Journal of Official Statistics
Vol. 8. No. 3, 1992, pp. 277–291

Anatomy of the Survey Interview

Wendy Sykes and Martin Collins[1]

Abstract: This paper sets out to describe in detail the verbal interaction that takes place between interviewer and respondent during a survey interview. This is achieved by applying the technique of Interaction Coding to 68 tape-recorded interviews carried out during the annual British Social Attitudes Survey. The pattern of observed behaviour is compared with what should be expected to occur during an interview which is progressing in line with the intentions of the research designer, as expressed through the medium of the questionnaire. An appreciable amount of behaviour is found to lie outside these limits, mostly reflecting respondents' difficulties with questions or response tasks. Much of it would not seem to threaten the integrity of the survey data, but the extent of reliance on the initiative and skill of individual interviewers is considerable, creating the risk of variability between them.

Key words: Survey interview; interviewer-respondent interaction; interaction coding; interviewer skills; interviewer variability.

1. Introduction

1.1. Background and objectives

The answers which are recorded in a face to face survey interview are the product of an interactive process involving interviewers, respondents and the mediating instrument – the questionnaire. Since the studies of Hyman, Feldman and Stember (1954) much research and experimentation have been conducted into the possible effects on survey results of interviewer administered questionnaires, especially those delivered face to face.

Despite all this activity, little seems to have been written about what actually *happens* in a survey interview: the *mechanisms* whereby errors arise. Some of the process is unobservable; some of it requires skilled interpretation of non-verbal behaviour (see Suchman and Jordan 1990). But what is *said* by the individuals concerned can be explored fairly easily. This is the main purpose of this paper: to provide a *description* of the verbal interaction in face to face survey interviews.

The pioneering work in this field was part of a programme of research on interviewing carried out by Cannell and others at the University of Michigan (Cannell, Lawson, and Hausser 1975). Using a simple code frame to classify the interaction between interviewers and respondents, the researchers identified a number of patterns of behaviour

[1] Centre for Research in Marketing, City University Business School, Frobisher Crescent, Barbican Centre, London EC2Y 8HB, England.

Acknowledgements: This paper is derived from an invited contribution to the International Conference on Measurement Errors in Surveys, in Tucson, AZ, USA, November 1990. The initial research on which it is based was carried out while the authors were working at the SCPR Survey Methods Centre, funded by the Economic & Social Research Council. Further financial and moral support was generously provided by the Joint Centre for Survey Methods.

which, although apparently innocuous, might affect the quality of survey data. For example, interviewers were found to be inadvertently providing reassurance to respondents who said they were unable to remember details of incidents about which they were being questioned.

The technique of Interaction Coding has since been elaborated and applied in various ways: to examine interviewer adherence to training (Cannell et al. 1975) and to test questionnaire design (Morton-Williams 1979; Morton-Williams and Sykes 1984; Oksenberg, Cannell, and Kalton 1991). However, the early work at Michigan stands alone in attempting to provide a descriptive framework for understanding what happens in survey interviews and for generating ideas about how data quality might be affected.

This paper seeks to build upon the earlier work through the development of a more detailed classification scheme and through more detailed analysis of the interactions between interviewer and respondent. It also attempts to add to the description of what happens during an interview an element of comparison with what might be termed the "ideal" interview. The starting point for this comparison, described in detail in the next section, is the stimulus-response model of the interview process (Brenner 1980; Deutscher 1972) where the "ideal" interview is characterised as an uncomplicated series of questions and answers.

1.2. An ideal model of the survey interview

Interviews in general have been described as "any interaction in which two or more people are brought into direct contact in order for at least one party to learn something from the other" (Brenner, Brown, and Canter (eds.) 1985, p. 3). Fowler and Mangione (1990) have suggested several features that distinguish the survey interview from other types of interview, most importantly that individual respondents are of interest only as members of a population that is to be described and that the object is a quantitative description of the population or of sub-groups.

This objective implies that the information produced should meet strictures regarding the essential characteristics of "scientific" data. Only then can the analyses be performed that are the aim and purpose of survey research. According to Galtung (1970) scientific analyses can only be performed on data which satisfy the criteria of classification, completeness, and comparability (see Brenner et al. 1985). Survey data should then consist of answers obtained in response to a standard stimulus (the questionnaire) and, for each question asked, an answer and only one answer should be obtained from each sample member.

