CHAPTER 20

Pretesting Questionnaires for Children and Adolescents

Edith de Leeuw and Natacha Borgers
Utrecht University

Astrid Smits
Statistics Netherlands

20.1 INTRODUCTION

Society is becoming more concerned with children's issues and children's rights. In most of the Western world, it is now recognized that the voices of younger children and adolescents should be heard and there is a demand for research that focuses on children as actors in their own right. As a consequence, survey researchers are realizing that information on children's opinions, attitudes, and behavior should be collected directly from the children; proxy reporting is no longer considered good enough *if* children can be interviewed themselves. Survey methodologists now focus on methods for designing questionnaires especially for children and adolescents and on methods for interviewing them. Official government agencies acknowledge children and adolescents as respondents and have developed and implemented special surveys for them (Scott, 1997). Also, academic research institutes and health organizations realize the need for accurate data collected directly from children and adolescents on their perspectives, actions, and attitudes (Greig and Taylor, 1999). Market research firms now acknowledge children and adolescents as special respondents and have guidelines for interviewing them (e.g., Esomar, 1999). However, relatively little is known about children and adolescents as respondents, and pretesting for this age group is a neglected issue (Blair, 2000).

Children are not miniature adults. Their cognitive, communicative, and social skills are still developing as they grow older, and this affects their ability to answer survey questions (Borgers et al., 2000; Cynamon and Kulka, 2001; Zill, 2001). For surveys of adults, procedures to enhance response quality and the improvement of data collection methods are well documented (Biemer et al., 1991; Groves, 1989; Lyberg et al., 1997; Sudman et al., 1996). Still, even surveying adults is far from simple, and methodological studies have shown that adults may experience problems with certain questions, and that question characteristics affect the data quality in surveys (Krosnick and Fabrigar, 1997). Especially when questions are very complex and/or when information has to be retrieved from memory, adults have difficulty (Eisenhower et al., 1991; Tanur, 1992). Interviewing children and adolescents is both similar to and different from interviewing adults. With children as respondents, the same problems may be magnified, as a slight error (e.g., ambiguity) in the questionnaire may be more difficult to overcome or have a larger impact. Also, children may experience additional problems when responding to a question, and the questionnaire should be adapted to suit the cognitive, linguistic, and social competence of each age group. The usefulness of an answer to a question will depend on the age of the child and his or her verbal abilities, so pretesting questions for their suitability for specific age groups is highly advisable.

The age of 7 is a major turning point in the development of children. At this age, their language expands (Nelson, 1976), reading skills are acquired, and they start to distinguish different points of view (Gelman and Baillargeon, 1983; Selman, 1980). These are important prerequisites for the understanding of questions. With special care, children can be interviewed with structured questionnaires or complete self-reports from the age of 7 onward. At the age of 18, adolescents are generally treated as adults in surveys, as is reflected in definitions of adult populations for many surveys (e.g., ISSP).

In this chapter we discuss methods for pretesting questionnaires for respondents between 7 and 18 years old. We start with an integrative summary of empirical knowledge on the young as respondents. This section is organized around the major phases in child development and will serve as a conceptual framework for testing questionnaires for children and adolescents. We present guidelines for optimizing questionnaire testing methods for different age groups.

20.2 DEVELOPMENTAL INFLUENCES ON SURVEYING CHILDREN: REVIEW OF EMPIRICAL KNOWLEDGE

As children grow from infancy to adulthood, their thinking becomes more logical and their reasoning skills develop, memory and language develop, and social skills are acquired. Although there is considerable variation among children, depending on heredity, learning, experiences, and socioeconomic factors, consecutive stages can be discerned. The pioneer in child development research was Jean Piaget. Piaget's theory of developmental stages has provided the impetus for much psychological research and gives a useful framework for practical applications (Flavell et al., 1993). In the following sections, we discuss developmental

issues for surveying children based on Piaget's stages, but amended with modern insights derived from information processing and sociocultural perspectives (see also Gray, 2002). One should always keep in mind that the stages presented should not be seen as sharply distinct categories, but rather, as a moving scale: There are differences within age groups, and there may be overlap between groups.

20.2.1 Middle Childhood (7 to 12)

Piaget (1929) saw the age of 7 as a major cognitive turning point; around this age children make the important transition from *preoperational* to the more advanced *concrete operational* period. Starting at age 7, children are better at logical, systematic thought using multiple pieces of information. Language skills develop further and reading skills are acquired. Children begin to learn about classifications and temporal relations, but still have problems with logical forms, such as negations. They become much more capable of perceiving underlying reality, despite superficial appearance (Flavell, 1985), but still may be very literal in interpreting words.

Consistent with Piaget's early view that young children have problems with logical negations and abstract thought, Holaday and Turner-Henson (1989) found that children in middle childhood have difficulties with "vague" words because they tend to interpret words literally. For example, offering vague quantifiers in questions about the frequency of behavior produces difficulties for children because they need clear definitions, especially in early middle childhood (7 to 10). For this age group, simple yes/no questions about doing something are better understood. Negatively formulated questions make the intended meaning ambiguous for children (as they do for adults) and should always be avoided in children's questionnaires. Younger children in middle childhood have particular difficulty with negatively phrased questions, while older children and adults experience less difficulty (Benson and Hocevar, 1985; Borgers and Hox, 2001; De Leeuw and Otter, 1995). To understand what is required, a child should also grasp the intended meaning of a question. As a result of their literal interpretation, the distance between the intended meaning and the literal meaning of words can cause serious problems for children in middle childhood. This is even more pronounced when *depersonalized* or indirect questions are used (Scott, 1997). A clear illustration is the observation made by Scott et al. (1995) during pretesting that in reaction to questions using the term "people my age," some children tried to guess the age of the interviewer before answering!

