoid hypothetical questions

Evidence has shown that hypothetical questions such as "Would you use this resource in your class if it were available?" are not good for the prediction of behaviour. People are generally poor predictors

of their own behaviour because of changing circumstances and because so many situational variables intervene. Investigators are able to collect more valid data if they question respondents' about their past behaviour and present circumstances, attitudes, and opinions.

e. Don't overtax the respondent's memory

It is risky to ask the respondent to recall past behaviour over a long retrospective period. This is true especially when recurrent events or behaviours are concerned. No student, especially young students, will be able to answer reliably a question such as "In the last month how many hours of homework did you do on an average day?" because the time is just too long to remember what happened in detail. If such a question must be asked, a one-week recall period might be more appropriate for this type of event.

f. Avoid double negatives

Double negatives, either in the question or an answer category (or both), create difficulties for the respondent. For example a statement such as 'Student self-evaluation should not be allowed' followed by agree/disagree is problematic to answer for respondents who are in favour of students' self-evaluation, that is those who do *not* agree that students' self evaluation should *not* be allowed. It is usually possible to solve problems of this kind by formulating the initial statement in a positive way.

g. Avoid overlapping response categories

Answer categories should be mutually exclusive. It should not be possible to agree with or choose more than one category – unless the instructions explicitly allow the respondent to check more than one alternative. Examples of questions with overlapping categories are:

Do teachers generally receive their salaries:

(Check one only)

- usually on time 1
- sometimes a week late 2
- more than a week late 3

How old are you?

- under 20 1
- 20-30 2
- 30-40 3
- 40-50 4
- 50-60 5
- 60 or more 6

The categories of the first question could be made mutually exclusive by removing the qualifiers 'usually' and 'sometimes'. In order to avoid overlap in the second question it should be modified as follows.

How old are you?

- under 20 1
- 20-30 2
- 31-40 3
- 41-50 4
- 51-60 5
- 61 or more 6

h. Beware of ‘leading’ questions

A leading question is a question phrased in such a way that it seems to the respondent that a particular answer is expected. For example:

“Do you favour or oppose school on Saturday morning?”

might read in a leading question as:

“You wouldn’t say that you were in favour of school on Saturday morning, would you?”

or in a more subtle form:

“Would you say that you are not in favour of school on Saturday morning?”

Specifying the characteristics of respondents

Before beginning to write the questionnaire it is important to consider the characteristics of the respondents.

A clear definition of the target population helps to adapt question wording and response formats and also helps to ensure that respondents have experienced what is being asked of them, or at least have sufficient knowledge to be able to respond to the questionnaire items.

In deciding on the sample design, and the population from which the sample is to be drawn, it is helpful to consider whether the population is of individuals, households, institutions, transactions, or whatever. The source from which the data are to be collected is not necessarily identical to the population definition. For example, if a mail questionnaire is sent to school presidents asking about school finance, the population is of schools and not of school presidents.

One of the most important considerations for the researcher is whether respondents consist of a heterogeneous or a homogeneous group. The former consists of individuals who differ from one another in some way that might influence the phenomenon of interest. A heterogeneous group may consist of people from different ethnic backgrounds, different levels of income, and different urban or rural areas. By contrast, homogeneous groups consist of individuals from similar socio-spatial backgrounds.

Research has shown that response rates are usually higher for homogeneous or select groups (for example, high school teachers, university professors, physicians) because they are more likely to identify with the goals of the study. Beyond this distinction, it is known that interest and familiarity with the topic has a positive effect on response rates.

A checklist for reviewing questionnaire items

The following list of questions provides a framework for reviewing each item that is to be included in a questionnaire.

1. Will the item provide data in the format required by the research questions or the hypotheses?
2. Is the item unbiased?
3. Will the item generate data at the level of measurement required for the analysis?
4. Is there a strong likelihood that most respondents will answer the item truthfully?

5. Do most respondents possess sufficient knowledge to answer the item?
6. Will most respondents be willing to answer the item, or is it too threatening or too sensitive?
7. Does the item avoid 'leading' respondents to a specific answer?
8. Is the language used in the questionnaire clear and simple – so that all respondents are able to understand all of the questions?

EXERCISES

1. Explain the uses of closed, open, and contingency questions.
2. Draft five closed and open questions related to some aspect of educational research.
3. Formulate a contingency question with accompanying instructions.
4. Here are some 'bad' questions which contain some of the problems presented in the above discussion. List the main problems and then redraft each question to address these problems and explain the changes that you have made.

...

... EXERCISES

Question 1

How many teachers are there in your school who have been at the school for at least five years and who are involved in special initiatives outside the normal class activities at least once per week?

..... teachers

Question 2

Do you enjoy studying English and Mathematics?

Yes

No

Question 3

If you could attend university which subjects would you like to study?

.....

Question 4

In the last six months how many times did you teach your students to read expository materials?

.....

Question 5

Sometimes teachers do not give me sufficient attention.

Definitely	Mostly	Mostly	Definitely
Disagree	disagree	agree	agree
1	2	3	4

Question 6

What is the condition of each of the following in your school?

	Bad	Good
Lighting	1	2
Water	1	2
Canteen	1	2
Water taps	1	2



4

Examples of questions

In the following discussions some examples have been presented of the main types of questions that are often used in educational planning and educational research data collections. These questions cover the areas of student background, teacher characteristics, school location, learning/teaching activities, and attitudes. The examples related to attitude scales include a discussion of the principles of Likert scaling and the method used for the calculation of the discrimination power of attitude scale items.

Student background

Demographic questions are designed to elicit information from respondents concerning their personal characteristics and social background. This type of information is important for explaining variations in educational outcomes and behavioural patterns. The most frequently used demographic questions focus on gender, age, level of education, income level, marital status, level of parents' education, religion, and ethnic background. A number of these areas cover sensitive and personal issues and therefore need to be handled carefully.

I. Gender and age

Data on student **gender** is critically important for examining issues of gender equity in all school systems. This information can be gathered from the class attendance register or can be asked as part of a student questionnaire.

While a question on gender can be asked irrespective of the student's age, for younger students the question can be more suitably phrased in the following manner.

Are you a boy or a girl?

(Tick one box only)

- Boy
- Girl

If the same information is to be obtained from teachers, the question form and the phrasing of the response alternatives can be more direct.

Your sex:

- Male
- Female

Whether the question is phrased as a standard question or a more direct enquiry, it is always important to seek advice concerning the wording of this type of question so as to respect local customs and culture.

Student age is another important variable for explaining the structure and evolution of an education system, and for examining the educational development of students over time. Data on student age can also be obtained from the class register or from a question included in a student questionnaire.