To meet these aims, questions are carefully constructed to be understandable and to have fixed and common meaning. They must also represent feasible response tasks that people are willing to undertake. Then, all answers should be in the same frame of reference or dimension and to the same level of specification.

The efforts of the questionnaire designer are supported by training interviewers to deliver the questions as worded and in the order specified. They are also trained to probe *open* questions in standardised and non-directive ways which preserve the integrity of the question design. Any interaction unrelated to the question and answer process is discouraged.

If the aims of the questionnaire designer and the interviewer trainer were fully realised in practice, the outcome would be the ideal interview: a straightforward stimulus-response process consisting of

question, answer, question, answer, and so on. In this paper, we will be looking at the kinds of departures from this ideal which are commonly met in actual interviews.

1.3. Departures from the ideal

The almost robotic simplicity of this ideal may seem implausible in the context of a social interaction, even one as constrained as a survey interview. In particular, politeness appears to demand that an interviewer should, on occasion, acknowledge or confirm receipt of an answer. We should perhaps accept as a norm the somewhat extended sequence of question (including scripted introductions and explanations), answer, acknowledgement, question, answer, acknowledgement, and so on.

This concession to reality should not, however, be granted lightly. The line between a neutral acknowledgement and positive feedback is ill-defined, and we have already referred to Cannell's finding that interviewers are not always discriminating in giving such feedback if it is not scripted. The extended sequence is, therefore, a reasonable model against which we can compare observed events. It is not the ideal.

Much of what is regarded as acceptable in survey interviews is, in fact, outside even these limits. Interviewer training programmes include standardised ways, developed over the years, of dealing with breakdowns in the question and answer process. These hiatuses occur because questionnaire design is an imperfect process and because respondents are not trained in their role within the survey interview. So, for example, misunderstandings of questions are countered by interviewers being trained to repeat the offending item, and respondent difficulties with certain response tasks are met with standardised explanations. Such solutions may be generally accepted, but they do reflect the incidence of problems – departures from the ideal. Our concern here is as much to chart the incidence of these normal problems that call upon the standardised strategies taught to interviewers to help preserve the reliability of survey data, as to explore deviations from standard practice.

1.4. Interaction coding

The methodology used to explore the events taking place in survey interviews is Interaction Coding, introduced into social psychology by Bales (1950) and into survey research by Cannell (1973), Cannell et al. (1975), and Cannell, Miller, and Oksenberg (1981). Subsequent developments have come from, amongst others, Brenner (1982), Dijkstra and Van der Zouwen (1982), Morton-Williams (1979), and Sykes and Morton-Williams (1987).

The basis of the procedure is the division of "a meandering and sometimes apparently seamless flow of behaviour into separate units or chunks" (Clarke 1983). Each unit may consist of a word, phrase, or sentence which is then classified in terms of its main function (e.g., question delivery, request for information, response to a question). Partitioning of the behaviour flow and the assignment of codes to the units of behaviour is carried out simultaneously. Evaluation of the way in which a function is performed is included in the coding of some behaviour units (e.g., giving a response adequate for coding). Other characterizations go even further, for instance, ascribing a failure to provide an adequate answer to misunderstanding of the question or response task. Finally, some absences of expected behaviour are coded, for example, a failure to probe or omission of a question.

The code frame used, of which a full version is given in the Appendix, was evolved

through listening to and examining the transcripts of a large number of interviews (not those used in this analysis). It was refined in discussion and application.

1.5. *The interviews analysed*

Our analysis is based on 68 interviews carried out in the 1984 British Social Attitudes Survey. The survey, one of a series carried out annually by Social & Community Planning Research (SCPR), involved a probability sample of the adult population of Great Britain, selected from the Electoral Registers. The study interviews were obtained with haphazardly selected respondents: 31 men, 36 women and one whose gender was not recorded; mostly aged under sixty (51).