Another important factor for participating in questionnaire research is memory. In middle childhood the variety and effectiveness of memory strategies increase. Many studies have shown clear increases with age of the amount of information that can be kept and manipulated in working (short-term) memory (Swanson, 1999). Around age 10 to 11, the memory capacity of children is at the same level as adults (Cole and Loftus, 1987) and the constructive processes used by children seem to function much like those of adults (Kail, 1990). When questions

are clear and concrete about the here and now, even young children (7 to 10) are able to give informative responses (Amato and Ochiltree, 1987). The still developing memory capacity also has consequences for the number and order of response alternatives. A limited number of response categories gives better results (Hershey and Hill, 1976). Holaday and Turner-Henson (1989) advise not more than three before the age of 10, but even with older children more than five is not advisable (Borgers et al., 2004). Scott et al. (1995) found good results using graphical representations (e.g., smiley faces) as response categories.

Retrospective questions pose extra problems for young children because of their still developing memory capacities. If the question is immediately recognizable and concerns salient and meaningful experiences (e.g., class outing, visit to pediatrician), even children in early middle childhood (7 to 10) can answer correctly, as their memory for salient issues is remarkable (Brainerd and Ornstein, 1991). However, several studies have shown that unreliable responses appear if these children are not involved or interested in the subject (Holaday and Turner-Henson, 1989; Vaillancourt, 1973). Younger children in particular are prone to construct scripts or event representations of familiar routines if they do not clearly recollect atypical events (Brainerd and Ornstein, 1991; Ceci and Bruck, 1993) or when more complex questions are asked (De Leeuw and Otter, 1995). Furthermore, there are developmental differences in reality monitoring, and under certain circumstances young children (early middle childhood) have more difficulty distinguishing between imagined events and those actually experienced (Ceci and Bruck, 1993; Johnson and Foley, 1984). This is corroborated by Saywitz (1987), who found that 8- and 9-year-olds tended to have less complete recall and more embellishments than did 11- to 12-year-olds.

Provided that extreme care is taken, diary-type methods can be used. The diary method is minimally demanding in terms of cognitive processes and memory, and uses the "here and now" type of question, which is especially appropriate for children (Amato and Ochiltree, 1987). Otter (1993) showed that using the diary method to measure 9-year-old children's leisure-time reading yielded good response quality, produced reliable and valid data responses, and was superior to self-administered paper-and-pencil questionnaires. Structured diaries were also used successfully to collect information about peer interactions of children in their final year at primary school (Ralph et al., 1997).

In addition to an age-related increase in working capacity, an age-related increase of processing speed has also been established. Kail (1993) found a steady decrease of reaction time and increase in processing speed with age; on six different tests, children of 8 and 9 took twice as long as adults, and children of 10 and 11 took 1.5 times as long. By age 17, they performed almost as fast as adults (Gray, 2002). Holaday and Turner-Henson (1989) therefore advise giving children more time to answer survey questions. One way to do this is to use longer introductions to a question. That has a positive effect on response quality, as shown by Borgers and Hox's (2002) finding that the number of words in introductions to questions was positively related to the reliability of children's responses.

Suggestibility has been a topic of much debate in the field of children's testimony in the past 20 years. For an overview, see Ceci and Bruck (1993) and the special issue of *Law and Contemporary Problems* (2002). There appear to be two aspects of suggestibility relevant for survey research (Bob Belli, personal communication): one is suggestibility resulting from cognitive factors, including potential alteration of memories for a past event, as discussed above. The other aspect is suggestibility that results from social and motivational factors, such as seeking to please the interviewer. According to Maccoby and Maccoby (1954), children as old as 8 years will assume that an adult knows everything already. In addition, they are afraid to say something wrong or foolish, especially in a situation that resembles school (Delfos, 2000). As a consequence, young children may react to the demand characteristics of the interview situation by responding in socially desirable ways (La Greca, 1990), or fall back on other response strategies, such as yea-saying (Maccoby and Maccoby, 1954). During middle childhood the structure of self-concept changes, and in late middle childhood (10 to 12) children start comparing themselves to others, and from approximately 10 years on, the effect of peers will be more present (Kohlberg and Puka, 1994). Furthermore, they become aware of the possibilities of putting on a facade and deceiving others intentionally (Selman, 1980). This is clearly illustrated by several methodological studies on children as respondents. Borgers and Hox (2001) reanalyzed questionnaire data from five studies and found that on sensitive questions the younger children had less item nonresponse than older children, while on nonsensitive issues this was reversed, indicating that older children prefer avoiding a socially undesirable answer. Van Hattum and De Leeuw (1999) found that a more private setting (CASI) resulted in fewer social desirable answers for children in late middle childhood.

20.2.2 Adolescence (12 to 18)

After the age of 11, children enter the stage of formal operations (Piaget, 1929). In this stage in early adolescence (roughly 12 to 16 years of age), cognitive functioning is well developed, including formal thinking, negations, and logic. There is a shift in emphasis from the real to the possible, from what is to what might be, and the young adolescent can manipulate ideas about hypothetical situations (Conger and Galambos, 1996).

At the beginning of adolescence memory capacity is fully developed and the constructive processes function much like those of adults. During adolescence memory processing increases rapidly, and by the age of 16 it approaches adult speed (Kail, 1993). Also, social skills are further developed. Selman (1980) calls this the stage of social and conventional system role taking (roughly 12 to 16), in which the young adolescent attempts to understand another person's perspective by comparing it to that of the social system in which he or she operates. Young adolescents in this age group are context sensitive and may have their own norms. After the age of 12, peers become more and more important, and numerous studies have shown that conformity to peers and peer pressure increases dramatically in early adolescence (Gray, 2002).

From 16 years onward, adolescents can be regarded as adults with respect to cognitive development and information processing. But resistance to peer pressure is very low and older adolescents have their own group norms and social norms. The social context of the survey (e.g., classroom, presence of siblings or friends, type and age of interviewer) can be important, especially in interaction with special topics (e.g., health, social networks). For example, in a drug survey among U.S. high school students, the more private data collection method worked best, resulting in more openness and increased reporting. Even the physical proximity (measured physical distance) of other students influenced the openness of answers (Beebe et al., 1998).

20.2.3   Summary

In surveying children, language ability is an important issue for the comprehension of questions. *Comprehension* is the first step in the question–answer process that has to be checked in pretesting questionnaires (cf. Tourangeau and Rasisnski, 1988). As reading and language skills are still developing in middle childhood (7 to 12), the understanding of words has to be checked very carefully for this group. Extra attention should be paid to complexity of wording, negations, and logical operators. As children can be very literal, depersonalized or indirect questions should be checked very carefully.