There are various ways in which information on student age can be collected. One way is to ask for the age at a specific reference date; another is to ask for the actual birth date. In this latter case, depending on the degree of accuracy needed, the respondent can be asked to specify the year, or the year and month, or the year, month, and day as in the following example.

What is your date of birth?
(Please write the corresponding numbers in the spaces below)

Day	Month	Year
--	--	----

Research has shown that the most accurate way to obtain information about age is to ask both date of birth and age at last birthday:

How old were you at your last birthday?

Age:

2. Socio-economic background: occupation, education, and possessions

Another type of background characteristic concerns the **socio-economic background** of the student. Indicators can be developed using information obtained directly from the individual, or by using either objective or subjective responses. Some common indicators of student socio-economic background are the parents' level of income, their occupational status, their level of education, and the personal possessions in the home. The measurement of parent income is always a difficult task in all countries of the world – and for many different reasons. Most school-aged children cannot answer such questions accurately. Similarly, adults sometimes have difficulty answering income questions because they are not in an occupation with a regular salary or because questions in this area represent an invasion of personal privacy. It is usually not useful to include a question on parent income when the respondent is under 15 years of age. For this reason, parents' level of education, occupational status, and home possessions are the most frequently-used proxy indicators of household wealth.

a. Parent occupations

Parent occupations are usually grouped into 'occupational status' categories based on levels of education, skill, and income. The categories are then ranked from lowest occupational status to highest. The categories used to group the occupations must reflect the range of occupations existing in society, and they must also be comprehensible to the respondent. Terms such as white-collar, blue-collar, professional, skilled, semi-skilled, unskilled are not easily understood by younger children if left undefined.

The following is an example of a set of questions directed to students on their father's occupation. The question begins with a filter to ascertain that questions 11 and 12 are asked only

of students whose father is working. In Question 11 the open responses will later be coded by the survey researcher. The third question in the series helps check for consistency between responses.

10. Does your father work for pay now?

Yes, he now works full time 1 (GO TO 11)

Yes, he now works part time 2 (GO TO 11)

No, he is now looking for work 3 (GO TO 13)

No, he is not working at present
(unemployed, retired) 4 (GO TO 16)

11. What is the occupation of your father (or the male person responsible for your education)?
(Please, describe as clearly as possible)

.....

.....

.....

.....

12. In your opinion, how can the occupation of your father be defined? (Please tick only one box)

professional and managerial

clerical and sales

skilled blue-collar

semi-skilled and unskilled

Before the non-compulsory level of education is reached (but after the age of 10 years), it is often recommended to use a classification of occupations with no more than 4 categories. For example, (a) professional and managerial, (b) clerical, (c) skilled blue-collar, (d) semi-skilled and unskilled. Each category should have a definition and/or examples that correspond to the type of occupation.

Obtaining useful information on parent's occupation is very difficult in self-administered questionnaires. For most children under the age of 14 years an interviewer or data collector should assist with this question so as to improve the quality of the children's responses. Some examples of interviewer probes are presented below.

Response:

My father works in a factory

Probe:

What kind of machine does he operate?

Response:

My father is a teacher

Probe:

What level does he teach (or, alternatively, what age students does he teach?)

Response:

My father works in a shop

Probes:

Does he own the shop? Or does he manage the shop?
Or does he work for someone else in the shop?

In the absence of interviewers, the following guidelines can improve response quality in self-administered questionnaires: (i) avoid simply asking the name of the place where the parent works since this is insufficient in detail. For example, a response 'in a hospital' could mean the father is a doctor, a nurse, an administrator, or a janitor; (ii) avoid asking vague job titles, such as 'engineer', (iii) look into the job classifications used in a recent census or national

population surveys and see if they can be adapted for use in a student questionnaire, (iv) when asking the mother's occupational status, remember to include the option 'housewife' (or home duties), and (v) a combination of open and closed questions is often more effective since it permits a check to be made of the consistency of the responses.

b. Parent's education

Open-ended questions that ask directly for the number of years of a parent's education are very difficult to answer because they imply that students remember not only the level of education completed by their parents, but also the sequence of levels of education, and that they know the duration of each of these levels. For these reasons, questions on parent's education should be given in multiple choice format.

What is the highest level of education that your father (or the male person responsible for your education) has completed?

*(Please tick **one** box only)*

- Never went to school
- Completed some primary school
- Completed all of primary school
- Completed some secondary school
- Completed all of secondary school
- Completed some education/training after secondary school
- Don't know

In this case the only task required is to recognize the correct information, rather than to remember it. Once the responses to this question are collected, the first six options can be ranked during coding, from one to six, or they can be converted into the number of years corresponding to each option (perhaps using a median value for the second, and fourth option).

Where possible, the researcher should make a prior request for information from the parents. This could be achieved by asking the students to consult with their parents – or by having the students take the following question home for completion by parents.

**How many years of academic education has your father/
guardian completed?**

- years of primary school
- years of secondary school
- years of post secondary academic education

c. Possessions in the home

Possessions in the home has become a useful alternative approach to collecting information on socio-economic background from students. The number of books in the home also provides information about possessions that are more closely linked to the educational level of parents. Variables based on this information usually yield a strong relationship with educational outcomes – even when reported by younger students.

The items included in the list must relate to the context in which the questionnaire is administered, and to the level of development and characteristics of the society. It is important that the list include possessions that denote high, medium, and low economic status in order to discriminate among students with different socio-economic backgrounds.

One of the most important possessions related to the social (and educational) climate of the home is the number of books. This information is usually collected in approximate categories – which must be defined with a detailed knowledge of prevailing societal conditions.

About how many books are there in your home?*(Do not count newspaper or magazines. Please tick one box only)*

- None
- 1-10
- 11-50
- 51-100
- 101-200
- More than 200

The data gathered using this kind of question are at best very approximate. However, experience has shown that these data are generally highly correlated with educational outcomes.

An alternative approach is to develop a 'checklist' of possessions. The items presented on the list need to acknowledge the general economic development level of the countries where data are being collected. For example, the following list of items would be less useful in wealthy countries (because most homes have all of the items) and much more useful in developing countries where a summarization of the total number of items in each home would provide a good level of discrimination among relatively poorer and relatively wealthier homes.

Which of the following items can be found in your home?*(Please, circle one number for each line)*

	No	Yes
Radio	1	2
TV set	1	2
Video cassette recorder	1	2
Telephone	1	2
Refrigerator	1	2
Car	1	2
Piped water	1	2

When investigating home conditions the possibility that some students do not live at home while attending school needs to be considered. In this case the term 'home' may need to be replaced with a more generic expression such as 'the place where you stay during school week'.

