QUESTIONNAIRE DEVELOPMENT

Prue Anderson
George Morgan

Contents

1. Introduction

This module describes the process of developing questionnaires for national assessments.

A questionnaire is a set of items designed to obtain information from a person. The kind of information can vary widely, and may include data on personal characteristics, work qualifications and practices, working conditions and resources, or background information about the person, and his or her attitudes, beliefs, or opinions on certain issues.

A national assessment seeks to obtain a reliable estimate of student achievement (measured in a specially designed test) and information about key variables associated with differences in achievement (measured in a questionnaire). Tests collect information about student performance, and questionnaires, when used in conjunction with achievement tests, collect data about variables that might be associated with or help explain differences in levels of student performance. For example, questionnaire data can suggest that schools with no libraries are associated with poor student performance or that schools where teachers regularly participate in professional development programs are associated with high student performance. These data suggest different ways educational resources might be usefully directed to improve student learning.

A good questionnaire collects data about variables that policy makers want accurate information about, variables that they can possibly affect and are willing to influence, and variables that research evidence indicates can impact on student achievement.

The information obtained in questionnaires is useful if policy makers can and do use it to guide policy decisions. The most common mistake of most questionnaires is to

collect too much information that is of little value or cannot be used. Policy makers are generally only interested in information about a few key variables. Moreover, even when there may be good scientific reasons for collecting some kinds of data, consideration of the political and social consequences of collecting the data might indicate that a national assessment is not the most appropriate mechanism for doing so.

Information can often be collected from sources other than questionnaires in countries that keep accurate and reliable records about characteristics of schools, teachers, and students. It is worthwhile finding out if government records are a useful source of information that may be cheaper and easier to access than administering questionnaires.

2. Designing Questionnaires

Questionnaire design should clearly describe the kinds of data that will be collected, how the data will be analysed and reported, and how findings might contribute to improving education.

The main steps in questionnaire design are:

– Deciding the purpose of a questionnaire and how the data will be used.

– Developing a blueprint that specifies respondents, focus areas, item types, and coding or scoring and administration protocol – completed by an interviewer or self-completed.

– Writing items, panelling to review and refine items, and designing the layout of the form to make it easy for respondents to use and efficient for data entry people to process the data.

– Specifying a data analysis plan for processing information collected and creating measurement variables and indicators for subsequent statistical analysis.

– Pre-testing or field testing questionnaires to establish the suitability of items and response categories.

– Analysing the pre-test questionnaire data, refining questionnaires, and producing the final questionnaires for administration in a national assessment.

The development of questionnaires should parallel the development of tests for the national assessment. Box 2.1 provides details on the steps in questionnaire development and the people involved.

Questionnaires and instructions for their administration should be prepared and pre-tested or field tested at the same time as the tests. This means that questionnaire blueprints should be developed at the same time as the test blueprints, and that questionnaires should be written and panelled at the same time as test items are being written and panelled.

Questionnaire Content

A questionnaire should collect information about key variables that might help to explain differences in the performance of students on an achievement test. However, a myriad of intertwining variables can conceivably affect students' performance. A questionnaire can focus on only a few.

Policy makers usually want to know about variables that are associated with important educational issues in their country such as the language of instruction, disparities in the distribution of educational resources, or attitudes to the education of girls.

It is also quite possible that policy makers will not know what variables to investigate. They may give a long list of variables that are drawn from personal experience and observation or that they think 'ought' to be in a questionnaire. This needs to be reduced to a few focused requirements that are likely to be of use in shaping the content of the questionnaire.

