

Karol J. Krótki, University of Alberta

A completed questionnaire from a social survey is both an incomplete and overperfect record of what transpired between the respondent(s) and the interviewer. The recorded answers abstract from the interview dynamics and simplify the complexity of answers into categories acceptable to the supervisors of the interviewer, even in the case of open ended questions, and certainly so in the case of pre-coded answers. The written entries throw no light on the process through which the respondent arrived at his answers, except when the interviewer offers marginal comments, or supplementary questions ask for the recording of any documents that might have been consulted (e.g., tax statements in the case of income). Completed questionnaires give no information on the "cost" of a particular question in terms of time and strain on the goodwill between the two parties. There are no means on the questionnaire to indicate how far the interviewer was forcing the conversation and answers to his questions into moulds strange, ill understood by the respondents and away from what the respondent may have wanted to talk about.

Disciplining the respondents' answers does not end with whatever forcing the interviewer engaged in. It is continued through the processes of mechanical and manual editing of data often in line with plausible and preconceived values for the different parameters. The existence of mathematical and theoretical models lends credence and support to such activities and culminates in further adjustment to data, both at the micro and at the macro level. However, this paper is not concerned with these latter influences.

In an earlier paper techniques of analyzing taped interviews were assessed (Krótki, 1973). Briefly, with a survey of any size (that is excluding short journalistic or psychological interviews with individuals), transcription seems to be the quickest and least expensive step leading to categorizations, coding systems, and aggregations.¹ This will be unwelcomed news to those who experienced the painful and bothersome procedures of transcription and were looking forward to some alleviation from their travails. In this paper we move a step further and ask what hypotheses can be tested through tape recording on the assumption that the taped interviews can be suitably processed. Such processing would include transcribing, categorizing relevant features of the taped interview, coding, quantifying wherever appropriate, and aggregating for suitable comparisons with other data, including questionnaires from interviews not tape recorded. There are dozens of little tricks of trade, ignorance of which makes taped interviews that much less valuable, all of which are also outside this paper.²

Verbatim recording of interviews and the subsequent use of the recordings can have a number of purposes and can be done in a variety of ways. Transcription on location during the interview even by skilled stenographers loses one third to one half of the interview due to conscious and unconscious selection of what to write (Bucher et al., 1965a; 1965b). When note taking is considered more obtrusive than tape recording during clinical, counselling, and psychological interviews the value of the tape lies in the possibility of almost limitless replay and unhurried consideration of each detail of the conversation (Brody et al., 1951; Eitzen, 1952). Tape recordings are being used to record hot news and immediate reactions. Such recordings, however, are seldom done on a large scale with structured questionnaires. In this paper we are interested not only in repetitive interviews leading to structured records, but also in the limited aim of evaluation, i.e., not in recording to produce substantive data. Audio-visual recordings are used when it is felt that some of the dimensions of the interview dynamics escape audio tapes (Kantner & Zelnik, 1969).

In the process of evaluation some data correcting procedures may be refined, but we resist the endless temptations to fork out in the pursuit of anthro-cultural and socio-psychological curiosities that arise in almost every minute of many an interview. Our purpose is to report on the testibility of objective and quantified hypotheses through tape recorded interviews. A number of hypotheses in six taxonomic groups have been chosen and their potentialities or actual experiences, where already available, are reported upon.³

The first question that arises with tape recordings is how far the process distorts the interview. The distortion can be two-fold: relatively to other non-taped interviews and relatively to actual "facts" as otherwise reportable. The overwhelming opinion of participants seems to be that there is little distortion (brief summary in Krótki, 1973). These considerations pose the first three hypotheses, all three methodological (M) and tape (T) related. They are followed by other methodological (M) hypotheses that concern survey (S) questions as such arising without tapes.

(MT 1) Tape recording does not distort the content of the interview relatively to interviews not tape recorded, though as suggested in hypotheses (MT 2) the effect of recording is not entirely neutral.

A rigorously conducted experiment involving a priori interpenetrated samples, or less certainly a posteriori matched groups, should provide definite answers anthropological culture by culture.⁴ Instead, we have in literature less rigorous enquiries ending impressions.