Memory and processing time is a second important issue. In middle childhood (7 to 12) both memory capacity and speed are still developing. Therefore, complexity of the question and number of response categories should be examined carefully in pretests. Retrospective questions may pose extra problems, and young children are prone to construct scripts of familiar routines if they do not clearly recollect events. In early adolescence (12 to 16) memory capacity is full grown, but memory speed is not. Even in this older age group, ample time for answering questions should be allowed.

In younger children, *suggestibility* is an important item. In early middle childhood (7 to 10), children have a tendency to please and are afraid of doing something wrong. This may result in more satisficing strategies and an inclination toward social desirability. In late middle childhood (10 to 12) children become less suggestible, but start to compare themselves with others. From the age of 12, peers become increasingly important, making adolescents increasingly sensitive to peer pressure and group norms. Sensitivity of topic and privacy of interview situation become important.

20.3   PRETESTING METHODS FOR QUESTIONNAIRES TO BE ADMINISTERED TO CHILDREN AND ADOLESCENTS

20.3.1   Setting the Stage: Survey Design Decisions

Designing and conducting quality surveys requires a careful decision process (e.g., Czaja and Blair, 1996; Lyberg et al., 1997). Designing surveys for children

and adolescents is no exception; however, with young respondents some design issues are of extreme importance and warrant extra attention. These include the question of proxy reporting, mode issues, question wording, and consent.

Self-Report versus Proxy   Before the age of 7, direct questionnaire research of children is not feasible, and one has to fall back on proxy interviews or on other forms of indirect data collection. Children younger than 7 do not appear to have sufficient meta memory skills to be questioned effectively and systematically (Memon et al., 1996) and experience severe problems in understanding more than very simple concepts (Riley et al., 2001). From the age of 7 on, children can be surveyed with structured questionnaires, and the older the child, the more reliable the answers (Borgers, 2003; Zill, 2001). Zill (2001) advises using an informed parent as informant for health-related issues until adolescence, and collecting information directly from a child on topics for which the child is the best informant. These include subjective phenomena, such as feelings, pain, but also questions on peer influence and peer behavior, and general questions in areas outside the scope of parents' knowledge. The latter is well illustrated by Blair (2000), who compared different protocols on children's food intake and checked these with validating information obtained through observation. Children aged 6 to 11 provided better information than their parents did. The main reason for the discrepancy was the faulty assumption by parents that the children had eaten the food taken with them to school (Blair, 2000). In general, the decision to rely on self or proxy reporting is made a priori on theoretical or practical grounds and differs from country to country and from topic to topic.[1] However, a pretest could provide useful data to guide this decision process. The study reported by Blair (2000) is a good example of this procedure.

Survey Mode   Data on very young children are usually collected through observational and assessment studies performed by specially trained interviewers and through interviews with caretakers (Borgers et al., 1999; Zill, 2001). From 4 to 6 years of age, children can be interviewed, but not easily. The interview resembles a qualitative open interview with a topic list, the form is play and talk, and much attention should be given to nonverbal communication and communication of the rules and expectations (Delfos, 2000). Interviewing such young children is a special skill, which is outside the general frame of survey research. However, special handbooks on this topic have been published for counselors, social workers, and law officers. Although these books often focus on very sensitive topics such as sexual abuse, they give guidelines that are extremely useful for

---

[1]Population definitions for general surveys differ and start at either 18 or 15 years: Nordic countries such as Finland and Sweden have 18 as the lower age limit in official statistics, while the United Kingdom and the Netherlands use 15. In labor force surveys in Europe and the United States, persons 15 years and over are eligible as respondents. The International Social Survey (ISSP) uses 18 years as the lower limit; the European Social Survey (ESS) uses 15. For special surveys children as young as 10 (e.g., Level of Living in Sweden) or 12 (e.g., Crime Victimization Survey and Survey of Program Dynamics, United States) are eligible in official statistics, but permission is needed.

any interview with very young children (see, e.g., Aldridge and Wood, 1998, and Wilson and Powell, 2001). Other forms of special data collection techniques with young children are playing assessment, drawings, story completion, and puzzle tasks (for a description, see Greig and Taylor, 1999). Because at such a young age the child usually is not able to give detailed information on general background characteristics and facts about family, health, and schooling, caretakers are interviewed as proxies.

Starting at the age of 7, structured questionnaires can be used either during a survey interview or through self-completion. Which particular mode is chosen depends on design constraints, such as research topic and budget, and on the literacy of the intended population. From the age of 7 to 8 years old, educational researchers start to use simple self-administered questionnaires in the classroom. When literacy is a problem, a combination of methods is often used, with an instructor reading the questions aloud and the pupils recording their responses on a self-administered form (Borgers et al., 2000). Also, in individual or household surveys, a combination of methods can be used when asking sensitive questions of young respondents. For instance, Scott (1997) used a combination of a Walkman with prerecorded questions on tape and an anonymous self-administered questionnaire. If the budget allows, computer-assisted self-administrative methods, such as CASI or Audio-CASI, have advantages both in school surveys and in household surveys of young respondents (Hallfors et al., 2000). Children and adolescents are good respondents in computer-assisted surveys, and even ordinary schoolchildren as young as 8 years can successfully complete electronic questionnaires and enjoy the process (Van Hattum and De Leeuw, 1999).

A pretest can provide useful information to guide the mode decision. A good example is the study of Helweg-Larsen and Larsen (2001, 2002), who observed both standard mainstream and special education students 15 to 16 years old while they completed a pilot version of a Danish health survey. They found that the special education students took longer and read at such a slow rate that they lost grasp of what had just been asked in the text. Based on these observations and subsequent focus groups, the researchers decided to use Audio-CASI technology for their main study.