A number of issues have to be taken into account when using socio-economic background data in educational research. One, that has already been mentioned, is the limited capacity of young children to provide an accurate report of this kind of information. Another issue has to do with the confidential nature of the information requested. The kind of questions that can be asked in this area without risking the validity of the data depends on local culture and customs. Yet another issue is related to social changes that have occurred over the last two or three decades. In particular, the increasing participation of women in the labour force requires a revision of occupational status scales (and of their meaning). Existing scales were mostly developed with data from samples of adult males in developed countries. Similarly, the increased incidence of divorce and of single-parent families requires a major revision in terminology when asking questions that refer to the father and/or to the mother.

Teacher characteristics

Among the teacher's characteristics of interest in educational data collection are gender, age, education, and years of teaching experience. At the school level this information can be collected either from teachers themselves or from school heads. However, asking teachers to answer a question such as 'How many years of education have you completed?' provides very little information. Such a question neither specifies whether pre-service training is to be included in the 'years of education', nor provides information on years of grade repetitions (if any) or on whether part-time years of attendance were converted into full-time years equivalent.

In seeking information on a teacher's educational background, questions should distinguish between academic education and pre-service teacher training and they should ask the respondent to specify how many years were attended for each level of education. Clear instructions also need to be provided on how to treat grade repetition and part-time attendance.

1. How many years of academic education have you completed?

(Do not count grade repetition years. Part-time years should be converted into full-time years. For example, two half-years equals one full year)

- ___ ___ years of primary school
- ___ ___ years of lower secondary school
- ___ ___ years of upper secondary school
- ___ ___ years of post secondary academic education

2. How many years of pre-service teacher training have you received altogether?

(Please, circle one number only)

- a. I did not receive any teacher training
- b. I have had a short course of less than one-year duration
- c. I have had a total equivalent of one year
- d. I have had a total equivalent of two years
- e. I have had a total equivalent of three years
- f. I have had a total equivalent of more than three years

The first question reveals how many years of academic education were completed altogether and the level of academic education that was reached. The distinction between the different levels of education helps to verify and improve the precision of responses. The second question covers different scenarios, ranging from no teacher training, up to more than three years of teacher training.

School location

The location of a school is often a key issue in data collections because physical location is often strongly related to the sociocultural environment of the school. In addition, the degree of physical isolation of a school can have important impacts on decisions related to staffing and infrastructure costs.

Consider the following question on school location.

In what type of community is your school located?

(Please tick one box only)

- A geographically isolated area
- A village or rural (farm) area
- On the outskirts of a town/city
- Near the centre of a town/city

In the above example, the first and second response categories are not mutually exclusive: a village or rural area may also be in a geographically isolated area. A second problem with the formulation is the ambiguity in the fourth response category – it is not clear what is meant by ‘Near the centre of a town or city’.

The following example shows a reformulation of the question that should improve the quality of the information obtained:

What type of community is served by your school?

(Please tick one box only)

- A village or rural community
- A small town community
- A large town
- A city

The second, third and fourth alternatives in the above question discriminate between urban centres of different sizes. One could add indications for population size for these response categories. For example, a small town community (between 50,000 and 150,000 inhabitants); a large town community (greater 150,000 and under 1 million); and a city of 1 million or more people. The number of inhabitants per category will depend on the demography of the country, and can be determined by looking at the geographical population distribution as reported in the most recent census data. If a dichotomy of urban/rural is to be made during the data analyses then category 1 could be used for rural, and categories 2, 3 and 4 could be combined for urban.

In a less densely populated country the following question may provide more accurate information concerning school location. The aim here is to identify the location of the school with respect to important external services.

How many kilometres by road is it from your school to the places in the list below?:

- (a) the nearest health centre/clinic kilometres
- (b) the nearest asphalt/tarmac/tarred road. kilometres
- (c) the nearest public library kilometres
- (d) the nearest secondary school. kilometres
- (e) the nearest city kilometres
- (f) the nearest regional capital. kilometres

The list of items for which the distance in kilometres is asked can vary according to the focus of the survey and the characteristics of the country. Whatever items are used, the number of kilometres can be summed for all items and then divided by the number of items as a measure of the degree of isolation of the school.

Learning, teaching, and school activities

Questions on activities, whether addressed to students, teachers, or school heads, usually employ either a 'yes-no' response format, or they ask for an evaluation of frequency or importance. Some examples for students, teachers, and school heads have been presented below.

1. Student reading activity

How often do you read these types of publications for personal interest and leisure?

Publication type	Rarely	Less than once a month	One or two times a month	About once a week	Two or three times a week	Most days
Mystery	1	2	3	4	5	6
Romance	1	2	3	4	5	6
Sport	1	2	3	4	5	6
Adventure	1	2	3	4	5	6
Music	1	2	3	4	5	6
Nature	1	2	3	4	5	6

In a question of this kind it is important to set the time points defining the scale so that they make sense in relation to the specific activity of interest, and to the purpose for which the data are collected. In general the expressions 'never' and 'always' should be avoided, as they are extremes that respondents tend to dislike. This is why in the above example, categories were formulated as 'rarely' and 'most days'. The categories are defined so that they do not overlap, and are not too distant from one another.

2. Teacher activities

During the school year, how often do you teach comprehension of each of the following kinds of text?

(Circle one number per line)

Kind of text	Almost never	About 3-4 times a year	About once a month	At least once a week	Nearly every day
a. Narrative text <i>(that tells a story or gives the order in which things happen)</i>	1	2	3	4	5
b. Expository text <i>(that describes things or people, or explains how things work or why things happen)</i>	1	2	3	4	5
c. Documents <i>(that contain tables, charts, diagrams, lists, maps)</i>	1	2	3	4	5

The categories in the above example are specified on the assumption that teaching takes place regularly. Therefore, if a teacher employs a specific teaching strategy more than a few times per year, he/she probably does it on a monthly, weekly, or daily basis. A time scale needs to be constructed in relation to the variable(s) of interest.

The question above provides a good example of how questions and their components should be specified. Suppose the question was formulated as: "How often do you teach reading of the following

kinds of texts?" In this case the information obtained would be much less useful because reading is a complex activity, made up of skills that range from basic decoding to sophisticated inferences. It is therefore necessary to specify what aspect of reading is being taught (for example, 'understanding') and the segment of teaching time on which teachers should base their answer. Should they think of an average class in an average year? If so, how representative is that year and class of their teaching experience? In asking these kinds of questions it is important to specify the class and year (for example, 'in your class this year').

3. School head activities

The school head's report of education-related activities within a school is very important. There is ample research evidence which shows that schools that foster a wide range of educational and cultural activities outside the classroom also have more effective reading programmes.