This writer can add to them his own surprise how unobtrusive is the obtrusiveness of the microphone. Protruding right in the middle of the conversing group it seemed to him to have been quite neutral in the many Moroccan homes, where the taping operation was observed during the population census of 1971. To respond objectively and definitely to hypothesis (MT 1) it would be extravagant to study each substantive characteristic. It might be enough to obtain some objective indicator such as the number of interviews conducted during the same unit of time by the same interviewer in comparable circumstances when taped and when not taped. Another possible indicator could be the proportions of "refusals" and "no response" on days when taped and when not taped. During the 1971 population census of Morocco some of the work of 43 interviewers was taped.⁵ The taped/not taped comparison was possible for only 30 interviewers, the work circumstances of the other 13 not being comparable. Of the 30 "comparables", 15 did less work on days when taped, 12 did more, and 3 did the same amount of work. It would appear that the taping had no effect on the length of the interview.

(MT 2) Tape recording improves the quality of the interview relatively to "true" facts.

In survey evaluation respondent's answers are sometimes compared with such records as family documents or employer's data. Provided the numbers of recorded and non-recorded interviews are large enough a macro comparison is then possible between the parameters obtained from the two (clearly, no micro or case-by-case comparison is possible). There is apparently little to choose between the validity of the recorded and not recorded interviews when compared with other validating data. Such distortions or losses as are found, appear to be related to social class: lower classes reply more accurately when taped, higher classes less accurately, at least in some British surveys (Belson, 1964; 1967).

(MT 3) Transcriptions, other forms of transformation from audial to optical record, and translations done twice by labour with same training disclose variable discrepancies without evidence of bias. High level labour is necessary to disclose biases.

Evidence available to date is uncertain and experiences are inconclusive, though the first part of the proposition is logically unassailable: there is no way to choose between two alternatives of equal quality.

The foregoing three hypotheses concern methodology of tape recording. The next nine concern the respondent-interviewer reactions that affect the methodology (M) of survey (S) taking. Some subtle hypotheses that would escape any attempt at quantifying from the tape are not tackled: e.g., the differences between the respondent's and interviewer's perceptions of the interest in and the nature of their interview.

(MS 1) The interviewer tends to conduct shorter and shorter interviews as the time of the day (particularly under trying climatic conditions) and the length of the interviewing campaign extends.

It is hypothesized that with time the interviewer:
becomes more impatient,
tends to ask leading questions,
interruptions not immediately relevant replies,
selects more understandable words,
adopts a more efficient mode of questioning.

Average daily production is a good indicator, but it cannot distinguish between the effects of boredom (the first two items) and the learning process (the last two items). The coding from tapes should differentiate (as it did not in Morocco) between features that indicate flagging "good behaviour" and features that suggest improvements in interviewer's competence. The interviewer can speed up the interview through the vertical method hypothesized in (MS 5). It appears that the vertical method arises as the day wears on, rather than with the dragging on of the campaign. Interviewers do start every morning with good intentions.

(MS 2) The type of the interviewer's interventions envisaged in hypothesis (MS 1) is more frequent with respondents that are illiterate, female, older or of lower socio-economic strata.⁶ There is little in the Moroccan data to affirm any such pervasive bias.

"To be sure, one can hear in the tape recorded conversations impatient asides from a census-taker confronted with an illiterate woman who insists on a lower age for herself than the period of her declared residence in the city. A stammering old man's attempt to supply ages for his family are ignored by the (interviewer) and the child to whom he turns for answers. Some interviews include almost no questions concerning age, this information being calculated by the (interviewer) who reads the family civil paper. but such interviews are the exception, and evidence of pervasive cross-interview biases is weaker than expected" (Davis, 1973:47)

(MS 3) There is no difference in the mode of obtaining data and their reliability when answers are offered by one person only, by one person supported by documentary evidence, by several persons, particularly when each "self" replies for himself/herself.