Consent One of the strictest codes for doing research with human subjects is the *Declaration of Helsinki* of the World Medical Association, set out to provide moral, ethical, and legal principles to biomedical researchers. Recent amendments of this declaration now include the issue of children and informed consent (Greig and Taylor, 1999; University of Essex, 2002). It states that "... when a subject is a minor, permission from the responsible relative replaces that of the subject in accordance with national legislation. Whenever the minor child is in fact able to give a consent, the minor's consent must be obtained in addition to the consent of the minor's legal guardian" (World Medical Association Declaration of Helsinki, paras. I.9, 11). Esomar, the world association for research professionals in opinion polling and market research, states in its guidelines that first of all, a researcher should conform to any relevant definitions in any national code of conduct and/or

in national legislation, and second, that in the case of children under 14, explicit permission should be asked of a parent, guardian, or other person to whom the parent has conferred responsibility (Esomar, 1999). Legislation may vary from country to country regarding the age at which children can legally give their consent. For example, in the United Kingdom it is 16 (University of Essex, 2002), and as a consequence the British Market Research Society prescribes that consent of a parent or responsible adult must be obtained with children under 16 (Market Research Society, 2003). In Sweden, permission from parents is required until 18, even for social surveys such as the ESS (S. Svallfors, personal communication).

Still, permission of a parent or guardian is not enough. The declaration of Helsinki prescribes that the minor's consent must be given, too, if the minor is able to do this. Professional research organizations such as the Society for Research in Child Development also require that researchers inform the child about the study and obtain permission of the child in addition to the consent of the legal guardian (Goodwin, 2002). This implies that the information presented to the child should be given in clear language and at a level the child can understand. To verify this, a cognitive pretest of the wording and phrasing of the consent statement should take place for the relevant age groups. Research in this area is scarce, but an exception is the work of Abramowitz et al. (1995), who investigated the capacity of children in middle childhood (7 to 12) to give informed consent to participation in psychological research, using vignette descriptions followed by open interviews. Their main finding was that children could describe the purpose of the studies fairly accurately, and that the child's consent was not influenced by knowing whether parents had given their consent. However, many children had difficulties recounting the potential risks and benefits of the studies.

Wording of Questions and Response Choices When developing and evaluating questionnaires for children, a researcher should start by following the basic rules for general questionnaire construction and evaluation as outlined in handbooks such as Converse and Presser (1986), Dillman (1978, 2000a), Foddy (1996b), and Fowler (1995). These include good advice to use simple words, avoid ambiguity, ask one question at a time, and so on. But one has to do more. Methodological studies on adult populations have shown that adults sometimes experience problems with certain questions, and that question characteristics affect the data quality in surveys (cf. Krosnick and Fabrigar, 1997). Evidence for interaction effects between respondent characteristics and question characteristics has been found by Borgers and Hox (2001), De Leeuw and Otter (1995), Knäuper et al. (1997), and Schwarz et al. (1999). These studies show that the less cognitively sophisticated respondents are more sensitive to more difficult or cognitive demanding questions than the more cognitively sophisticated respondents, resulting in more item nonresponse and less reliable answers for respondents lower in cognitive ability.

With children as respondents, these problems are magnified. In addition, children experience specific problems when responding. Not only their cognitive, but

also their communicative and social skills are still developing, and this affects different stages of the question–answer process. Therefore, special care should be given to the construction of questionnaires for children and adolescents. Pretesting of the questionnaire is certainly necessary to examine the adequacy of question wording and response options for different age groups. This is still a new field, and few publications about procedures and results are available. Levine and Huberman (2002) describe how they effectively used cognitive interviewing (think-aloud with probing) with children aged 9 to 14 to test questions from the U.S. National Assessment of Educational Progress. Hess et al. (1998a) describe similar positive experiences with adolescents aged 12 to 17 when pretesting the youth part of the U.S. Survey of Program Dynamics.

Design Decisions and Pretesting   When surveying children, many design decisions may be guided or informed by using pretesting methods. A first step is consulting with experts in the field. The next step is evaluating the procedures and questionnaire using cognitive testing methods, using the intended respondents as informants. Cognitive pretests will enable the researcher to discover which wordings or questions are problematic for young respondents and why, thereby suggesting improvements in questionnaires for children. In the next section, well-known cognitive methods for pretesting with adults (e.g., Esposito and Rothgeb, 1997) are reviewed for usability with children. In addition, we discuss how these methods can be optimized for children.

20.3.2   Focus Groups

Different pretest methods have different strengths (Presser and Blair, 1994). The strength of focus groups is the interaction within the group; the participants stimulate each other to discuss topics and explain ideas (Morgan, 1997). As a consequence, a wide range of information can be gathered in a short time; however, this information is not always very detailed. Focus groups are useful for generating ideas and topics for questions, evaluating the data collection procedures planned, and evaluating the acceptability or sensitivity of certain topics, but for a detailed evaluation of the questions, in-depth interviews are more useful (Campanelli, 1997; Snijkers, 2002).

The usefulness of focus groups in the design phase of a survey is well illustrated by Scott et al. (1995), who conducted a series of six focus groups with children aged 11 to 15 in the United Kingdom. The decision to add a Young Person's Survey to the British Household Panel challenged the researchers to develop a way of interviewing children in their homes in privacy. Because of potential literacy problems, the researchers opted for prerecorded Walkman interviews with a paper self-completion response booklet. The goal of the focus groups was to help develop structured questions and to fine-tune the Walkman method. The focus groups took place in a neutral setting, the interviewer's home. Groups were separated by gender and by age groups (11 to 13 and 13 to 15) and lasted about two hours with a snack break at half time. Each focus group started with a general

open discussion on health and health-related issues. This served as a warm-up but also provided information on the typical language use and on sensitivity of topics. This was followed by trying out formats for semistructured questions thought suitable for these age groups (e.g., response card with a range of smiley faces). Question formats were presented and discussed in the group. In the last phase of the focus group, Walkmans were handed out together with a short self-completion booklet. According to the researchers, the focus group discussions were very productive for identifying appropriate wordings, question formats, and response options for the development of the Young Person's Questionnaire. The Walkman test showed that children did not experience any technical problems when using a Walkman and provided useful feedback on voice type (Scott et al., 1995; see also Scott, 1997). A subsequent test of the redesigned procedure during the pilot phase of the Young Person's Survey was reported to be very successful (Scott, 1997).