Does your school have any special programs or initiatives for reading outside normal classroom activities?

(You may tick more than one)

- | | |
|---|---|
| Extra-class lessons in reading | <input type="checkbox"/> |
| Extra-individual tuition in reading at school | <input type="checkbox"/> |
| Special remedial reading courses | <input type="checkbox"/> |
| Other | <input type="checkbox"/> (specify) |
| None | <input type="checkbox"/> |

This question is a modified form of a yes/no question. Ticking a response category corresponds to answering 'yes', and leaving it blank corresponds to answering 'no'. It is important to include the category 'other' if the items listed are not exhaustive.

The following question asks respondents to rank different items related to the work of school heads.

Please rank the following activities in order of importance in your work as a school head
 ('1' is the most important activity, '6' is the **least** important activity)

	Importance ranking
(a) evaluating the staff
(b) discussing educational objectives with teachers
(c) pursuing administrative tasks
(d) organizing in-service teacher training courses
(e) organizing extra-class special programs
(f) talking with students in case of problems

This kind of question makes it impossible to score all items as 'very important', and the respondent is forced to a ranking in order of importance.

The advantages of rank order methods are basically that it is easy for respondents to understand the instructions, and the questions force discrimination among objects. One of the disadvantages is that forced responses, may not yield a real degree of preference or attitude, but rather information that the respondent prefers one object over another.

Remember that *ranking* and *rating* are two different processes. Ratings are assigned independently to each item. Ranking requires that a set of items be placed in order, thus providing a comparison of each item to all others.

Attitudes, opinions, and beliefs

An attitude is often defined as a state of readiness, or a tendency to respond in a certain manner when confronted with particular stimuli. Social psychologists consider that attitudes arise from deeply rooted personality characteristics and value systems within individuals, and that they become manifest in the form of opinions. The main difficulties in measuring attitudes are that (a) the object of an attitude can range from the very specific to the very general, (b) attitudes are not static, and (c) attitudes are both shaped and changed by socio-demographic circumstances and life experiences.

In the field of educational research, the measurement of attitudes has become an important issue in attempts to monitor the 'affective' (that is non-cognitive) outcomes of schooling. The most popular approach to attitude measurement has been via the use of attitude scales.

Attitude scales usually consist of a number of attitude statements which are presented to respondents with a request that they should indicate whether they agree or disagree. Scaling techniques are deployed to order respondents along some underlying attitudinal continuum.

1. Likert scaling

Likert scaling is the most frequently applied attitude scaling technique in educational research. It consists of six main steps.

Step 1 Determining the attitude to be measured

In the field of educational planning some of the more important areas of attitude measurement include pupil attitudes towards school, teachers, and school subjects. In addition, given the importance of retaining good teachers within school systems, there

has been a growing interest in measuring the sources of teacher satisfaction with, and attitudes towards, their professional work as teachers.

Step 2 Listing possible scale items

Here, a set of statements, or a series of items, are devised that express a wide range of attitudes, from extremely positive to extremely negative. The statements are designed to reflect favorably or unfavorably on the object of the attitude.

One common approach for constructing these statements is to organize a discussion focussed on the stimulus for the attitude (for example, the quality of school life) with individuals or small groups representative of the target population to whom the scale will be administered. The various negative and positive comments and statements made during this discussion may be selected and edited for use as stimuli in the attitude scale. Another approach is to ask a sample of respondents to respond to a set of open-ended statements related to the attitude being investigated. These responses are then used to construct attitude statements.

Each statement is followed by an agreement scale on which respondents are requested to indicate the degree to which they agree or disagree with each statement.

Although the scale may have only two choices (agree/disagree), more choices may sometimes permit a finer distinction in the intensity of the attitude. Generally, Likert scales have 5 categories (strongly like, like, neutral, dislike, strongly dislike). Occasionally, the neutral or middle category may be omitted, forcing respondents to express an opinion for each statement.

It is usually recommended that an equal number of positive and negative statements be used. For positively worded statements the scoring categories are as follows below.

Strongly agree	agree	uncertain	disagree	Strongly disagree
5	4	3	2	1

For negatively worded statements the scoring is reversed so that 'strongly agree' would be scored as '1', and so on, with 'strongly disagree' being scored as '5'.

Step 3 Administering items to a sample

In this step a sample of respondents, selected randomly from the population to be studied, is asked to indicate attitudes with respect to the list of items drawn up in Step two. For this trial-testing phase a sample of around 150 to 250 respondents from a wide range of environments is normally required so as to provide stable statistical analyses in the following steps.

Step 4 Computing a total score

The researcher calculates a total score for each respondent, by summing the values of all items. Take the following example adapted from a 'Quality of School Life' scale designed for 14 year-old students. The scale aims to measure the attitudes of students towards school in terms of their 'well-being'.

Suppose respondent X had the following response pattern for the above 5 items: Agree, Disagree, Strongly Agree, Neither agree nor disagree, Disagree. The total score computed for respondent X would be:

$$4 + 4 \text{ (Negative item)} + 5 + 3 + 4 \text{ (Negative item)} = 20$$

1. School is a place where I usually feel great

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

2. The teachers at my school are often unfair

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

3. I really like to go to school

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

4. Going to school makes me feel important

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

5. School is a place where I sometimes feel depressed

- Strongly agree
- Agree
- Neither agree nor disagree
- Disagree
- Strongly disagree

Step 5 Analyzing the item responses (pre-testing the items)

In this step it is necessary to determine a basis for keeping some items for the final version of the measurement scale, and discarding others. This can be done either through correlational analysis or by item analysis that yields a discrimination coefficient for each item. The discrimination power of an item is a measure of its ability to differentiate the high-scoring respondents (clearly positive attitudes) from the low-scoring respondents (clearly negative attitudes).

Many standard computer packages, such as 'Reliability' in SPSS facilitate this by calculating the correlation of each item with the total score. As a general set of benchmarks: items with a correlation (with the total score) of under 0.3 are considered to be 'weak', and 'good' items would have a correlation (with the total score) of around 0.5 or higher.

Items are often discarded if they show negligible or no variation across respondents. For example, if almost all respondents answered that they 'strongly agree' with item one ('school is a place where I usually feel great'), then this item is simply adding a constant to all scores.

Step 6 Selecting the scale items

The final list of items for the attitude scale is selected from among those trial items that have (a) high discrimination, and (b) a range of mean response scores. The need for high discrimination has been mentioned above. The need for a range of mean response scores arises because this results in more reliable measurement of the respondents along the full range of the total scores.