Box 2.1
Components of Questionnaire Development

Component		**People involved**
1. Purpose	Clarify the purpose and potential use of the questionnaire data.	Policy makers, key stakeholders, and test development manager.
2. Blue-print	Design the questionnaire blueprint to specify respondents, focus areas, item types, coding and administration protocol.	Test development manager, subject experts, data analyst, experienced item writers, experienced teachers, policy makers, and key stakeholders.
3. Items	Write questionnaire items.	Test development manager and item writers.
	Panel questionnaires to refine for clarity and usefulness.	Test development manager and item writers.
	Review questionnaires.	Test development manager, policy makers and key stakeholders.
4. Data analysis plan	Specify the plan for processing information, for creating measurement variables and indicators, and for types of analysis.	Data analyst and test development manager.
5. Pre-test	Design, produce, and proofread questionnaires for pre-testing.	Test development manager, item writers, design and layout professionals, and proofreaders.
	Write administration instructions for pre-testing of questionnaires and train administrators.	Test development manager and item writers.
	Pre-test questionnaires at the same time as tests are being pre-tested.	Test development manager, logistics manager, and test administrators.
6. Final question-naire	Analyse pre-test questionnaire data.	Test development manager, data analyst.
	Refine questionnaire and administration instructions on basis of pre-test data and feedback from pre-test administrator.	Test development manager, item writers, data analyst.
	Produce questionnaire for final test.	Test development manager, item writers, design and layout professional, proofreaders.

Policy makers may not appreciate that analysing and reporting on questionnaire data are expensive and require technical expertise. Resources are invariably limited, especially when it comes to questionnaires, so questionnaires need to be short, concise, and highly relevant. Data collected in a questionnaire must also be technically acceptable if they are to be used to explain student performance. Basing the length of the questionnaire directly on models used by other national assessments is not recommended, but these models may be useful as a rough guide. Each country has its own needs and these must determine the nature of the questionnaire that is appropriate. Moreover, most questionnaires administered in national assessments collect far too much data, which are rarely analysed properly because of limited resources. Sometimes the data are not even processed.

The test development manager, or person responsible for the production of the questionnaire, may need to give policy makers some guidance about key variables that are likely to provide useful information. To do this, he/she may need to do some research and present policy makers with some relevant examples to help them refine their list of variables. Policy makers need to consider how they might use the information they will collect. Identifying this will help to further refine the list of variables to be addressed.

The main components of questionnaire development, together with the people involved in each component are set out in Box 2.1.

Since questionnaires will be designed to tap into things that respondents are likely to know about, the topics included are likely to vary for students, parents, teachers, and head teachers. The following are suitable topics for questionnaires for each of these groups of respondents:

Student Questionnaires

– gender, age, language background (usually collected on the front of the test booklet);

– educational background, such as years at school, periods away from school;

– opportunities to attend school;

– expectations of success, personal or family attitudes about the value of school;

– perceptions of classroom environments such as sense of safety, friendliness of other students, or support from teachers.

Parent Questionnaires

– nationality, gender, language background;

– home environment, such as access to books, desks, lights;

– family background, such as education of parents, language spoken at home;

– attitudes to education, such as commitment to sending children to school, perception of the value and relevance of education, or perceptions of the quality of education;

– study resources provided at home for children;

– affordability and accessibility of education for children;

– expectations of educational achievement for children;

– involvement with schools such as participation in the classroom or on committees;

– the nature of school reports about children's progress, and their value;
– financial support for school in the form of payment for textbooks and fees.

Teacher Questionnaires

– gender, age;
– first language;
– nationality;
– teaching conditions, such as class size, access to resources, percentage of students who have textbooks, access to replacement teachers when sick, and assistance with difficult students;
– educational experience, teacher qualifications, number of years in this school;
– professional engagement with learning such as access to and interest in professional development, interest in teaching, time spent preparing classes;
– teaching methodology, such as language of instruction, use of assessment, style of teaching;
– satisfaction with working conditions such as tenure, rates of pay, level of supervision;
– relationship with school community such as interactions with parents, involvement in school committees, participation in local community events;
– distance from teacher's home to school.

Head Teacher Questionnaires

– age;
– educational experience and qualifications;
– school environment, such as quality of buildings and facilities and availability of resources;
– school records such as fluctuations in student numbers, the extent of student or teacher absenteeism, the frequency of students moving schools;
– professional engagement with school leadership such as access to and interest in professional development, interest in education;
– leadership style, use of time;
– satisfaction with working conditions such as tenure, rates of pay, level of supervision;
– relationship with school community such as interactions with parents, participation in local community events.