Using focus groups of young persons in the design phase of special surveys may provide useful information, and although it is still in the pioneering phase, its use is growing. Different approaches may be used for different purposes. For instance, Spruyt-Metz (1999) used focus groups to pretest a self-administered questionnaire on health and risk behavior among Dutch high school students aged 12 to 17. She was interested primarily in question interpretation and the meaning of important concepts and used open-interview questioning. Cannell et al. (1992) used focus groups of U.S. adolescents to test the acceptability of health-related sensitive topics (e.g., cigarette smoking). They presented subjects with potential questions, and stimulated group discussion by giving specific probes on the understanding of the question, how one would react, and on whether or not one would answer it, or answer it truthfully.

A rather unorthodox but fruitful application of focus group techniques was employed by Watson et al. (2001) in New Zealand, who used *post pilot focus groups* to evaluate the usability of Multimedia-CASI techniques. Following completion of a questionnaire, students aged 12 to 18 participated in structured focus groups. Each group consisted of six to 10 students of the same gender and took about 40 minutes. Open-ended questions were used to stimulate discussions about available time, use of headphones and computers, but also question the difficulty and emotional burden of the questions. The focus groups revealed two important themes. First, the students were very positive about the multimedia computer interface, especially the audio component. In the eyes of the respondents, the computer made everything easier. The second perceived advantage was privacy. Students appreciated the computer but also emphasized how important it was that nobody else could read the screen.

Focus Groups with Children and Adolescents   Compared to general adult focus groups, focus groups for children and adolescents appear to be more structured and more centered around specific tasks. Whether this is inherent for groups with children and young adolescents, or whether this is the result of the specific topics in the studies cited above, is unclear. The researchers do not describe in

detail if and how the focus groups were adapted to the younger respondents. However, general publications about interviewing children (e.g., Delfos, 2000; Wilson and Powell, 2001) emphasize the importance of a well-designed protocol for open interview situations and the extreme importance of explaining clearly what is expected of the child. This is also stressed by Morgan et al. (2002), who wrote one of the first methodological articles about focus groups with children.

Although in most countries children and adolescents are acquainted with group discussions from classroom settings, they will not know what a *focus* group is and what its rules are. Therefore, it should be made very clear to them what is expected and also that a focus group is not school or a test situation. Also, during the focus group itself, the participants sometimes need to be reminded of the rules. For instance, Morgan et al. (2002) wrote simple rules on a flip chart in the beginning and left them on display during the entire session. Examples of these rules were: Everyone gets a chance to speak, speak one at a time; you do not have to put up your hand to talk (this is not school). Of course, explaining the rules is important when conducting focus groups with adults, too. But young respondents are still developing the cognitive and social skills for meta-communication (see also Section 20.2) and compared to focus groups for adults, the moderator has to pay more attention to meta-communication.

In general, many issues and good practices for focus groups with adults are common to conducting focus groups with children and with adults; it is a question of translating these good practices to the needs of younger age groups (Morgan et al., 2002). Through the setting and the explicit verbal and nonverbal behavior of the moderator, the researcher has to create a different interaction-stimulating environment for each age group. In the following paragraphs we discuss optimal focus group settings for different age categories, emphasizing the special needs of each group. We will not discuss the general rules for conducting good (adult) focus groups; for a thorough introduction we refer to Morgan (1997) and Stewart and Shamdasani (1990); for a quick overview, see Cheng et al. (1997) and American Statistical Association (1997). However, as certain topics, such as group size and homogeneity, are recurrent methodological issues in focus group setups for developing questionnaires (Bishoping and Dykema, 1999), we will comment explicitly on these topics.

Group Size Young children need more attention than older children, and as a general rule, the younger the participants, the smaller the group should be. For children in early middle childhood (ages 7 to 10) a group size of about five is optimal. To increase motivation and keep the attention of these young children, one moderator should constantly attend to motivating the children and keeping the conversation going. A second moderator will be necessary for general practical assistance in running a group of young children (see also Greig and Taylor, 1999; Morgan et al., 2002). More grown-ups in the room will disrupt the balance of power in the group, and it is advisable to have note takers in a separate room and to videotape the entire session for nonverbal cues and interactions (Annon, 1994).

In late middle childhood and early adolescence (ages 10 to 16) group sizes may range from 5 to 8 (Scott et al., 1995). A second moderator will no longer be needed for practical child-care issues and may be replaced by a note taker or observer. In late adolescence (16 to 18), group size may increase to 8 to 10 participants, only slightly less than in adult groups (cf. Bishoping and Dykema, 1999).

Group Homogeneity Group composition is an important consideration in focus groups. Homogeneity in age with small age bands (e.g., ages 7 and 8, 9 and 10) is recommended (Morgan et al., 2002). In early adolescence this is crucial, as the eldest will in general look down on the youngest, who has just left primary school. Therefore, age homogeneity should be strictly enforced, with the 12- and 13-year-olds separated from the older children (cf. Scott et al., 1995).

Whether or not groups should be homogeneous with respect to gender is age-dependent. Before the age of 10, gender homogeneity is not necessary, but in late middle childhood and early adolescence it is advisable (Greig and Taylor, 1999; Scott et al., 1995). In late adolescence much depends on the topic of the study and on culture. For instance, Spruyt-Metz (1999) varied the composition of focus groups of Dutch adolescents. She used both all-girl and all-boy groups, but also added mixed-gender groups to stimulate discussion. According to Spruyt-Metz (1999), having opposite sex members in the group may reduce acting-out behavior and make the group more task-oriented. Only for the adolescents of Turkish and Moroccan origin were the groups gender homogeneous, because of cultural taboos on discussing many of the topics in the protocol with members of the opposite sex. The findings of Bishoping and Dykema (1999) are helpful in deciding on gender homogeneity for focus groups with late adolescents and young adults (16+). They review extensively the importance of sociopsychological factors in focus groups for adults and conclude that sex segregation has negative effects, especially on disclosure of emotions and personal information, for men, while for women all-female groups enhance their input.

Scott et al. (1995) note that their focus groups were homogeneous in terms of socioeconomic status. But this could be country specific and dependent on the school system and whether or not there are large status differences between schools as there are in the United Kingdom.