In writing attitude statements it is recommended that items be worded positively **and** negatively so as to avoid the 'response set' – which is the tendency to respond to a list of items of the same format in a particular way – irrespective of content.

a. Problems in the design of rating questions

- Error of proximity: the tendency to rate items similarly because they are near to each other in the questionnaire.
- Central tendency error: the tendency to rate most items in the middle category (when the middle category is offered). Such respondents either dislike extreme positions, or lack knowledge.
- Error of leniency: the tendency to give high ratings to most items by liking or agreeing with everything.
- Error of severity: the opposite to the error of leniency: respondents who dislike, or disagree, with most items.
- Halo effect error: the tendency to rate a particular statement according to how respondents feel about it in general. For example, giving a very low rating to statements such as 'I enjoy reading', 'I like to borrow library books', and 'I prefer to read something every day' because of a dislike for the reading teacher.

b. Assumptions in Likert Scaling

It is important to note that the following assumptions underly this scaling technique:

- That there is a continuous underlying dimension which is assessed by total scores on the attitude scale and that each item contributes in a meaningful way to the measurement of this dimension.
- That a more favorable attitude will produce a higher expected score, and vice-versa.

- That items are of equal value in that they each provide a replicated assessment of the dimension measured by the total score on the scale.

Other scaling techniques that rely on attitude statements are the Thurstone Scale (1928) and the Guttman Scale (1950). The Osgood Scale (1957), also referred to as the semantic differential technique, is composed of pairs of adjectives to measure the strength and direction of the attitude.

EXERCISES

- I. Here are three questions on student age. Discuss their suitability for students of primary, secondary and post-secondary education level.

How old are you?

.....

What is your date of birth?

.....

What is your date of birth?

(Please write the corresponding numbers in the spaces below)

Day Month Year

-- -- --

...

... EXERCISES

2. Specify ten items that would be appropriate to include in a question on home possessions in order to measure the socio-economic background of pupils in your country.

Which of the following things can be found in your home?
(Please circle one number for each line)

	No	Yes
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2
.....	1	2

3. Consider the indicators (a) to (d) presented below.
- (a) Teacher years of teaching experience.
 - (b) Primary school and grade enrolment by gender.
 - (c) Instructional time per year for Grades 1, 3, and 5.
 - (d) Pupils' interest in reading.
 - Decide whether one or more variables are required for each indicator.
 - Decide if one or more questions are required for each variable.
 - Write the questions.
4. Draft ten attitude statements (each with 5 scale response categories) that could be used to construct a scale for measuring student attitudes towards mathematics.



Moving from initial draft to final version of the questionnaire

This section looks at the ordering of questions in the questionnaire, the training of interviewers and administrators, pilot testing, and the preparation of a codebook. It gives advice on how to design the layout of the questionnaire, including instructions to respondents, interviewer instructions and introductory and concluding remarks. Guidance is provided on how to trial test and then use the results of this to improve the final form of the questionnaire.

Two widely-used patterns of question sequence

Two widely-used patterns of question sequence in questionnaire design have been found to motivate respondents to co-operate and fully complete a questionnaire. They are called the *funnel sequence* and the *inverted funnel sequence*.

The characteristic of the funnel sequence is that each question is related to the previous question and has a progressively narrower scope. The first question can be either open format, or multiple choice. It should be very broad, and is used to ascertain something about the respondent's frame of reference on a topic. This ordering pattern is particularly useful when there is a need to prevent further specific questions from biasing the initial overall view of the respondent.

- 1. Would you say that the general quality of education provided by primary schools in your community is:**
 - very good
 - good
 - uncertain
 - bad
 - very bad

- 2. How would you rate the overall quality of the primary school attended by your child?**
 - very good
 - good
 - uncertain
 - bad
 - very bad

- 3. Do you think your own child is receiving a good primary school education?**
 - Yes
 - No

- 4. Given the opportunity, would you have your child attend another school in your area?**
 - Yes
 - No

In the inverted funnel sequence, specific questions on a topic are asked first, and these eventually lead to a more general question. This sequence requires the respondent to think through his or her attitude before reaching an overall evaluation on the more general question. Such a question order is particularly appropriate when there is reason to believe that respondents have neither a strong feeling about a topic, nor a previously formulated view.

The placement of items in a questionnaire requires careful consideration. Good item placement can increase the motivation of respondents – which in turn results in more valid data.

General guidelines for item placement

1. Non-sensitive demographic questions should be placed at the beginning of the questionnaire because they are easy to answer, non-threatening, and tend to put the respondent at ease.
2. Items of major interest to the research study should be placed next since there is greater probability of the respondent answering or completing the first section of the questionnaire.
3. Sensitive items that cover controversial topics should be placed last so that potential resentment that may be provoked by these items does not influence responses to other questions.
4. Items on the same topic should be grouped together. However, care should also be taken to prevent one item influencing responses to later items.
5. Items with similar response formats should be grouped together when several different response formats are being used within a questionnaire.
6. Section titles should be used to help the respondent focus on the area of interest.

Covering letters and introductory paragraphs

If the questionnaire is to be mailed, or distributed, for a respondent to complete, it is important to have a covering letter. The purpose of such a letter is to explain the object of the survey, and to encourage respondents to complete the questionnaire. In an interview, one of the tasks of the interviewer is to persuade the respondent to co-operate. In a self-administered questionnaire, the covering letter is the only instrument for overcoming resistance. For this reason, the covering letter is important, and should do the following:

- Identify the organization conducting the study (for example, the Ministry of Education).
- Explain the purpose of the study.
- Assure the respondent that information provided will be managed in a strictly confidential manner and that all respondents will remain unidentified.
- Explain WHY it is important that the respondent should complete the questionnaire.
- Provide the name and contact numbers of the Principal Researcher.

The following additional information should also be included in both the introduction to the questionnaire and the covering letter:

- Brief detail on how the respondent was selected (for example, 'Your name was randomly selected ...').
- Expression of appreciation for the respondent's help.
- Estimate of questionnaire completion time.

EXAMPLE COVERING LETTER

Date: 25 September 2000

To: Participants in Reading Literacy Teacher Questionnaire

As someone currently involved in the teaching of reading literacy, the Ministry of Education would greatly appreciate a few minutes of your time to respond to the enclosed questionnaire.

The results of this study will determine the reading literacy levels of primary school students and this information will be used as part of a review of teacher pre-service and in-service training programmes.

You were randomly selected from a pool of currently employed primary school teachers. You will not be identified by name. All information provided by you will be treated as strictly confidential.

The questionnaire should only take 15 minutes to complete. Please return it in the enclosed postage-paid envelope by 20 December 2000.

Your participation is very much appreciated and will allow us to focus on critical issues related to the teaching of reading literacy as determined by experienced teachers.