Each interview lasted approximately one hour and involved 151 questions and question parts. Items were mostly pre-coded and included single and multiple response questions which were either read aloud or administered with a showcard. Topics covered a wide range of social, moral, legal, and political issues. The questionnaire, which had been thoroughly pre-tested, was designed by researchers with considerable experience. The interviews were carried out by 12 well trained and experienced interviewers. In other words, it can be assumed that all reasonable steps had been taken to meet the highest professional standards.

1.6. *Coding the data*

The coding exercise was carried out by one coder who had been thoroughly trained in the use of the code frame. This involved a verbal description of the code frame and its application, followed by a practical session with feedback from the researcher.

To test the reliability of the code frame, 20 interviews were each coded by two trained coders working independently. The researcher then compared the two coders' output and resolved differences between them (sometimes without agreeing with either). The result, illustrated in Table 1, was a high degree of agreement, especially between the researcher and the coder eventually assigned to the task.

Table 1. Code frame reliability test

	%
Researcher and both coders agree	88
Researcher agrees with final coder	5
Researcher agrees with other coder	1
Researcher agrees with neither	6
Total	100

Base: All behaviour units identified

2. The Survey Interview in Practice

2.1. *The overall pattern*

In reporting our results, we look first at units of behaviour in isolation. As Table 2 shows, 45% of the total of some 38,000 units of behaviour were contributed by the interviewer, 55% by the respondent. In terms of function, about half (51%) arose from the core activities of asking the questions (including reading scripted introductions and instructions) and providing response. This is the behaviour that would be expected under the most naive conception of an interview as a simple to-and-fro stimulus-response process. The other half of the verbal activity lay outside these basic limits, in the form of subsequent interaction between interviewer and respondent (with two-thirds of the activity provided by the latter).

Much of this subsequent activity has to be expected and forms part of a more realistic view of what will happen during an interview. It reflects the fact that the interview is an interaction between two humans rather than an automated process. So, we find

Table 2. The incidence of verbal behaviour (percent)

	Contributed by:	
	Interviewer	Respondent
Stage of process:		
Initial stimulus	28	
(Interjection)	(1)	
Initial response		23
First interaction	9	1
Subsequent interactions	7	31
Total behaviour	45	55

Base: All units of verbal behaviour (37, 906).

activity which represents only polite conformity to social norms: interviewers acknowledging a response or respondents indicating that a response is complete. Some behaviour represents only a trivial departure from the question-answer process, for example, respondents thinking aloud before answering. Some of it is even anticipated in the research design, for example, the interviewer using a *scripted* probe in order to obtain a more detailed response at some questions. As we will show, just over 80% of all verbal behaviour fell within these extended limits of what we can expect to observe in a smoothly running interview: 84% for the interviewer and 80% for the respondent.

Not all the remaining verbal behaviour is necessarily damaging or in any sense "wrong." A respondent may ask that a question be repeated or give a clear indication of a failure to understand. The interviewers may quite properly, in line with their training, repeat a question or use a non-directive probe to clarify an answer. This behaviour does, however, lie outside the formal framework provided by the questionnaire and implies weaknesses in that framework. Several authors have shown that it is at this stage that problems of interviewer variability can arise, due to the dependence on the interviewer's initiative and skills, especially in probing. (See, for example, Collins 1980; Shapiro 1970.) The level of occurrence of such behaviour, even with a highly structured and well-developed questionnaire, illustrates the importance of training interviewers to deal correctly with departures from the pattern envisaged by the designer of the questionnaire.

2.2. *The initial stimulus*

In looking at the results in detail, we can break the overall interview process into constituent parts, beginning with the delivery of the initial stimulus: the question together with any scripted introduction or instructions regarding the response task. Here, as Table 3 shows, events were reassuringly close to the ideal implicit in the provision of a script in the form of the questionnaire.

There were a few departures from this script. Thus, of all the occasions when an introduction was used, 4% involved the interviewer providing their own unscripted introductory or linking phrase. Of all the question-asking occasions, 4% involved changes to question wording beyond those specifically allowed at repetitive questions. On the same percentage of occasions the interviewer, instead of waiting for a response, volunteered an unscripted intervention.