Questinnaire Blueprint

A blueprint is required to guide the development of a questionnaire. It describes the content of the questionnaire; identifies the respondents; lists key variables to be addressed; specifies the format of items, kinds of response categories, and the administration protocol.

Box 2.2 provides an example of the blueprint of a questionnaire used in Papua New Guinea (PNG) to collect information about students' attitudes and values to school and

their local community. Recent reforms in education and new curriculum materials being introduced to schools had emphasized teaching students to value their local community and acquire skills that would assist them in constructively contributing to village life as adults. Papua New Guinea's policy makers wanted to collect information about students' expectations and perceptions of school and their community. The questionnaire was administered to all the students who sat the national assessment tests.

Box 2.2
Attitudes and Values Questionnaire Blueprint

Focus areas Attitudes to school Beliefs about life in PNG Perceptions of local community	
Number of questions 10 15 15	
Respondents Grade 3 Grade 5 Grade 8 Grade 5 Grade 8 Grade 5 Grade 8	
Response categories yes or no yes or no yes or no	
Part II Attitudes to school Beliefs about life in PNG Perceptions of local community	
beliefs about personal achievement, intended length of schooling, personal future plans attitudes to education: teaching in vernacular, compulsory eduction, role of school, education of girls, roles of women perceived level of co-operation in the local community: support for school; local involvement in community events; and sharing of resources	
perceptions of helpfulness of teachers, friendliness of students, bullying, willingness to make friends from outside village attitudes to community, personal intention to stay in local community or reasons for going perceived attitude of local community to girls and women	
attitudes to personal hygiene habits attitudes to conflict resolution and fighting perceived level of constructive employment in local community, and use of peaceful means to resolve problems	
attitudes to alcohol and drugs perception of problems caused by drug and alcohol use in local community	

Questionnaire Items

In deciding on the number of items in a questionnaire, consideration has to be given to the amount of time available to answer questions, the resources available for analysis, and the complexity of the analysis required. It is preferable to have a short limited questionnaire that is properly analysed and provides useful information than a long comprehensive questionnaire that is never processed.

The number of items needed to measure a specific variable depends on the nature of the variable. Some variables such as gender or age can be measured directly. Other variables such as socioeconomic status tend to be constructed from a number of variables, such as level of parental education, job status, location of home, and ownership of property. A raw variable is the data derived from a direct measure. An aggregated variable combines data from two or more items to represent a construct. Policy makers generally find results of analysis based on raw variables easier to interpret than those based on aggregate variables.

Decisions about whether a raw variable is a sufficient measure or whether an aggregated variable is required to support a construct should be based on good research practices and conditions in the country. Both national and international surveys have used aggregated variables. In Vietnam, for example, a variable "school quality" based on four variables was used (see Box 2.3).

Box 2 3
Weights Assigned to Variables Constituting School Quality in Vietnam Study

	Weight
Total school resources	.81
Total classroom resources	.76
% of pupils in full-day school	.70
Amount of parental contribution	.63

Country specific issues are relevant in deciding how many items are needed to measure a variable. For example, in a country where conditions of teacher education are fairly uniform, and all teachers have at least two or three years of tertiary education in recognized institutions, a single raw variable measuring years of tertiary education may be sufficient. However, in a country where conditions of teacher education vary widely, the quality of teaching institutions is extremely uneven, and many teachers may have been given on-the-job training, a number of raw variables may need to be aggregated to represent a construct of teacher education that reflects the situation in that country. Similarly, in a wealthy country, study resources at home may be measured by a single raw variable regarding access to the internet, but, in a poor country, study resources at home may be better represented as an aggregate of raw variables including access to a desk, a chair, a lamp, pencils, paper, and text books.