Yours sincerely,
xxxxxxx

If a good covering letter is enclosed with the questionnaire, the introductory paragraph on the questionnaire itself may be shorter and contain some instructions for responding to questions. The IEA Reading Literacy Questionnaire for Teachers had the following introductory paragraph.

The following questionnaire is part of an International study of Reading Literacy and attempts to identify differences in English instruction. It is recognized that teachers are likely to respond quite differently to the enclosed questions.

Please answer all questions in such a way as to reflect most clearly your teaching practices. Most questions will require you to circle your selected response. Others will require you to write down a number. Do not leave blanks.

We thank you for your contribution to this important research.

Drafting instructions for answering questions

Writing instructions for answering questions is a very important part of the questionnaire layout. If the questionnaire is to be administered by an interviewer, then the instructions will be addressed to him or her. Such instructions are usually written in capital letters, as follows.

Who was your employer on your last job (PROBE FOR CORRECT CATEGORY)

- Private
- National Government
- City
- Self-employed
- Public, non profit
- Other
- Doesn't know

In a mailed or self-administered questionnaire, it is very important to provide clear instructions because there is no resource person to help clarify respondents' queries. Instructions can be for a single question or for a set of questions.

INSTRUCTIONS TO A RESPONDENT FOR A SET OF QUESTIONS

INSTRUCTIONS: For each of the following questions, please mark the answer that comes closest to the way you feel about learning mathematics. There is no right or wrong answer. Answer the questions in the order in which they appear on the paper. Thank you for your co-operation.

INSTRUCTIONS TO A RESPONDENT FOR A SINGLE QUESTION

About how many different teaching positions have you held during your life? (*Count only those teaching positions that you have held for at least one full academic year*)

The following examples provide illustrations of different instructions given for the same question. In the first example, the instructions relate to an interview. In the second example the instructions relate to a self-administered questionnaire. Note that the question is multiple choice, followed by an open ended contingency question.

INTERVIEW FORMAT

- I. Thinking about government facilities provided for schools, do you think your neighborhood gets better, about the same, or worse facilities than most other parts of the city?

Better (ASK A)	1
About the same	2
Worse (ASK A)	3
Don't know	8

- IA. If **better** or **worse**:: In your opinion, what do you think is the main reason why your neighbourhood gets (better/worse) facilities?

.....
.....
.....

SELF-ADMINISTERED FORMAT

I. Thinking about the government facilities provided for schools, do you think your neighborhood gets better, about the same, or worse facilities than most other parts of the city?

- Better 1 (answer 1A below)
- About the same 2
- Worse 3 (answer 1A below)
- Don't know 8

1A. If **better** or **worse**: In your opinion, what do you think is the main reason why your neighbourhood gets (better/worse) facilities?

.....

.....

Training of interviewers or questionnaire administrators

Frequently, the testing of a questionnaire is undertaken by interviewing respondents – even if the final version of the questionnaire is to be self-administered. This implies, however, that the questionnaire administrators and the interviewers receive an appropriate level of basic training before setting out to pilot the questionnaire.

All questionnaire administrators and interviewers should be given written instructions so as to ensure that each respondent receives the questions in the same format and with the same instructions. If the interview is to be administered to young children or adults who cannot read, the interviewer should be given a card on which this information is written. The interviewer should be instructed to read the card to each respondent in the same way.

The interviewer should be instructed on the amount of direction to give to the respondent. This can range from very little to a large amount. If possible, it is useful to have the interviewer participate in simulation exercises by both answering respondent questions and interviewing respondents. Having a few people observe the simulation helps in giving comments and suggestions for improving an interviewer's techniques.

If the questionnaire is to be self-administered, then the data collector should be given instructions on how to introduce the questionnaire as follows.

1. Say: 'I am Mr/Mrs. from the (*National Research Centre for Educational Planning*), and we are interested in knowing your views on the (*education profession*)'.
2. Say: 'We would greatly appreciate your completing this questionnaire, which should only take 10 minutes. The directions for filling it in are given on the front page'.
3. Hand the respondent the questionnaire (and a pen or a pencil).
4. Clarify the questions that the respondent(s) may have about the instructions.
5. If there are questions about particular items, simply respond: 'Just answer the question as you interpret it'. Alternatively if more guidance is necessary the interviewer or administrator could be instructed to 'clarify all questions about the items'.
6. Note on the back of this sheet any questions respondents had about items, or any comments or remarks concerning the questionnaire (for example, too long, too hard to understand, too difficult).
7. Thank the respondent when he/she completes the questionnaire.

Pre-testing the questionnaire

Pre-testing the questionnaire is an essential step before its completion. The purpose of the pretest is to check question wording, and to obtain information on open-ended questions with a view to designing a multiple choice format in the final questionnaire. Pre-testing has a number of very important advantages.

1. Provides information on possible ethical problems overlooked previously.
2. Helps determine if the research questions or hypotheses are appropriate.
3. Helps determine if the levels of measurement are appropriate for the selected variables.
4. Provides a check that the population is appropriately defined.
5. Provides information on the feasibility and the appropriateness of the sampling method.
6. Helps determine sample size by allowing estimation of variance from the pre-test sample.
7. Provides additional training for interviewers, instrument administrators, experimenters, coders, and data editors.
8. Helps determine the length of the questionnaire.

After training the interviewers and questionnaire administrators, the next step in pre-testing is to select a small pilot sample of respondents that covers the full range of characteristics of the target population. In the field of education this usually implies that the

pilot sample includes appropriate gender balance and covers a range of richer/poorer and rural/urban communities.

Pre-testing should never be carried out on a 'convenience sample', (for example, the researcher's friends or family, or schools in one neighbourhood of the Capital city). For interview questionnaires 50 interviews will provide solid material for verifying question wording, sequencing, instructions and general quality of the instrument. However, larger samples of around 200 are required to calculate various statistics such as discrimination co-efficients.

Note that even questions 'borrowed' from existing questionnaires need to be pre-tested to ensure that they will work as required with the 'new' respondents. This is particularly the case with questionnaires administered to schoolchildren and with questions that are translated from other languages.

The first version of the pre-test questionnaire often contains considerably more questions than the final questionnaire. This can be upsetting for the respondents – especially if many questions are asked in an unstructured and open form so that the amount of time required to complete the questionnaire is considerable. If absolutely necessary, the questionnaire could be divided in two or three parts (of equal length and answering time) for the first tryout, so that each respondent answers only a fraction of the questions. For each form at least 50 respondents should be asked to participate. The information collected in this first pre-test should provide sufficient information to produce a second version of the questionnaire for final pre-testing.