For all age groups it is advised to avoid having close friends, or even classmates in one group, as this may have affect group dynamics. It may stimulate concentration lapses in younger children (Morgan et al., 2002) and inhibit open interactions (Scott et al., 1995). Especially in adolescence, when peer pressure is heavy (Gray, 2002), one should avoid selecting children from the same peer groups or school classes and preferably mix children from two or more schools.

Session Duration The younger the child, the shorter the attention span. In early middle childhood (7 to 10) the attention span is still limited, and this has consequences for the scheduling of a session. One should have short periods of discussion (around 20 minutes) alternated with play activities (Delfos,

2000). Morgan et al. (2002) used two 20-minute sessions separated by a short refreshment break; they also advise keeping the (tape) recorder running during the breaks to catch relevant remarks.

According to Delfos (2000), children 10 to 12 can have longer periods of discussion (30 to 45 minutes), alternated with refreshment breaks. Scott et al. (1995), who studied children aged 11 to 16, used focus groups that lasted approximately two hours. Although the attention span of these older children is longer, the moderator should carefully monitor the process and stimulate participation. Group discussion can be alternated with other activities, such as making lists of important points (Morgan et al., 2002), showing pictures, or having children handle survey material (Scott et al., 1995). Adolescents can handle discussion periods of one hour, after which a refreshment break is definitely needed. This is as long as most adult focus groups. Still, one has to remember that young adolescents are not adults. They need more time to think, as their mental processing speed is still lower (cf. Kail, 1993).

General Setting Notably with the younger children (7 to 10), the setting should be chosen with careful consideration of the demand characteristics of the room. The moderators should always be on the same eye level as the children (Annon, 1994; Delfos, 2000). Annon (1994) also notes that when a one-way mirror is used, it should not be on the same level as the children, as it may distract them. In setting the scene, it is also important to pay attention to the power balance. Morgan et al. (2002) explicitly chose an informal arrangement, in which all participants sat on soft mats on the floor in the middle of a pleasant light room in a community center. Furthermore, to reduce the hierarchical adult–child relationship, all used first names and all had colorful buttons with their names.

To promote group cohesion with these young children and to clearly communicate that interaction and participation are the goal of the session, group games are advised as warming up. Morgan et al. (2002) used a ball game to introduce the group members to each other; a ball was thrown to a group member, who had to state his or her name, favorite food, animal, and so on, and then throw the ball to another participant. This is also very useful to assess the cognitive and verbal development of the children and to tune into the child's language (Cares, 1999).

Similarly with children in late middle childhood (10 to 12), the setting should be chosen with consideration of the demand characteristics of the room, and the moderators should be on the same level as the children. However, one should avoid treating children this age as little ones, as they feel quite superior to the younger children in primary school. Warm-ups and informal introductions remain extremely important and age-related games play an important role in this. When moderators and children draw special name labels together, this helps to get acquainted and to reduce the authority imbalance (Hill et al., 1996). Nonverbal communication is an important part of controlling the group process and at regular times and after each subtopic, the moderator has to structure the session by summarizing and asking for additions from the children (Delfos, 2000).

For adolescents it is extremely important that the setting itself has no relationship at all with school or youth centers. It should be new and neutral territory for all, so that none of the adolescents is in a power advantage. Especially for the younger adolescents (12 to 16), careful monitoring of the group process is recommended, and shy adolescents should be encouraged. One way to do this is alternating the verbal discussions with other tasks. For instance, let each one individually write down what he or she thinks is important. The moderator can ask the quieter group members what they have written and so reduce dominations of the group by the more boisterous ones. Compared to adult focus groups, more time should be dedicated to warming up and acquainting the members with the rules and goals of a focus group. All focus groups are vulnerable to group pressure and conformity effects, but adolescents are more sensitive to peer pressure than younger children and adults. With adolescents, moderators have to be even more attentive to group processes, and give feedback when necessary. Finally, the moderators should realize that they themselves are *not* young (even if they are 22) and that fashions, music, and fads change very quickly (personal communication from M. Isacson) Moderators should never try to be one of the group, as in participant observation, and should never transcend their older adult identity (cf. Morgan et al., 2002).

20.3.3 In-Depth or Cognitive Interviews for Testing Questionnaires

Cognitive interviewing in the context of pretesting questionnaires is a form of in-depth interviewing used to find out what goes on in the head of a respondent when answering questions. The cognitive interview in questionnaire testing should not be confused with the cognitive interview in the context of law and child-witness literature. The cognitive interview of a child witness is a special structured interview taking the respondent step by step back to the event, and is explicitly designed to get more reliable reports on past events (e.g., Memon and Koehnken, 1992; Memon et al., 1996). To pretest questionnaires thoroughly, cognitive interviews are used to investigate the total question–answer process and discover sources of confusion and misunderstanding. This method is widely used as a pretest method to investigate the understanding of questions by *adults* and has proven to be successful in identifying potential problems in questions and in suggesting solutions for these problems (Campanelli, 1997; Presser and Blair, 1994; Willis et al., 1999b).

Potentially, cognitive pretesting of questions could also be a successful method with children and adolescents. It relies heavily on think-aloud procedures, which come very naturally to children. Young children often talk aloud in a noncommunicative manner during play or when performing tasks. According to the Russian developmental psychologists Vygotsky (Gray, 2002), this is a natural and necessary phase in the acquisition and internalization of language and verbal thought. Furthermore, think-aloud procedures are often used as an educational tool in primary and secondary schools, especially in teaching mathematics (Kraemer, 2002; P. Lynn, personal communication). Strangely enough, one of the first studies using cognitive testing procedures with young respondents (age 10 to 21)

reported that think-aloud procedures were problematic and that most teenage respondents lacked the ability or the motivation to articulate their thought processes spontaneously (Stussman et al., 1993). Blair (2000) also reports problems using think-aloud protocols with young children (6 to 11). However, both studies gave standard think-aloud instructions for adults, and the procedures were not adapted for younger respondents. Stussman et al. (1993) suggest that traditional cognitive interviewing techniques need to be modified for the young, with more attention to nonverbal communications and more probes. In addition, Blair (2000) comments that more introduction and explanation are likely to be necessary for children to be good respondents.