Decisions about whether to use a single raw variable or an aggregate variable to obtain a measure also depend on beliefs about the significance of the possible raw variables. For example, in measuring teaching experience, if its quality varies extensively depending on where teachers are employed, and there is a belief that this might also affect student performance, then information about where the teacher has worked should be collected as well as information about the length of time spent teaching. If there is a belief that the number of years of teaching experience might affect students' performance, regardless of where this experience was gained, then a single raw variable is probably sufficient.

Item Format

Forced-choice items are a great deal easier, faster, and cheaper to process than open-ended items. As they provide a limited number of categories from which to select a response, data processing is simply a matter of entering the respondent's selection, which can be immediately entered into a computer for analysis. Open-ended items, on the other hand, give respondents a space in which to write their answer, and responses have to be processed by hand before they can be entered into a computer for analysis, a complex and time-consuming process. Some national assessments have included open-ended questions that were never analysed.

Questionnaire data are often summarized for reporting. For example, responses to a question about the time students take to travel to school each day may be categorized into a few large groups, such as, less than one hour, between one and two hours, and more than two hours. In an open-ended version of this item, some students will give times in minutes and others in hours, others may write 'a long time' or 'quickly' and others will give an illegible answer. The range of responses will be large. Someone has to try to categorize all these responses, including making decisions about how to classify responses such as 'a long time'.

It is better to use forced-choice items if it is possible to make a good guess about the likely range and differences in the categories of most respondents' answers. If there is some uncertainty about this, then more finely differentiated categories than are required for reporting purposes may be used. Once data are entered, decisions can be made about which categories give little information and can be combined or dropped (e.g., if no one selected them).

It may be feasible to use open-ended items if the questionnaire is being administered to a small sample and resources are available to classify the responses by hand. It is sometimes useful to pre-test or field test items as open-ended with the intention of using the information obtained in the try-out to generate categories for a forced-choice version of the item in the final administration.

Language of the Questionnaire

The language used in the questionnaire should be a language that respondents are most likely to be able to read and write fluently. However, this needs to be

balanced with economies of scale. Generally, questionnaires are administered in the same language(s) as the test material.

Respondents

Given that some background information about students is always collected on the front page of test booklets, the selection of questionnaire respondents depends on what policy makers want to know and how feasible it is to obtain this information reliably and efficiently.

Some problems that may be associated with respondents are:

– students may be too young to fill in a questionnaire reliably or accurately;

– lack of resources may limit the administration of questionnaires to a small group, such as teachers or head teachers, rather than thousands of students;

– many parents may be illiterate or unreliable in returning questionnaires;

– teachers and head teachers may not be motivated to fill in a long questionnaire, or may feel too threatened to answer questions honestly.

Whatever decision is made about respondents, the sample selected for a questionnaire should be representative of the population. If the questionnaire is being administered to students, the same sample that has been drawn for the test should respond to the questionnaire. Sampling experts should be consulted about required sample sizes for administration to teachers, head teachers, and parents.

Questionnaires Administration

Questionnaires are typically responded to in writing or are administered in an interview. The latter requires a trained interviewer to ask the questions and to write down interviewee responses (perhaps applying codes provided in the questionnaire form).

In large-scale assessments, most questionnaires are written and administered in groups to minimize cost. Questionnaires should contain instructions on how questions should be answered. The instructions might include reasons for collecting the information.

Collection of questionnaire data may be done under supervision of a field worker, who collects the forms after they have been completed by respondents.

Data Analysis Plan

A data analysis plan specifies what kind of information will be provided by each item in a questionnaire, and how the information will be used in analysis. The provision of expert statistical assistance in designing the plan will increase confidence that data can be analysed meaningfully and that the findings will be rigorous and defensible.

The plan should show:

The measurement characteristics of the variables. Nominal or categorical variables have no intrinsic order to the categories (e.g., the categories of gender).

Ordinal variables have a clear ordering of categories such as number of years of teaching experience. The way the questionnaire data can be analysed depends on the measurement characteristics of the variables. See Appendix A for definitions of the measurement characteristics of variables.