Extreme instance can arise, though not on the Moroccan and Gambian tapes, when a whole village assists at the interview and objects loudly and chorally to the claim of an elderly female respondent to be still on the reproductive side of menopause. More generally a "self" takes more time to give information about himself than about others. In Morocco in 1971 questions concerning age (3.1 vs 2.5), birth place (5.5 vs 3.6), migration (4.8 vs 3.8), length of residence

(2.7 vs 2.1) provoked all more utterances in respect of the responding head of the household than for his spouse (Davis, 1973, tables B3 through B10). Whether this is due to the initial difficulty of understanding the question when posed for the first time or due to there being more information available to the respondent about himself (with consequent greater reliability of self-answered questions?) cannot yet be said.

(MS 4) The questions are not asked as formulated because of the tendency to slip, particularly in the case of answers, into generally prevailing and known definitions, no matter what the "official" survey definitions.

High marks can be given in this respect to taped interviews.⁷ An anthropologist working on these data (Susan Davis) has several interesting examples in this regard.

(MS 5) To ask the same question for all members of the household (vertically) is faster than all the questions for each member separately (horizontally).

High marks to taped interviews. For the Moroccan data the hypothesis has been confirmed for the age question and for the migration question (Davis, 1973). However, the faster, vertical method is liable to confusion of answers between members of the household. For a relevant finding see hypothesis (MS 1).

(MS 6) There is no objective way to differentiate between replies to questions that were not understood and where the answer was unknown to the respondent.

Subjective impressions can be formulated by a listener, but he longs to have an audio-visual taping to check out his impressions. No ideas on quantification, from transcriptions alone, have arisen.

(MS 7) The path from the initial response (eh?) through all the intermediaries to the finally recorded answer is a tortuous one and the number of utterances in between is inversely proportionate to characteristics listed in hypothesis (MS 2).

High marks can be given to taped interviews for their detective ability in this respect, but real benefits accrue only under special circumstances. If the distribution of female respondents is bi-modal - some shy in the presence of strangers, some more talkative than male respondents - the fact may be interesting, but cannot be easily used to improve the operations of the next survey. If the length in the interview varies with illiteracy of respondents who are evenly distributed over the country (Davis, 1973:35), nothing can be done about it. But if illiteracy is concentrated in certain strata, then it should enter significantly into the cost functions and influence sampling design. Strata by themselves were not a useful variable in Morocco (Davis, 1973: 49).

(MS 8) The use of a foreign language, literary rather than colloquial vernacular, and superior or educated accents increases the number of utterances during an interview.

High marks go to taped interviews in this respect provided transcriptions have been suitably coded. Whether the more subtle question of rapport and understanding between the two parties can be equally assessed is less certain.

(MS 9) Number of lines on transcription vary with socio-economic data and they also differ with the type of question.

Results of hypothesis (A1) and (MIG 2) indicate that interviewing in lower socio-economic strata (villages and shanty towns) is more expensive than in higher strata. However, strata are compounded with literacy. For survey design the finding on strata can be used, but not that on literacy. High marks go to taped interviews because of their ability to measure the "cost" of each question in terms of time or in terms of space on the transcription. The extraordinary similarity between the number of lines taken by the different subjects in table 1, except for household composition in hypothesis (C2), is remarkable for such different anthropological cultures as Gambia and Morocco.

To estimate completeness of coverage of a survey, taped interviews are a weak instrument, but at least two hypotheses can be suggested.

(C1) Under no circumstances is there ever an indication that whole dwellings are missed, but there is a slight possibility that other households within a given dwelling are being discussed in a given household in such a manner that indications of missed households are obtained. This is so especially, when definitional problems arise in multi-household dwellings and/or multi-family households. There is no information that this possibility has been used.

(C2) Questioning about the whereabouts of age and sex groups particularly vulnerable to omissions is less intense (shorter in duration) than the importance of these groups (their predilection to be omitted from the age structure) would justify. Visitors are also enquired about less than the de facto and de jure problems would justify.