Think-aloud procedures with young respondents can work well, as Hess et al. (1998b; see also Zukerberg and Hess, 1996) showed. They conducted cognitive interviews with adolescents aged 12 to 17 to evaluate question understanding, task difficulty, and question sensitivity for the youth questionnaire in the U.S. Survey of Program Dynamics. The researchers developed a detailed protocol beforehand that included probing questions. They report that during the interviews they found a greater need to probe than they typically do during cognitive interviews with adult respondents. This corroborates the conjecture of Stussman et al. (1993) that the young need more extensive probing.

Levine and Huberman (2002) also used think-aloud techniques successfully to test questions on background information from the U.S. National Assessment of Educational Progress questionnaires with children aged 9 and 13 to 14. Levine and Huberman (2002) developed a detailed protocol with special probes for the cognitive interviews, and interviewers were trained to use them. The young respondents were given a special instruction and explanation of the procedure. Each think-aloud was preceded by having the respondent read the specific question aloud. This facilitated the detection of language and comprehension problems and served as a warm-up for the think-aloud. During the think-aloud the young respondents were continuously encouraged in a neutral manner, and probes were used frequently.

Unique in the Levine and Huberman study is that validating information was available based on responses by parents and teachers, which enabled comparison of revised questions with original questions. It is encouraging that Levine and Huberman (2002) showed that revised questions had a lower error rate.

Cognitive Test Interviews with Children and Adolescents

Using cognitive interviews for pretesting of children's questionnaires is possible and can result in worthwhile information, provided that the procedures are adapted to the special needs of children and adolescents. In the following paragraphs we discuss necessary adaptations to the general setup and protocol for in-depth interviews with adults. To accommodate different age groups, adaptations have to be made to all phases: arrival, introduction, start of the interview, interview proper, and ending (cf. Snijkers, 2002).

Arrival

In early and late middle childhood (7 to 12) special attention has to be paid to this stage. The child will be accompanied be a parent, caretaker, or

teacher, and both child and caretaker have to be welcomed and introduced to the interviewer, and time has to be taken to make the young child feel at ease. With children, the arrival stage includes many aspects of the introductory stage, too. Confidentiality and background information (why is the study done, etc.) have to be explained briefly to both parent and child. Therefore, some of the general procedures that with adult respondents are discussed in the introduction of the interview, are now introduced at the arrival stage when the parent or caretaker is still there [e.g., explaining videotaping, obtaining permission to record the session (both parent and child should give permission)]. In early adolescence, more often than not a caretaker will still accompany a child, and as a consequence, the arrival will take more time. With older adolescents, the situation more resembles the usual situation with adults. The arrival takes less time, with confidentiality and consent discussed during the introduction. However, in many countries, consent of a parent or caretaker is needed even for older adolescents (16 to 18) and should be obtained before the session.

Introduction

For a successful cognitive laboratory interview the introduction is crucial. In general, one has to take more time to explain what the rules are and what is expected than with adults. The importance of this is illustrated by Presser et al. (1993), who asked youngsters preinterview questions on what a survey was. They found that neither younger (6 to 8), nor older children (9 to 11) had a clear idea what a survey was and what the goals and rules of a survey were. More explanation of question asking and answering is needed with children than with adults.

Starting the Interview

Because the situation is completely new, the interviewer has to explain the procedures carefully, give clear examples, and practice the required tasks before the interview starts. For instance, one can rehearse thinking aloud using simple age-appropriate examples (e.g., a simple arithmetic task, a simple puzzle, sorting objects, etc). Extra time should be reserved for explanation and practice exercises, as part of a short training-phase before the real interview starts.

The Interview Itself

In general, the same rules of thumb for duration are valid as for focus groups. However, the estimates given for focus groups are the maximum possible. Because of the lively nature and potential for interaction, focus groups are in general more relaxed and demand less concentration than does an individual in-depth interview. Especially with the youngest age group, one has to watch the child carefully and react to drops in attention.

Different interviewing techniques for different age groups are advised. Think-aloud is very natural for young children (7 to 10), who often still read aloud. Levine and Huberman (2002) explicitly asked 9- and 13-year-olds to start by reading the question aloud. Not only did this stimulate them to think aloud, it also provided clues for further probing. For example, when a child could not read or pronounce a word correctly, this could indicate a comprehension problem.

During the think-aloud the interviewer has to be continuously alert, reinforce the child, and start up the process if the child stops for a moment (ask "Why do you stop"; if tired/not concentrating, suggest a short break). Both Hess et al. (1998a) and Stussman et al. (1993) recommend that the interviewer probe more frequently than with adults, and it is advisable to prepare a probing protocol and train interviewers to use frequent probes (Levine and Huberman, 2002).

It is important to make sure that the child feels completely at ease. Although thinking aloud is quite natural for young children, they will *not* perform well when they feel uncomfortable or watched. Young children can be very open in a situation they trust, but become completely shy and introverted when they find themselves in an unknown situation (Scott, 1997). In some cases it is therefore better to have a parent or caretaker present at the interview. Only when a young child feels comfortable will he or she perform well.

Paraphrasing is a technique that should not be used with younger respondents. Especially in young middle childhood (7 to 10), paraphrasing a question will not work, since children this age tend to repeat a question literally.

Late adolescents (16 to 18) may feel very embarrassed when asked to do a think-aloud. But paraphrasing combined with direct probes (e.g., "What does this word mean?" "What do you think it means?") may give good results in this age group. For adolescents, it is important for the interviewer to reinforce them and reassure them that this is not a school test and that not the adolescent but the questionnaire is being evaluated! Adolescents often lack confidence and may be unsure about themselves and their performance. Reassurance and frequent reinforcement is far more important for this group than for adults (cf. Hess et al., 1998b).