– How data from a number of variables will be aggregated to produce a new variable, and how the new variable will be used. For example, an index of 'poverty' might be constructed from variables such as household income, location of home, number of children, and parents' education level. How the variables will be aggregated to represent 'poverty' would be considered in the design plan.

3. Writing questionnaires

It should be clear from the way an item is constructed what information is required. Furthermore, it should be within the competence of respondents to provide the information. Thus, one would not ask young students to recall how many days they have been absent during the school year as they would be unlikely to be able to do this reliably. At best, they might remember how many days they were away last week.

The wording in items should be as simple and clear as possible. Vocabulary should be familiar and sentences short and direct. It is important that all the respondents should be able to read the questionnaire.

The first part of a questionnaire item can be a question, an incomplete sentence, or a statement that respondents evaluate.

The style in which the respondent is addressed should be consistent. One or other of the following may be used:

– refer to 'you'; for example, 'How old are you?'.

– refer to 'I'; for example, 'I come to school by …'

Questions

Questions should be clear and unambiguous. The following question is ambiguous.

How long have you been a teacher?

This question confuses the time that has elapsed since training with the time engaged in teaching. Anyone who has left teaching and returned, such as women who took time out to raise their families will be unsure how to answer the question. It is not clear if this is a measure of teacher experience or of time elapsed since training. There are at least two questions here:

When did you complete your teacher training?

How many years of teaching experience do you have?

The second question is still problematic, as the way years of teaching experience might be measured is not clear. For example, should five years experience, working part-time, one day a week, be counted as five years or as the equivalent

of one year? If almost all teaching positions in the country are full-time, then this is not an issue, but if many are part-time it is. The question might read:

How many years of full-time (or equivalent full-time) teaching experience do you have?

It may be necessary to define what is meant by 'equivalent'.

Statements

Items that open with a statement usually require respondents to make some kind of evaluation of the statement, such as 'strongly agree', 'agree', 'disagree', or 'strongly disagree'.

It is preferable to avoid negative statements as they can be confusing as shown in the following item.

	strongly agree	agree	disagree	strongly disagree
I do not like school.	A	B	C	D

Students who do like school should select the categories of either disagree or strongly disagree to express their liking for school. Young children often find double negatives difficult.

It is preferable to keep statements as neutral as possible. It is better to have a statement that says, 'I like school', rather than one that says, 'I love school'. Students can express greater fondness for school by selecting strongly agree for their response.

Statements should focus on one issue. Thus statements such as 'I work hard and do well in my school work' should be avoided. Students who do well at school without working hard will not know which response to select. Students who work hard may strongly agree with this statement, although they may not do well in their school work. The statement is better expressed as two statements: 'I work hard at school'; 'I do well in my school work.'

Response Categories

Good response categories have the same meaning for all respondents. The following response categories are likely to have different meanings for different people.

How many text books do you have in your grade?

A. none
B. a few
C. some
D. many

The response categories for the item should be quantified so that the meaning is clear.

How many text books do you have in your grade?
A. none
B. 1 to 4
C. 5 to 20
D. more than 20

Sometimes response categories may have a different meaning for different respondents, but this is part of the information sought, as the following item shows:

How good is your school library?
A. no library
B. poor
C. adequate
D. good
E. excellent

If the item is about the respondent's level of satisfaction with the school library, regardless of any objective measure of its quality, then this is a good item. If the item is combined with items that quantify, for example, approximately how many shelves of books or computer facilities are in the library, then the respondent's perception can be compared with more objective measures of the extent of the library facilities.

Response categories also need to take into account the level of accuracy of answers respondents are likely to be able to give. Respondents are unlikely to know the number of books in a library with any kind of accuracy, unless it is very small.

Response categories need to cover all possible responses. If there are a few major categories and many minor ones, it is preferable to list the major ones and include an option of 'other'. Pretesting helps identify the major categories.

Response categories should not overlap or leave gaps. Both mistakes are shown in the question

How long have you been teaching at this school?
A. less than 5 years
B. less than 10 years
C. more than 10 years

Teachers with less than 5 years teaching experience do not know if they should select the first or the second option. Teachers with 10 years teaching experience do not have an option to select.