This second version of the questionnaire will then be administered in one single form in order to further verify the functioning of the items and answer categories, as well as that of the questionnaire overall structure, layout, and accompanying instructions.

This process of pre-testing has a number of goals:

- **To reformulate or eliminate ambiguous or superfluous questions**

Take, as an example, a question soliciting information on how often the respondent speaks a language other than the official national language, at home. If a high percentage of students respond that they speak a language different than the national one, in an area where demographic statistics show that the presence of foreigners is low, it may be that the data are reflecting the use of dialects and not language. In this case it would be necessary to reformulate the question to reflect this observation.

- **To provide material to design the answer categories for open questions that need to be closed**

Take, for example, a question on the age of teachers. This could be asked during pre-testing in open format with a view to formulating categories for a closed question in the final version. The pretest could find that the ages of teachers in a particular country fall fairly evenly into categories that cover 5 years: 20-25, 26-30, 31-35, 36-40, 41-45, and 45 or more. For the purpose of the question it would be sufficient to have six categories of teacher age in a closed format for the final version. A similar exercise could be employed for open qualitative questions. For example, the following main reasons could be identified in a pre-test survey on teacher absenteeism: 'own health problems', 'family sickness', 'maternity leave', 'other family matters', and 'on training course'. This question could be closed using these five reasons, and adding a sixth category for 'other'.

- **To determine whether the questionnaire is balanced in its structure, and to discover whether instructions were properly followed**

The design of the layout should be guided by concern for the convenience and comprehension of the respondent, and in consideration of the subsequent work of the data processors, who will have to enter the data using computers. From the perspective of the data processors it is more practical to have numbers, rather than boxes, for the answer categories. Alternatively, the numbers can be placed next to each box, in such a way as to not confuse respondents and yet making it easy to enter and check the numerically coded answers.

The following steps cover the process of pre-testing and the main points to be examined during piloting:

Basic steps in pre-testing

1. Select a sample similar in socio-economic background and geographic location to the one that will be used in the main study. This sample will not be included in the final survey. Make sure you have a sufficient number of copies of the questionnaire for the pre-test.
2. Instruct interviewers or questionnaire administrators to note all respondents' remarks regarding instructions or question wording.
3. Administer the questionnaires.

4. Debrief the interviewers and check the results:
 - a. Is each item producing the kind of information needed?
 - b. What role is the item going to play in the proposed analysis?
 - c. Are the questions meaningful to the respondents?
 - d. Are respondents easily able to understand the items?
 - e. Can respondents use the response format for each item?
 - f. Did the interviewers feel that they were receiving valid information?
 - g. Was the question order logical and did the interview flow smoothly?
 - h. Did some parts of the questionnaire arouse suspicion?
 - i. Did other parts of the questionnaire seem repetitive or boring?
 - j. Were interviewers able to read the questions without difficulty?
 - k. Were respondents able to follow all instructions?
 - l. Was the questionnaire too long?

Reliability and validity

I. Validity

Validity concerns the degree to which a question measures what it was intended to measure (and not something else). Generally, there are three main types of validity related to the use of questionnaires: content, empirical, and concurrent validity.

- Content (or face) validity refers to whether a panel of judges or experts on the topic agree that the statements do relate to what they are supposed to measure. If agreement is obtained, then the instrument has content or face validity.
- Empirical (or predictive) validity is usually tested using a correlation coefficient which measures relationships between questionnaire responses and other related behavioural characteristics or outcomes. For example, a researcher could test the validity of an intelligence test by comparing scores on the test with the students' grade point average on a range of school subjects.
- Concurrent validity consists of measuring the degree to which a variable correlates with another measure, already validated, of the same variable. An example of concurrent validity is given by a study designed to test the validity of questionnaire items for use with 10-year olds in various countries (Wolf, 1993). The study compared the answers given by the children in the questionnaire with those given by the mothers who were asked the same questions either through an interview or a mail-questionnaire. The results showed high concurrence between the responses of children and those of their mothers to questions related to home conditions, such as questions on father's occupation, student age, and where the child studied in the home. However, considerable disagreement was observed on questions that were retrospective or prospective in nature, such as questions on how long the child had attended pre-school or how much more education the parents wished to have for their child. The conclusion was that it made sense to ask 10-year olds about their present life situation, however questions about the past or the future should be avoided as much as possible.

2. Reliability

Reliability concerns the consistency of a measure. That is, the tendency to obtain the same results if the measure was to be repeated by using the same subjects under the same conditions.

There are two general approaches to establishing the reliability of a questionnaire. The first is to ask the question again in a different part of the questionnaire in the same or slightly altered form, but in such a way as to yield the same information. This is a consistency check, but does not take into account variations in day-to-day variations. A second, and better approach, called Test-Retest, is to readminister a questionnaire to the same group of individuals several days later and to compare the results that were obtained.

This second approach was used in a small study of 9-year olds in Sweden (Munck, 1991). A correlation coefficient that described the strength of the relationship between responses at two times of administration was calculated. Three items from the study are presented below along with the Kappa coefficient.

Item	Kappa
Are you a boy or a girl?	0.98
Do you speak Swedish at home?	0.77
How often do you read for somebody at home?	0.41

Although the kappa for the question on gender seems high (0.98), for such a question one would expect the value to be 1. On a question like this, agreement can be increased through more careful supervision by the person who administered the questionnaire. The relatively low coefficients for the second two questions suggest that multiple data sources on many questions may be required for children at this age.

The analysis of trial data should also focus on producing frequency distributions of responses for each variable. Frequencies can be used to eliminate or modify questions that produce unusual response distributions. Items in which the rate of non-response or of 'don't know' responses exceeds 5 percent of the sample should be examined. Such high rates are usually indicative of ambiguities that are still inherent in items or inadequacies in the response categories. If the variable that the problematic item is measuring is central to the study, then further developmental work might be needed.

As changes are made, it is usually desirable to trial test the questionnaire again. It is not unusual that at least three trial-test studies are required before the questionnaire is adequate for its purpose.

The codebook

A codebook should be prepared in order to enter the data into a computer. The codebook is a computer-based structure file designed to guide data entry. It contains a field for every piece of information which is to be extracted from the questionnaire – starting from the identification code which allows each respondent in the sample to be uniquely identified.

In the codebook, each question/variable is identified by a name and is defined by a number of acceptable codes, or by a range of valid values for open-ended questions.

A coding scheme should be prepared for the closed and open ended-qualitative questions. The coding scheme is a set of numerical codes which represent all response categories, with additional codes to enter missing data (that is, questions left blank by the respondent) and not-applicable data (that is, questions that were not supposed to be answered by certain respondents).