The departures from the questionnaire

Table 3. Delivery of the stimulus

Introductions:	Base: All uses of introductions	(994)
		%
	Read as scripted	96
	Unscripted introduction	4
	Total	100
Questions:	Base: All questions asked	(8,875)
		%
	Read as worded	88
	With permitted abbreviations	8
	With minor changes	2
	Routing error	1
	Reading interrupted	1
	Total	100
		%
	Waited for response	96
	Unscripted interjection	4
	Total	100
Task explanations:	Base: All scripted explanations	(862)
		%
	Delivered as required	100

framework are few and would mostly be judged to be non-prejudicial. But they *are* departures and provide early signs of the propensity of the framework to break down during the interview. And they are coming from the fully trained and instructed member of the partnership between interviewer and respondent.

2.3. *The initial response*

Turning to the respondent's initial response to the stimulus (in Table 4) we begin to find rather more departures from the desired pattern. Encouragingly, 84% of the initial responses were – as required – adequate answers which the interviewer was able to code. But a sizeable minority of responses did not conform to the theoretical strictures and indicated, for the most part, difficulties with comprehension or with ability to answer.

Note again that we are not saying these departures are necessarily harmful. Indeed, it may be that we should prefer the respondent to seek clarification rather than provide an answer to a question that has not been understood. It was, however, the intention of the research designer that the question stimulus should yield an appropriate response. Any other outcome creates the

Table 4. The initial response

	%
Answer adequate for coding	84
Requests repeat	2
Requests clarification	1
Requests task explanation	1
Thinks aloud, then answers	1
Inadequate answer (misunderstanding)	4
Inadequate answer (other)	2
Unable to answer	2
Answer appears invalid	2
Total	100

Base: All initial responses (8,605)

need for further interaction which may be more harmful.

2.4. The first interaction

Since interviewer and respondent typically take turns in speaking, it is not surprising that most (93%) of the first interactions following on from the initial response are from the interviewer. We have already referred to two kinds of behaviour at this stage which have to be expected in a social interaction: acknowledgement that an adequate response has been given (including rhetorical confirmation of a response) or a *scripted* probe used to elicit a codeable response, especially to an unfolding question. As Table 5 shows, such behaviour accounted for about two-thirds of interviewer activity at this stage.

The remaining 34% of interviewer activity mostly comprised various attempts to help respondents provide adequate answers. Most of the behaviour – repeating questions or using non-directive probes or clarifications – is what interviewers are trained to do when a problem arises. It is not, in that sense, wrong but the decision to use it and the quality of its delivery depend on interviewer initiative and skill. It is not part of the fully scripted process and provides an opportunity for interviewers to behave in inconsistent ways. This is the point at which we believe, from other studies, that interviewer variability is likely to be introduced.

Of the small amount of respondent behaviour at this stage, 29% consisted of a slightly belated answer to the question, following immediately on an initial failure. Other behaviour indicated problems with the response task: requests for repeat of the question or for clarification. But the largest category suggests discomfort with the constrained role implicit in a survey interview, with respondents feeling the need to explain their answers. In a tiny minority of cases they went so far as to seek from the interviewer an evaluation of their answer.

2.5. Subsequent interactions

The final stage of the interview includes everything that goes on between interviewer and respondent after the first interaction. In line with interviewer training, most such behaviour arose from the respondent.

Looking first, however, at the small amount of interviewer behaviour, Table 6 shows a pattern like that in the first interaction. Just over 70% was acknowledgement or scripted probing, part of our expectation of the interview. Most of the rest again consisted of attempts to help the respondent, the only new category being a small number of attempts to control digression or other extraneous behaviour by the respondent.

In respondent behaviour, the main category was that of providing a delayed but adequate answer to the question. Other behaviour showed increasing diversity. At the stage of the initial response to the stimulus, respondent behaviour was spread over 9 categories; at the first interaction it was spread over 11 categories; these later interactions cover 13 different categories of behaviour. As might be expected, more of what might be termed "closure" activity took place: confirmation of the given response or an indication that the answer to a multiple response question was complete.