Questionnaire items that open with a statement that respondents are asked to evaluate may have many different response categories. It is important that the categories do not overlap. Some examples of response categories that have been used in the teacher questionnaire for the Third International Mathematics and Science Study (TIMSS) are:

– yes, no
– strongly agree, agree, disagree, strongly disagree
– almost every day, once or twice a week, one or twice a month, never or hardly ever

– not at all, a little, quite a lot, a great deal
– not important, somewhat important, very important

Managing Sensitive Issues

Some issues are sensitive, such as whether teachers have a second job to supplement their salary. If most respondents are unlikely to answer a question honestly, it should be left out. Policy makers may be very interested in this information, but there is little point in collecting unreliable data. Sometimes it may be possible to collect related information that is not as sensitive.

Including questions about sensitive issues may offend respondents who may refuse to answer the rest of the items or to return the questionnaire. If there is any concern about the sensitivity of issues, it is better to leave them out.

Questionnaire Layout

There are two critical considerations in the layout and design of questionnaires:
– ease of use for the respondent; and
– ease of use for data processing.

Questionnaires are easy to use when they have the following characteristics:
– simple, consistent way of answering questions;
– uncluttered presentation;
– separate questions are easily identified;
– response categories are clearly associated with each question;
– consistent use of headings, fonts, layout;
– response categories are coded for data entry.

Response categories can be set out in many different ways. They may be in a vertical column or in a horizontal row. Respondents may circle an alphabet letter or number or tick a box to indicate their selection. It is preferable to be consistent in the style of response.

An example is provided in Box 3.1 of an item in which responses are not clearly identified with response categories.

Box 3.1

The following example shows a poor alignment of boxes and response categories.

How long does it take you to get to school most days?

Less than 15 minutes	15 minutes	30 minutes
45 minutes	1 hour	more than 1 hour

The boxes are placed between the response categories instead of being clearly aligned with just one category.

Reviewing Questionnaires

Writing questionnaires is much more difficult than it looks. It is essential that all items are carefully scrutinized and revised to ensure they are clear and unambiguous. Asking a panel to review the questionnaire will help in this task. Members of the panel should include item writers; someone familiar with the characteristics of the respondent population; and someone who is able to ensure the items are culturally appropriate.

It is useful if panel members attempt to complete the questionnaire as though they were respondents. This will help to identify where categories of response might be unclear, overlap, or fail to include some kinds of response.

Panel members should critique the items, especially for clarity of wording and the suitability of response categories. They should ensure that wording is as simple and clear as possible, that the style of items is consistent, and that items are presented in a logical order with appropriate instructions.

Panel members should also check that items match the questionnaire blueprint and ensure that each item provides the required information. They also need to check that there are an appropriate number of items to measure each of the variables with sufficient precision.

Once the questionnaire has been refined, policy makers should be given the opportunity to review it. Policy makers need to approve the items, especially if they touch on politically sensitive issues. They also need to check that the items will provide useful information.

After questionnaires have been refined they should be pre-tested or field tested along with the test materials.

Pre-testing provides the opportunity to improve the quality of items and reduce the time and cost of processing data from the final questionnaire. Items that do not work (e.g., where respondents are confused) can be dropped, and categories of response can be expanded or contracted.

Following administration of the questionnaire, the administrator should collect feedback from respondents (students or teachers) about items that are unclear or do not contain appropriate information. They should check that no items are considered offensive because they touch on sensitive issues.

More formal statistical analysis of responses may indicate that response categories need to be more finely differentiated. For example, if most students select a particular response category for an item, it is better to split the category into more finely differentiated categories to obtain more precise information.

If the range of possible responses to an item is potentially very large and hard to anticipate, the item should be left open-ended in the pre-test. Responses can then be classified and used to generate categories for a forced-choice item in the final questionnaire.

4. Coding questionnaire responses

Response categories must be coded for data entry. Coding may be alphabetical, or numerical.