The coding scheme for closed questions is easy to prepare. Codes are usually assigned sequentially to the set of response alternatives. They are often already printed on the questionnaire itself to identify each alternative, or next to the box to be ticked by the respondent, as in the following examples.

Are you a boy or a girl?

Boy 1

Girl 2

Are you a boy or a girl?

₁ Boy

₂ Girl

The coding scheme for the above question will be '1' for 'Boy', '2' for 'Girl', '8' for 'Not Applicable' and '9' for 'Missing'. It is customary to assign missing data to the highest possible value. That is '9' for one-digit questions, '99' for two-digit questions, etc. The values of '8', '88' etc. can be used to code 'Not Applicable' data.

The following table gives an example of a codebook format.

CODEBOOK FORMAT				
Variable name	Coding instructions	Column numbers	Missing	Not applicable
IDNUMBER	Respondent identification number Code Actual Number (001-528)	1-3		
Q1	Highest grade completed 1=1-8 2=9-11 3=12 4=13-15 5=16 6=17+	4	9	8
Q2	Gender of teacher 1=Male 2=Female	5	9	8
Q3	Hours worked per week current/last job Code in actual hours	6-7	99	88

Note that each variable is identified by its name, question content, the coding scheme employed, column numbers, missing and non-applicable values and any other special coding rules employed on a variable-by-variable basis (for example, Q3). From the information contained in the codebook, any researcher should be able to reconstruct the computer-stored data files from the completed questionnaires.

Open-ended quantitative questions are usually entered by recording the number supplied by the respondent. When preparing the codebook attention should be paid to the number of fields needed to enter such questions. Some programs as, for example, the DataEntryManager software (DEM), allow the researcher to specify a range of valid values for open-ended quantitative questions, so that an internal filter is provided to check for the entry of non-valid data. Suppose, for example, that an open question asks for the number of years of teacher education. When preparing the codebook it is possible to specify the maximum and minimum number of years required to become a teacher in a given education system. The programme will later block or signal the entry of data that fall outside this range.

The preparation of a coding scheme for qualitative unstructured items is a more laborious task. First, it is necessary to analyze the responses obtained in order to identify a set of categories that provide a meaningful classification system. Then each category is assigned a numerical value. The set of numerical values into which responses will be coded is the coding scheme. Codes for missing and not applicable data should be prepared following the same criteria as for closed questions.

EXERCISES

1. Explain the difference between funnel and inverted funnel sequences.
2. Explain the concepts of validity and reliability.
3. List three aims that a good covering letter should address.
4. State the main objectives of a trial-testing programme.

...

... EXERCISES

5. Specify a coding scheme for the following three questions.

Are you a boy or a girl?

(Tick one box only)

- Boy
- Girl

About how many books are there in your home?

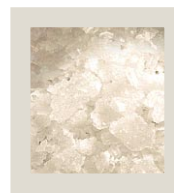
(Do not count newspaper or magazines. Please tick one box only)

- None
- 1-10
- 11-50
- 51-100
- 101-200
- More than 200

Does your school have any special programs or initiatives for reading outside the normal classroom activities?

(You may tick more than one)

- Extra-class lessons in reading
- Extra-individual tuition at school
- Special remedial reading courses
- Other



6

Further reading

Centre for Educational Research and Innovation. 1996. *Education at a Glance. OECD Indicators*. Paris: OECD.

Converse, J.M.; Presser, S. 1986. *Survey questions: handcrafting the standardized questionnaire*. California: Beverly Hills, Sage.

Foddy, W. 1993. *Constructing questions for interviews and questionnaires*. Cambridge: Cambridge University Press.

Guttman, L. 1950. A problem of attitude and opinion measurement. In: Stouffer S.A. (ed.), *Measurement and prediction*. New Jersey: Princeton University Press, Princeton

Johnstone, J.N. 1976. *Indicators of performance of education systems*. Paris: IIEP/UNESCO.

Johnstone, J.N. 1988. Educational indicators. In J.P. Keeves (ed.), *Educational research, methodology and measurement: an international handbook*. New York: Pergamon Press, pp. 451-456.

Likert, R. 1932. A technique for the measurement of attitudes. In: *Archives of Psychology*, 140:52.

Munck, I. 1991. *Plan for a measurement study within the Swedish IEA Reading Literacy Survey and some results for population A*. Stockholm: Institute of International Education, University of Stockholm.

OECD. 1995. Definitions, explanations and instructions. In: *UNESCO, OECD, EUROSTAT data collection instruments for the 1996 data collection on education statistics*. Paris: OECD.

Oppenheim, A.N. 1992. *Questionnaire design, interviewing and attitude measurement*. London: Pinter Publishers Limited.

Osgood, C.E.; Suci, G.; Tannenbaum, P. 1957. *The measurement of meaning*. Urbana, Illinois: University of Illinois Press.

Payne, S.L. 1951. *The art of asking questions*. Princeton University Press, Princeton, New Jersey.

Schleicher, A.; Siniscalco, M.T.; Postlethwaite, N.T. 1995. *The conditions of primary schools. A pilot study in the least developed countries*. A report to UNESCO and UNICEF.

Sheatsley, P.B. 1983. Questionnaire construction and item writing. In: P.H. Rossi; J.D. Wright; A.B. Anderson, *Handbook of survey research*. New York: Academic Press.

Thurstone, L.L. 1928. Attitudes can be measured. In: *American Journal of Sociology*, 33, 529-554.

UNESCO. 1976. *International Standard Classification of Education*. Paris.

Wolf, R.M. 1993. Data quality and norms in international studies. In: *Measurement and Evaluation in Counselling and Development*, 26: 35-40.

Since 1992 UNESCO's International Institute for Educational Planning (IIEP) has been working with Ministries of Education in Southern and Eastern Africa in order to undertake integrated research and training activities that will expand opportunities for educational planners to gain the technical skills required for monitoring and evaluating the quality of basic education, and to generate information that can be used by decision-makers to plan and improve the quality of education. These activities have been conducted under the auspices of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ).

Fifteen Ministries of Education are members of the SACMEQ Consortium: Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe.

SACMEQ is officially registered as an Intergovernmental Organization and is governed by the SACMEQ Assembly of Ministers of Education.

In 2004 SACMEQ was awarded the prestigious Jan Amos Comenius Medal in recognition of its "outstanding achievements in the field of educational research, capacity building, and innovation".

These modules were prepared by IIEP staff and consultants to be used in training workshops presented for the National Research Coordinators who are responsible for SACMEQ's educational policy research programme. All modules may be found on two Internet Websites: <http://www.sacmeq.org> and <http://www.unesco.org/iiep>.