3. Interaction Sequences

3.1. Recording sequences

So far, we have described in fairly simple terms what goes on in a survey interview. We have identified a limited range of behaviour which we would expect to find in an interview which conformed closely to a stimulus-response process and we have

Table 5. The first interaction

By interviewer	Base: All interviewer behaviours	(3,445)
		%
	Acknowledgement	36
	Confirmation	18
	Scripted probe	12
	Question repeat	10
	Non-directive probe	7
	Confirmation of question meaning	5
	Clarification of task	3
	Unbiased information	1
	Failure to probe	6
	Directive probe	1
	Total interviewer	100
By respondent	Base: All respondent behaviours	(258)
		%
	Answer adequate for coding	29
	Requests repeat	3
	Requests task explanation	3
	Requests clarification	2
	Changes answer	2
	Thinks alound, then answers	1
	Explains response	46
	Digresses	7
	Other extraneous behaviour	2
	Inadequate answer	2
	Seeks evaluation of answer	1
	Total respondent	100

noted other types of activity which occur in practice. We have not yet taken account of the sequence of behaviour at a question.

To do this we have employed a complex recoding process which seeks to identify straightforward sequences – those which fit *broadly* with a stimulus-response model of the interview. This has been done for both single and multiple response questions. For the latter, of course, the acceptable sequence can be quite long.

At the first stage, behaviour codes were grouped into equivalence categories which seemed sensible in the current context.

A = Scripted introduction read as worded or non-directive introduction added by the interviewer.

B = Question read as printed or with only minor changes, or with permitted abbreviations.

C = Adequate explanation of the response task.

D = Question answered adequately for coding by the interviewer or thinking aloud before giving a codeable response.

E = Interviewer acknowledges adequate response or rhetorically confirms

Table 6. Subsequent interactions

By interviewer	Base: All interviewer behaviours	(2,583)
		%
	Acknowledgement	37
	Confirmation	16
	Scripted probe	18
	Non-directive probe	6
	Question repeat	5
	Confirmation of question meaning	4
	Clarification of task	3
	Unbiased information	2
	Control of extraneous behaviour	1
	Directive probe	5
	Failure to probe	2
	Total interviewer	100
By respondent	Base: All respondent behaviours	(11,929)
		%
	Answer adequate for coding	76
	Confirms response	5
	Requests repeat	2
	Requests clarification	1
	Indicates answer complete	1
	Requests task explanation	1
	Thinks alound, then answers	1
	Inadequate answer (misunderstanding)	4
	Unable to answer	2
	Answer appears invalid	2
	Inadequate answer (other)	1
	Digresses	1
	Explains response	1
	Total respondent	100

the given response.

F = Respondent confirms response.

G = Interviewer uses scripted probes adequately.

H = Non-directive probes, single word probes, or "expectant" silences at multiple response questions.

I = Respondent indicates answer is complete at a multiple response question.

At the second recoding stage, sequences or combinations of the first-order equivalence groups were themselves combined into groups:

J = Groups or sequences to do with "legitimate" question reading (B, BC, AB, ABC).

K = Groups or sequences to do with acknowledgements by the interviewer or confirmations by the respondent (E, EF).

L = Groups to do with legitimate ways

of getting full answers to multiple response or open questions (G, H).

M = Groups or sequences to do with closures of multiple response questions (I, KI).

Thirdly, longer sequences using codes constructed at the first two stages were defined:

N = JD (Question plus answer).

or = JDK′ where K′ includes all multiples of K such that JDK′ embraces JDK, JDKK, JDKKK, etc. (Question, answer, acknowledgements and/or confirmations).

O = LD (Probe plus answer).

or = LDK′ (Probe, answer, acknowledgements and/or confirmations).

At the fourth and final stage, a single composite code was defined to cover all straightforward completions (questions plus answers) of *multiple* response questions:

T = NO′, NO′M, ND′, ND′M, ND′K, ND′KM, ND′K′, ND′K′M′, N(DK)′, N(DK)′M. (As above, a prime mark indicates that all multiples of the character are allowed; (DK)′ indicates all multiples of the sequence DK, i.e., DK, DKDK, DKDKDK, etc.)