This possibility has not been tested. The question on household composition in Morocco, unusually expensive (57 lines in table 1) as compared with Gambia (17 lines), invites further exploration. The other relevant question, called in the table "family relations" called for 9 and 10 lines on the Moroccan and Gambian transcriptions respectively.

For the substantive discussion we selected three topics: age distribution, migration and occupation. Some results are available from Morocco and Gambia with regard to the first two. A thorough analysis would code utterances concerning age for use of documents (separating documents issued at the time of vital events

from documents completed retrospectively), the use of historical calendars, the family history method, the community comparative method, eye estimation, confident knowledge of respondents. There was no fertility in the Moroccan questionnaire and only few fertility questions in Gambia. Consequently, no experience in this all important field is available similar, e.g., to the embarrassment reported on pregnancies, when still invisible, in Niger (Pool & Pool, 1971) and the consequent possibility of underreporting.

(A1) The age estimation provokes exchanges for heads of households, particularly with characteristics of hypothesis (MS 2).

High marks. In Morocco confirmation has been obtained for all the stipulations, except for the socio-economic strata. The contribution of strata is compounded with the other variables. However, this hypothesis is useful for sampling design, because one can "get at" these other variables through geographic strata.

(A2) The age estimation provokes more exchanges for households with documents and for members of lower birth order (younger children).

High marks. Confirmation has been obtained in Morocco for both stipulations. It would thus appear that consulting documents adds to the expense, but it also is likely to add to the accuracy, except when the document itself is doubtful (Krotki, 1973).

(A3) The likelihood that the interviewer uses leading questions when enquiring about age is greater when the head of household or speaker has characteristics of hyp. (MS 2).

High marks go to taped interviews for their ability to detect questions inviting agreement.

(A4) The likelihood of use of historical reference dates is greater when head of household or speaker has characteristics of hyp (MS2).

High marks. There is an inclination among respondents in Morocco to prop their memories with references to personal events rather than societal events. This inclination was shared by respondents in Niger (Pool & Pool, 1971). The ratio of personal to societal reminiscences is higher the higher the proportions of the variables mentioned in the hypothesis (Davis, 1973: 40). For a slightly different experience in migration see hyp. (MIG 3). As suspected by the proponents of the historical calendar (Scott & Sabagh, 1970: 106) the link between the two series is weak.

(A5) The likelihood of broad and facile estimation of period of birth, of obvious guessing, of mistakes in mental arithmetics is greater when the head of household or speaker has characteristics of hypothesis (MS 2).

A hypothesis postulated "since Plato", but still of uncertain operationality.

(A6) Estimation of ages in decades or round numbers is more common for speakers with characteristics of hypothesis (MS 2).

One does not need taped interviews when the respondent is recorded on the questionnaire. When civil registration documents were originally issued on the basis of retrospective estimation the rounding disappears (deceptively). Other deceptiveness arises when interviewers calculate the age, repeat it aloud and implore the respondent to remember it from now on (Davis, 1973:31).

(MIG.1) Migration questions take up a high proportion of interview exchanges and time. The migration questions appear to be expensive in both surveys on table 1, even in Gambia where they were limited to the place of birth and tribe.

(MIG.2) The number of exchanges per migration question(s) is greater for respondents with characteristics of hyp. (MS 2).

High marks. In Morocco the complexity of questions other than on migration, was felt more keenly in the shanty towns than in the rural douars (Davis, 1973: 41). It cannot be said whether that would have been the experience with migration questions (if migration questions were asked from the rural questionnaire, which they were not) in view of the presumably simpler migrational history of the country people.

(MIG.3) The migration question "How long ago did you leave ..." is more easily answered if asked "How old (big) were you when ..."

High marks. The Moroccan tapes detect a tendency to switch to the unofficial form of the question (Davis, 1973: 47).

(MIG.4) The timing of migration questions caused more utterances than the geographic dimension of the moves.

Potentially high marks. Somewhat inconsistently with the outcome of hypothesis (MIG.2) and contrary to the expectation in hypothesis (A.3), there were more personal references than societal references on the timing of migration among the literate group than among the illiterates (Davis, 1973: 38).