20.3.4 Auxiliary Methods

Observation   Monitoring of standardized interviews and self-administered questionnaire sessions is a relatively quick method that can provide useful additional information during field tests and pilot studies. Coding schedules developed for interviewing adults (e.g., Fowler and Cannell, 1996; Lessler and Forsyth, 1996; Oksenberg et al., 1991) are mainly for *verbal* behavior: for example, "interviewer reads verbatim," "interviewer deviates slightly," "respondent interrupts," "respondent asks clarification." Coding schedules for children should have more emphasis on *nonverbal* behavior, since children, especially younger children in middle childhood, will have more motor (movement) behavior. An example is provided by Presser et al. (1993), who developed and tested three interview protocols to measure daily food intake for children aged 6 to 11. They videotaped all test sessions and applied an extensive coding scheme with specific nonverbal codes for the child (e.g., head shaking, nodding, smiling) added to the standard verbal coding scheme of Oksenberg et al. (1991) for interviewer behavior. Presser et al. (1993) found that in the younger group, the interviewer deviated twice as much from verbatim reading of the questions as in the older group, and used more probes, indicating more problems in the question–answer process. They

also found that younger children smiled about three times as much as older children. This could indicate that young children will smile or laugh to hide that they do not understand a question. However, the fact that it is possible to code overt children's nonverbal behavior reliably does not mean that the interpretation is necessarily clear. In the field of child interviews, there is little work on the interpretation of coded behaviors, and more research and development is necessary. The newly emerging field of usability testing with children (Hanna et al., 1997) is facing similar problems, forcing researchers to acquire more methodological knowledge about children as subjects (Markopoulos and Bekker, 2002).

There are few examples of systematic observation of children during pilot testing of self-administered questionnaires. Researchers generally only note down the time it takes to fill in a test or questionnaire, to acquire data to improve planning the major fieldwork. An exception is the work of Helweg-Larsen and Larsen (2001, 2002), who observed both standard mainstream and special education students, aged 15 to 16, while they completed a pilot version of a Danish health survey. The special education students who had learning problems took longer and read at such a slow rate that they lost grasp of what had just been asked in the text. It became apparent that students in special education, but also a number of mainstream students, experienced literacy problems.

Debriefing   Interviewer and respondent debriefing studies have proved to be useful for studying response errors in survey data (e.g., Campanelli et al., 1991), and the observations of trained interviewers may provide worthwhile information on difficulties encountered in interviews with children. Until now this promising area has not been explored.

In a comparison of computer-assisted self-administered questionnaires with paper-and-pencil questionnaires in Dutch primary schools, Van Hattum and De Leeuw (1999) used a form of teacher debriefing in which teachers were asked about their experiences, the experiences of their pupils, and problems encountered during data collection. According to the teachers, asking sensitive questions (e.g., about bullying) by computer was less stressful than paper questionnaires for their young pupils (aged 9 to 12). Teachers also reported the problems their pupils had understanding several questions (e.g., meaning of certain words) but did not report any problems with the computer itself.

There are several examples of the use of respondent debriefing in surveys of children. Helweg-Larsen and Larsen (2001, 2002) in Denmark, and Watson et al. (2001) in New Zealand, added special debriefing questions at the end of computer-assisted questionnaires for adolescents. Topics included the computer interface, as well as privacy issues. Hess et al. (1998a) included debriefing questions in a field test of the youth questionnaire of the U.S. Census Survey of Program Dynamics. Like Scott (1997), they used a combination of Walkman and self-administered questionnaire, and at their debriefing focused on reactions to the audiocassette and privacy issues. Based on the debriefing results, the procedures were slightly modified to reduce repetition of the answer categories on the taped interview.

20.4   CONCLUSIONS

We discussed above various methods for pretesting questionnaires for children and adolescents. For the clarity of this chapter, we discussed each method separately, but this does not mean that in survey practice only one method should be used. In our opinion it is not either–or; the methods discussed in this chapter complement and reinforce each other and should be used in combination. This is clearly illustrated in the study of Presser et al. (1993; see also Blair, 2000), who used a variety of methods when developing interview protocols for food intake aimed at children aged 6 to 11. Besides think-aloud pretests, they compared different interview protocols and videotaped these for behavior coding. The same videotapes were also used as starting points in debriefing interviews. Data from all sources were combined to devise a new interview protocol for food intake. Another good example is the study by Reynes (2002; see also Reynes and Lorant, 2001), who used a combination of pretest methods when adapting the Buss and Perry Aggression Questionnaire to young French children aged 8 to 10. Experts were used to check the simplified vocabulary and sentence structure; the questionnaire was then pretested on 8-year-olds to make sure that all questions were understood; and in the final phase a pilot study was done on a large sample of 8- to 10-year-olds ($n = 500$) to check psychometric properties such as the reliability of the aggression scale. Hess et al. (1998a) used a similar procedure and combined the results of cognitive think-aloud interviews with those of a full field pretest to investigate potential problems in a self-administered questionnaire of adolescents (12 to 17) as part of the U.S. Survey of Program Dynamics.

Watson et al. (2001) and Helweg-Larsen and Larsen (2001, 2002) followed a slightly different procedure when pretesting health surveys for adolescents in New Zealand (12 to 18) and Denmark (15 to 16): After having completed the questionnaire in a pilot study, the respondent immediately took part in postpilot focus groups to investigate their experiences of the survey. Helweg-Larsen and Larsen (2001, 2002) also used systematic observation during the pilot.

Usually cognitive laboratory methods are used in a *pretest*, which is followed by a pilot or field test and the final study, but cognitive laboratory methods can also be useful as a *posttest* to gain insight into problems encountered during data collection or data analysis. Questionnaire test methods can be extremely useful after a survey is completed and when unexpected results are found, or in ongoing or longitudinal surveys. The goal of the questionnaire posttests is to identify sources of measurement errors encountered in the data. A prime example is the study of Jakwerth et al. (1999), who used standardized in-depth interviews to investigate reasons for the high item nonresponse rates reported over the years for the U.S. National Assessment of Educational Progress in achievement tests, for eighth graders (approximately 13 to 14 years).

Although in most research disciplines the instrumentation is checked, the methods vary. For example, in test development for educational research, an instrumentation phase is always included in which psychometric reliability and validity of the test are estimated on a large sample, while a cognitive pretest

of the questionnaire is rarely employed. In survey research, cognitive pretests are being used increasingly and pave the way for the costly pilot phase. In our opinion a cognitive pretest should always be part of the test design stage. It is very cost-efficient and gives a thorough insight into what may be wrong with questions and test items and suggests ways to improve them.

ACKNOWLEDGMENTS