The recoding of the data set in accordance with these rules resulted in a new data set consisting of some primary codes, some first stage recodes, some second stage recodes, some third stage recodes, and some final stage recodes. It was then possible to identify all questionnaire items where the question and answer sequence was in line with the broad view of what should happen:

.N. for single response questions, where the points denote the beginning and end of a question and answer sequence, and

.T. for multiple response questions.

Table 7. The incidence of straightforward sequences

All closed questions	(8,853)	77%
Single response read out	(5,631)	77%
unfolding	(680)	73%
showcard	(1,871)	87%
Multiple response	(351)	32%

All other sequences were characterised at this stage as non-straightforward, to be explored in more detail.

3.3. *Straightforward sequences*

As shown in Table 7, straightforward sequences occurred on 77% of occasions when questions were asked. In general, single response questions generated more such sequences than did multiple response questions. Those where the respondent chose an answer from a showcard performed especially well in this respect, with 87% of all such questions executed in a straightforward manner. On the other hand, only one-third of multiple response questions (also using showcards) were completed in such a manner. (The very small number of open ended questions has been omitted from this analysis.)

3.4. *Exploring the non-straightforward sequences*

Non-straightforward sequences – those which departed from even a broad stimulus-response model – were of a bewildering variety. A few recurrent patterns could be recognised and grouped in further recoding of the data set:

P = Requests for clarification of the question or the response task, requests for repeat of the question.

Q = Clarification of the task, repeats of the question, other confirmations of question meaning.

R = Answers inadequate for coding due to misunderstanding of the question or task or for some other reason, answers which appear to be invalid.

S = Combinations of clarifications requested and given: Q′, (PQ)′, Q(PQ)′.

U = Combinations of inadequate answers and clarifications given by the interviewer: RQ′, (RQ)′.

We must stress again that some of these sequences – which we describe as non-straightforward – are not necessarily wrong: we would prefer that a respondent should seek clarification rather than offer a probably superficial answer to a question that has not been understood. But it is the function of questionnaire design to avoid such problems and they must, to that extent, be unexpected. Equally, contributions from the interviewer may be in line with their training, but only in so far as that training is concerned with dealing with breakdowns in the framework for data collection provided by the research designer. Such breakdowns would not, in the ideal world, occur.

For *single response "read" questions* the largest category of non-straightforward sequences was of cases where the initial answer was inadequate for coding by the interviewer. In most such cases, the situation was handled by interviewers using approved methods: a non-directive probe or a repeat of the question. There were, however, occasions when the interviewer accepted the answer and failed to probe, or used a directive probe or prompt.

The second most common category for such questions involved clarification by the interviewer, either volunteered or offered on request immediately after question delivery. In most cases, the respondent went on to give a codeable response but there were a few instances where the question still caused problems and further probes were needed to complete the sequence. In isolated cases, even after clarification, the respondent felt unable to answer and said so.

Another major category of departures from the intended pattern of behaviour was where respondents expressed inability to answer immediately following question delivery. In a minority of cases, interviewers probed and obtained an adequate answer. More commonly, no further effort was made by the interviewer to pursue an answer.

Other departures involved the tendency for respondents to give an answer but then feel the need to explain it to the interviewer. Finally, there were cases where question delivery did not conform to expectations: questions not read as scripted, interruptions by the respondent, and questions omitted in error.

For *unfolding questions* the main categories of non-straightforward sequences were very similar: sequences involving the need for probing or clarification; respondents explaining their answers or expressing an inability to answer. Another major category, however, was of cases where the question had not been delivered in the two stages intended.

Questions using showcards – which suffered from the lowest incidence of non-straightforward sequences – again yielded sequences involving clarification by the interviewer (usually resulting in an adequate answer) and sequences where probing was needed to get the answer. In a minority of cases, inadequate or unlikely answers were accepted by the interviewer or probed directively.

As with the other question types, explanations of answers given and claimed inability to answer accounted for most of

Table 8. The incidence of straightforward sequences for individual questions

Base: All closed questions	%
90%+ straightforward sequences	18
80–89%	36
70–79%	26
60–69%	13
50–59%	4
Under 50%	3
Total	100

the other common forms on non-straightforward sequences. In this case, problems with question reading were rare.