Alphabet codes usually require respondents to circle the letter for their response. This may not be suitable for younger students. Ticking boxes or shading circles can be an easier way for people with limited literacy skills to show their answers. If the items use this kind of layout, they should be coded numerically.

If numerical coding is used, the first response category is usually coded 1, the second category 2, and so on. It makes data entry efficient if the codes are printed on the questionnaire. This can be done in a small greyscale font as shown in Box 4.1. In the example, the response categories item are numbered under the boxes: walking is category 1, public transport is category 2, and so on. The student ticks the box that applies to him/her. The data entry person enters the number of the box that the student selected.

Box 4.1

Today I cam e to school by

walking ☐ 1 public transport ☐ 2 private transport ☐ 3 riding an animal ☐ 4

If respondents are given the opportunity to select more than one response category for an item, each category should be treated as a separate item for data entry and data processing. This way it is possible to keep track of which categories each respondent selected. The item in Box 4.2 is presented to the respondent as one question with multiple

Box 4.2

☐ u were away last week check one or more boxes to show your reason.
☐ I was sick.
☐ I had to help my parents.
☐ The weather was bad.
☐ I did not have food.
☐ My family had problems.
☐ I did not have a clean uniform or proper clothes to wear.
☐ It was not safe (peace and order problems).
1
Other _____
1

possible responses. However, it is treated as eight separate items in data entry. Responses to the first category (absent through sickness) are recorded as either 1 or missing, responses to the second are recorded as either 1 or missing, responses

to the third category (helping parents) are similarly recorded, and so on for each of the eight categories.

Preparing Questionnaires for Data Entry

Questionnaire data can be scanned with special equipment or entered manually. The design and layout of the questionnaire has to be customized if machine-scanning is to be used.

If data entry is being done manually, information can be entered directly from the questionnaire if response categories have been coded. However, it may be difficult for data entry people to maintain a high level of accuracy, especially if they are unfamiliar with this kind of work. Data entry accuracy is also likely to be compromised if the layout of the items varies extensively or if some items have a large number of response categories.

Data entry will be facilitated if raters write the code for the selected category in the left-hand margin, next to each item number. Data entry then becomes a simple matter of entering the codes written in the margin. Adding lightly shaded boxes in the margin for raters to write the codes makes the process more efficient.

Coding Missing or Ambiguous Responses

Sometimes respondents do not answer items or answer them ambiguously, such as selecting more than one response category when categories are mutually exclusive.

Collecting information about missing responses tells you whether there are some items respondents consistently failed to answer. For example, the questionnaire may be too long, so that items at the end are not answered, or an item may be too close to other items and easily overlooked. Collecting information about ambiguous responses will also tell you if an item is possibly unclear to many respondents or if the respondents do not understand how to complete the questionnaire.

The data entry person needs to know how to code missing or ambiguous responses. Codes used for missing or ambiguous responses should not be confused with the codes used for categories of response.

A letter of the alphabet may be used to denote missing or no attempt, such as X. The code for ambiguous responses could be a second letter such as Y. Multiple-choice items use codes of 9 for missing and 8 for selecting two or more options. These codes are usually not used for questionnaires as it is quite likely that some questionnaire items will have 8 or 9 response categories.

5. Linking questionnaires to test data

How questionnaire data are linked to test data will be guided by the needs of analysis and reporting. It is essential that all links are clearly and unambiguously

established prior to data collection. Any linking error discovered after data have been collected may be difficult or impossible to fix. This could also result in having to abandon some planned analyses.

Student Questionnaires. The easiest way to link student questionnaires to test data is to print the tests and the questionnaires in one booklet. The student records his/her name on the booklet and the test administrator ensures the student works in his/her own booklet for each of the test sessions.

If the tests and questionnaires are separate documents, one method of linking data is to overprint or label both tests and questionnaires with their students' names. Names will be taken from the school roll and should be identical for each label. Again, the test administrator needs to ensure that students work on the test and the questionnaire with their name on it.