(MIG.5) Controlling for variables in hyp. (MS.2), there are more leading questions for migration than for other subjects.

Potentially high marks. The hypothesis is based on the thought that the length and complexity of the migration questions in Morocco were an additional inducement to cut corners.

(MIG.6) Incomplete definitions concerning the urban-rural dichotomy and the "first city lived in" provoke much discussion.

Potentially high marks. It remains to be seen whether the general hypothesis (MIG.5) can be separated from the specific hypothesis (MIG.6). The operational consequences of the two hypotheses would be different.

(OCC.1) The number of exchanges between the interviewer and the respondent concerning occupations, is a function of whether or not the respondent is the subject.

Potentially high marks.

Table 1. - Lengths of exchanges between interviewer and respondent during the Gambian survey of 1972 and the Moroccan census of 1971

SUBJECT (1)	Question numbers		Average lines		Average time - M	
	G (2)	M (3)	G (4)	M (5)	min. (6)	sec.
Household composition	a	U2	17	57	2'	53"
Family relations	b,c	U4-U8	10	9		28"
Age	d,e	U9-U10	23	21	1'	15"
Migration	f,g	U12-U16	21	48	2'	44"
Education	h,i	U22-U25	15	14		43"
Employment	l,m,n o,p	U28-U31	21	17	1'	25"
Irrelevant exchanges			13	13		45"
Non-comparable			34	117	10'	13"
TOTAL			154	296	16'	54"
Age as % of total			15%	7%		7%
Migration as % of total			14%	16%		16%

The table is based on 62 interviews from Gambia and 25 urban interviews out of 34 measured out of 800 taped interviews in Morocco; the "non-comparable" questions in the last line are mainly literacy, fertility and housing⁸

(OCC.2) Questions concerning unemployment produce less unemployment reported in the aggregate than the residual unemployment arising out of questions on employment.

Asking where and when the subject worked seems to draw away attention from unemployment. Labour force participation can, probably, not be studied without taped interviews.

(OCC.3) Serendipitous benefits of questions on occupations are lost in analysis.

Exchanges concerning occupations results in little occupational detail, but give considerable insight into the quality of life as perceived by the respondents, the impact of God Almighty, the method and chances of gathering a bountiful harvest, the fertility, location and distribution of the farmer's land. These details could become inputs into an aggregated subjective social indicator on quality of life rather than the construction of a meaningful occupational distribution. Taped interviews are a potentially valuable instrument in this regard.

(OCC.4) The richness of detail on occupations varies with the socio-economic level of respondents and subjects.

It is not clear a priori whether the variation is direct or inverse and whether it occurs with both the respondent and subject or only one.

(OCC.5) Whether in a developed society or underdeveloped, female respondents tend to overassess the occupation of their menfolk towards occupations with greater prestige.

Can be done only with some kind of follow-up and case-by-case check. Provided the respondent is recorded on the questionnaire there is no need for taped interviews in this respect.

It will be seen from the above summary of suggestions that the record of past achievements of tape recorded interviews is so far rather modest. Their further potentialities are mixed. Much of the disadvantage is not inherent in the method, but in the difficult circumstances in which it is being tried. As a minimum, however, the possibility remains that with a continuing development a new, even if only additional, instrument of objective evaluation will become available. The role in refining data collection procedures is more certain.