For *multiple response questions*, half the non-straightforward sequences involved respondents indicating verbally that their answers were complete. We would now accept this as necessary behaviour which ought to be expected. Then, multiple response questions would compare less unfavourably with single response questions, although we would still show a higher proportion of non-straightforward sequences for the former.

The only other major category of non-straightforward sequences was that involving the need for clarification by the interviewer. Unlike the other three question types, few difficulties were encountered with answers needing probing.

3.5. *Results for individual questions*

Across all types of single response questions, examination of the incidence of non-straightforward sequences for individual questions shows that few questions were close to being problem free (see Table 8). About one in five questions yielded 90% or more straightforward sequences. At the other extreme, four questions seemed especially problematic, with fewer than 50% straightforward sequences. The worst case – with only 24% straightforward sequences – was a question at which respondents were asked to choose a code letter from a card to indicate the income from all sources of the household. Inability or reluctance to answer and the need for clarification of the question or the response task were widespread. A subsidiary finding is a slight tendency for the incidence of straightforward sequences to increase through a series of questions with the same format and response task.

4. Discussion

It may at first seem disturbing that only about half of the verbal interaction in these interviews was concerned with the core tasks of the interviewer asking and the respondent answering questions. We have shown, however, that much of the remaining behaviour is inevitable in a social interaction. Taking such behaviour into account, we suggest that in a smoothly running interview the interviewer should be expected to:

- Read scripted introductions;
- Read questions as instructed;
- Explain response tasks as instructed;
- Use *scripted* probes;
- Acknowledge or confirm responses.

These categories of behaviour accounted for 84% of interviewers' verbal input to the interview. Most of the rest would be deemed to be harmless; some of it even helpful. It is mostly in line with the way that interviewers are trained to cope with problems, e.g., by repeating questions or using unscripted non-directive probing to obtain an adequate answer.

Nevertheless, the fact that such behaviour is needed does show that problems have arisen. To this extent, the framework for data collection planned by the research designer and imposed through the questionnaire has broken down. And we should remember that this was a fairly simple, if

long, interview, using a pre-tested questionnaire and carried out by well trained and personally briefed interviewers. Not all surveys are this good.

In responding to deviations from the planned framework, interviewers have to apply skills learned in general training and through experience. Our work in the field of interviewer variability leads us to the firm belief that it is variation between interviewers in their use of these skills (especially probing) and in their approach to other shortcomings – helping the respondent by clarifying the question or the task, or helping the researcher by clarifying ambiguous answers – which is most often at the root of any problem. It is in this context that we find the fact that a sixth of interviewer behaviour is in these types of activity a matter for concern.

In parallel with the definition of the interviewer activities which ought to be expected in an interview that is progressing according to plan, we can define the forms of verbal behaviour which ought to be expected from a respondent as:

- Providing an adequate answer at any point, not necessarily immediately after the first asking of the question;
- Indicating that an answer to a multiple response question is complete.

This accounts for 80% of respondent verbal behaviour. Much of the rest indicates problems with questions or response tasks, recognised by the respondent and leading to requests for help. But there are also examples of inadequate or inappropriate responses which suggest the same problems, not recognised or acknowledged by the respondent. Finally, there are signs of discomfort with the highly artificial and constrained role set for the respondent by the rules of the interview.

Again, we ought to recognise that other more complex or less well developed questionnaires might pose greater problems for respondents. Even here, it is the 20% of respondent behaviour which deviates from our expectations that provides the opportunity for the interviewer to have some effect on the survey data.

Our deliberately limited objective in this paper has been to provide a descriptive background against which other research and evaluation can take place. The results do, however, also focus attention on two broad areas. The first is that research designers have to be aware of the propensity for the data collection framework of the scripted questionnaire to break down. This has implications for the thorough development and testing of the questionnaire, possibly using the same kind of coding and analysis as we have used here. The second, perhaps even more important area is that of interviewer training: to recognise when the planned framework has broken down and to apply corrective actions in a consistent manner.