If it is not possible to pre-label separate booklets and questionnaires, the student questionnaires need to have sufficient identifying information to allow them to be linked to test data. It is preferable to allocate a numerical identity (ID) to students and ensure they use the same ID number on each booklet. The test administrator will have to oversee this. (The data analyst will also need the list of student names and ID numbers.) Matching names may be used as a backup where there are ID errors.

It is not desirable to rely on students' names to match forms. Unless names are absolutely identical on each form and are entered identically by the data processing person, with no spelling errors, the computer cannot match them. Matching will then have to be done manually, a time-consuming and expensive exercise. Some students will make matching by names additionally complicated by using different names, such as shortened forms, family names, or religious names on different forms; writing illegibly on one or more forms; or failing to write their name on one or more forms.

Parent Questionnaires. Parent questionnaires may need to be linked to student data. The linking will probably be through the students' names. The same problems apply as outlined above. Procedures should to be set up to try to ensure consistency.

Teacher and Head Teacher Questionnaires. Teacher and head teacher questionnaires are usually linked to the grade and the school only. If the students' grade is known, the teacher information can be used in the analysis of student data. The test administrator should check that teachers and head teachers have provided this information on the questionnaire.

When questionnaires are returned from schools, each school's questionnaires should be stored in a separate bundle. This way, even if the school information was not supplied on some questionnaires, it has not been lost.

GLOSSARY

aggregated variable		two or more direct measures (raw variables) that are combined for analysis to create an aggregated variable to represent, for example, 'poverty'
blueprint		specifications about the criteria that questionnaire items must meet, including the variables to be addressed, level of precision required, and any other criteria or constraints regarding questionnaire development
categorical variable		a variable with no intrinsic ordering to the categories such as gender (being male or female), also called a nominal variable
data analysis plan		procedure for processing and interpreting data from the questionnaire
field test		another name for a 'trial test' in which assessment and questionnaire materials are tried out prior to final testing, usually involving a small sample (e.g., students, teachers, parents), to establish the quality and suitability of items.
final form		the booklet(s) of test(s) or questionnaires that are administered in the main data collection exercise
forced-choice item		the respondent is asked to select one of the response categories provided
grey-scale		a light grey shade of print
hand-recording		the recording of students' responses to items by humans (not machines)
item		a single part of a test or questionnaire to which an individual score or code is attached; it may be a question, an unfinished sentence, an instruction, or a statement
item panel		a small group of between three to six people who critically review and refine all aspects of items to ensure they are of high quality
rater		a person who hand-scores or records students' responses to items
missing responses		items that the student has made no attempt to answer (i.e., blank answers)
nominal variable		a variable with no intrinsic ordering to the categories such as gender (being male or female), also called a categorical variable
open-ended item		the respondent is asked to generate a response to an item (response options are not supplied)
ordinal variable		a variable in which response categories can be rank-ordered, such as not at all, a little, quite a lot, a great deal; 3 however, equal intervals between categories cannot be assumed
pilot testing		another name for a 'trial test' in which assessment and questionnaire materials are tried out prior to final testing, usually involving a small sample (e.g., students, teachers, parents), to establish the quality and suitability of items

policy makers		government officials who shape educational policies
pre-testing		another name for a 'trial' testing that is conducted prior to final testing, with a small sample of students, to establish the quality and suitability of items
proofreading		a detailed review of every aspect of a text to ensure that it is clear, consistent, and free of errors
raw variable		a variable that is measured directly such as age, gender, or distance travelled to school
respondent		person to whom a questionnaire is administered
response categories		the options given in an item for respondents to select or to be rated on.
test		an assessment instrument consisting of a number of items that students respond to under standardized conditions; the items are designed to allow students to demonstrate their knowledge, skills, and understandings
trial testing		an activity in which assessment and questionnaire materials are tried out prior to final testing, usually involving a small sample (e.g., students, teachers, parents) to establish the quality and suitability of items
variable		characteristic or attribute of the test population, such as age, gender, work environment, attitudes, beliefs