FOOTNOTES

- ¹An attempt to short-circuit the transcriptions of the 1971 Moroccan interviews through a form with mere check listings, resulted in eight hours work for the migration and age questions alone on an average questionnaire, not very much less than the transcription of a whole questionnaire (Krótki, 1973, table 2). For more general introductions to the Moroccan data see Krótki, 1972; Krótki & Quandt, 1972; Quandt, 1972.
- ²For example, the counting of transcription lines might give the same measure of "cost" as watch-measured timing (Krótki, 1973). For confirmation see table 1. The less expensive line counting requires some readiness to standardize transcriptions. Inter-language comparisons can be deceptive: French texts are typically one third longer than English, while texts in agglutinative languages can be still shorter.
- ³The thirty hypotheses listed in this paper rests on the work of researchers who in the early months of 1972 were engaged on the Moroccan data: Mohammed Abzahd, Mohammed Ayyad, Douglas Davis, Susan Davis, Karol P. Krótki III, Dona MacLaren, Anna Quandt. For earlier ideas on audio-taping credit goes to John Blacker, Christopher Scott, William Seltzer.
- ⁴Matching in this case means aggregate matching of two groups (experimental and control) on as many characteristics as possible. It is not the matching required for record linkage and related procedures.
- ⁵For facilitating the tape recording in Morocco thanks are due to Tayeb Bencheikh, the then Director of Statistics, Abdelmalek Cherkaoui, his successor, Abdessattar Elamrani-Jamal, the Census Director, and Mohammed Rachidi, the Director of the Demographic Centre.
- ⁶For the Moroccan data socio-economic strata were approximated by geographic data. When we say "lower socio-economic strata" we mean simply "villages and shanty towns".
- ⁷If social surveyors use "official" definitions that are different from "native" ones, they break a fundamental rule of survey taking.
- ⁸David C. Roberts, the Government Statistician of Gambia made generously available to several researchers sets of transcriptions of interviews from the pilot census. Thanks are due to Don Peirce, a graduate student at the University of Alberta, for work on the materials available to the author.
- ⁹Douar is an Arabic word, used in Morocco to describe villages, or tribal areas, or more generally areas inhabited by country people.

REFERENCES

- Belson, 1964 - William A., "Readership research in Britain". *Business Review*, 6 November.
- Belson, 1967 - William A., "Tape recording: its effects on accuracy of response in survey interviews". *Journal of Marketing Research*, 4:253-260, August 1967.
- Brody et al., 1951 - E.B., Richard Newman, and F.C. Redlich, "Sound recording and the problem of evidence in psychiatry". *Science* 113:379-80, April 1951.
- Bucher et al., 1956a - Rue, C.E. Fritz, and E.L. Quarantelle, "Tape recorded interviews, some field and data processing problems." *Public Opinion Quarterly* 20:427-439, Summer Review.
- Davis, 1973 - Douglas, Some effects on census interview content of interviewer and respondent characteristics. Haverford, Pa: Haverford College. Typescript.
- Eitzen, 1952 - D.D., "Objective recording procedures in counselling and research". *Marriage and Family Living* 14:225-8, August.
- Kantner & Zelnik, 1969 - John F. and Melvin, "United States: exploratory studies of Negro family formation - common conceptions about birth control". *Studies in Family Planning* 47:10-13, November 1969.
- Krótki, 1972 - Karol J., "Programme d'exploitation de l'exercice EREB/CETI". *Note d'Avis No. 21*, 20 May 1972. Rabat, Maroc: Centre de recherche et d'etudes demographiques.
- Krótki, 1973 - Karol J., Audiotaping of interviews for evaluation of social surveys. Annual meeting of the Statistical Sciences Association of Canada. Kingston, Ontario: Queen's University, June 1973. Reprinted as Discussion Paper No. ..., Population Research Laboratory, University of Alberta.
- Krótki & Quandt, 1972 - Karol J. and Anna, The CETI/EREB exercise. Programme of analysis. Census evaluation by taped interviews. Rabat, Maroc: Centre de recherche et d'etudes demographiques.
- Pool & Pool, 1971 - Janet E. and D.I., "The use of tape recorders to ascertain response errors in a KAP survey, Niger, West Africa." Annual meeting of the Population Association of America, Washington, D.C., April 1971.
- Quandt, 1972 - Anna, The Collection of tape-recorded interviews during the 1971 census of Morocco. Document Technique du C.E.R.E.D., No. 4E. Rabat, Maroc: Centre de recherche et d'etudes demographiques, May 1972.
- Scott & Sabagh, 1970 - Christopher and Georges, "The historical calendar as a method of estimating age: the experience of the Moroccan multi-purpose sample survey of 1961-63". *Population Studies* 24(1):93-109.