Appendix

Coding Frame

Question Delivery

Introductions to questions (*including stems to battery questions*)

Read as worded
Non-directive introduction made up
Not read as worded
Directive introduction made up

Questions (*including battery items*)

Read as worded
Acceptable abbreviations to repetitious questions
Read with minor changes only
Not read as worded
Reading interrupted
Skip error

Explanations associated with question delivery

Adequate explanation of task
Inadequate explanation of task
No explanation of task

Question Answering

Respondent

Answers adequately for coding or recording
Thinks aloud before answering
Changes answer
Answer possibly invalid
Requests clarification of question
Asks for repeat of question
Repeats question
Asks for information about survey
Accepts clarifications or offers of clarification
Explains response
Confirms response
Indicates answer complete
Extraneous but non-digressing behaviour
Cannot answer
Answer inadequate for coding(indicates misunderstanding)
Answer inadequate for coding (other)
Asks for evaluation of response
Refuses to answer
Digresses

Interviewer

Acknowledges response
Uses scripted probe
Clarifies respondent role
Confirms question meaning
Repeats question
Non-directive probe
Silences and noises as prompts
Gives unbiased information about survey, sponsors, etc
Controls digressions
Confirms response
Offers clarification
Digresses
Fails to control digression
Makes negative comments about survey, sponsors, etc
Evaluates response
Interprets question
Anticipates response
Incorrectly confirms response
Probe implies response inadequate
Directive probe
Fails to probe
Omits scripted probe

5. References

Bales, R.F. (1950). Interaction Process Analysis. Cambridge, Mass: Addison-Wesley.

Brenner, M. (1980). Assessing Social Skills in the Survey Interview. In Singleton, Spurgeon and Stammers (eds.) The Analysis of Social Skills. New York: Plenum.

Brenner, M. (1982). Response Effects of Role-Related Characteristics of the Interviewer. In Dijkstra and Van der Zouwen (eds.) Response Behaviour in the Survey Interview. London: Academic Press.

Brenner, M., Brown, J., and Canter, D. (eds.) (1985). The Research Interview: Uses and Approaches. London: Academic Press.

Cannell, C.F. (1973). Some Experiments in Dyadic Interactions and Response Accuracy in Survey Interviews. Michigan: Institute for Social Research.

Cannell, C.F., Lawson, S.A., and Hausser, D.L. (1975). A Technique for Evaluating Interviewer Performance. Michigan: Institute for Social Research.

Cannel, C.F., Miller, P.V., and Oksenberg, L. (1981). Research on Interviewing Techniques. Michigan: Institute for Social Research.

Clarke, D. (1983). Language and Action. Oxford: Pergamon.

Collins, M. (1980). Interviewer Variability: A Review of the Problem. Journal of The Market Research Society, 22, 77–95.

Deutscher, I. (1972). Public and Private Opinions: Social Situations and Multiple Realities. In Nagi and Corwin (eds.) The Social Context of Research. London: John Wiley.

Dijkstra, W. and Van der Zouwen, J. (eds.) (1982). Response Behaviour in the Survey Interview. London: Academic Press.

Fowler, F. and Mangione, T. (1990). Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Sage Applied Social Research Series, No. 18.

Galtung, J. (1970). Theory and Methods of Social Research. London: George Allen & Unwin.

Hyman, H., Feldman, J., and Stember, C. (1954). Interviewing in Social Research. Chicago: Chicago University Press.

Morton-Williams, J. (1979). The Use of Verbal Interaction Coding for Evaluating a Questionnaire. Quality & Quantity, 13, 59–75.

Morton-Williams, J. and Sykes, W. (1984). The Use of Interaction Coding and Follow-Up Interviews to Investigate Comprehension of Survey Questions. Journal of The Market Research Society, 26, 109–127.

Oksenberg, L., Cannell, C., and Kalton, G. (1991). New Strategies of Pretesting Survey Questions. Journal of Official Statistics, 7, 349–365.

Shapiro, M.J. (1970). Discovering Interviewer Bias in Open-Ended Survey Responses. Public Opinion Quarterly, 34, 412–415.

Suchman, L. and Jordan, B. (1990). Interactional Troubles in Face-to-Face Survey Interviews. Journal of the American Statistical Association, 85, 232–241.

Sykes, W. and Morton-Williams, J. (1987). Evaluating Survey Questions. Journal of Official Statistics, 3, 191–207.

Received February 1991
Revised January 1992