Journal of Official Statistics, Vol. 11, No. 2, 1995, pp. 147–160

Understanding the Standardized/Non-Standardized Interviewing Controversy

Paul Beatty[1]

The debate of how "standardized" the survey interview should be is examined, exploring the nature of the past development of standardization, reasons for current standardization practice, and shortcomings that are being addressed by recent calls for less standardized interviews. Various literature on interviewing – historical, methodological, and actual interviewing manuals – is examined. Viewed together, these materials show the evolution of the role of interviewer. Interviewers were first employed to "humanize" the research process, but practical concerns led to a drastic reduction in the amount of control they had in the research process. The notion that "rapport" was important allowed interviewers to retain some control of the interaction with respondents; when this idea lost popularity, an extreme move toward eliminating interviewer error began. The future challenge for researchers is to retain the benefits offered by standardization while resolving legitimate shortcomings of standardized interviewing as currently practiced. Possible strategies are discussed.

Key words: Survey research interviewing; data collection; open-ended questions.

1. Introduction

About fifty years ago, a controversy between two seemingly irreconcilable schools of thought in the survey field reached full force. Some researchers believed that obtaining useful survey information required loosely structured "in depth" interviews; others countered that more structured surveys were highly preferable to this cumbersome method. Lazarsfeld's (1944) "Negotiation" between the two viewpoints is widely regarded as a classic, suggesting common ground that future researchers could work together on. Ironically, today we are witnessing a controversy that is strikingly similar in some important ways. Some researchers argue that standardized interviewing is inherently riddled with interviewer-respondent communication problems;

[1] Office of Research and Methodology, National Center for Health Statistics, 6525 Belcrest Road, Hyattsville, MD 20782, U.S.A.
Acknowledgements: An earlier version of this paper was written while the author was a Research Associate at the Survey Research Center, University of Michigan, and presented at the May 1993 AAPOR meeting held in St. Charles, IL. This paper has benefited greatly from the input of many people. I would like to thank most prominently Charles Cannell for his insights, comments, and encouragement in pursuing this topic; Howard Schuman for comments and encouragements with early versions of this paper; and Beth-Ellen Pennell for involving me in earlier research which was the basis for this work. Views expressed are those of the author and do not necessarily reflect the views or practices of the Survey Research Center at the University of Michigan, or the U.S. National Center for Health Statistics.

consequently, social scientists should rely on more conversational interviews to obtain valid information (Suchman and Jordan 1990; Mishler 1986; Briggs 1986). Supporters of standardized interviewing defend the validity of the methodology in spite of some common pitfalls, arguing that standardization is an essential characteristic of modern survey research (Fowler and Mangione 1990). Recent criticisms of standardized interviewing have been the focus of increasing debate (Schaeffer 1991), and a clear-cut compromise between advocates of standardization and non-standardization is not apparent. These interviewing styles differ in many substantial ways.

If one understands the conflict between the two viewpoints, however, there are logical compromises. But understanding how this conflict evolved is a tricky matter. Looking at methodological pieces from the past on interviewing is insufficient, as they provide little insight into how interviewing was actually conducted; likewise, interviewer training manuals from the past explain little of the basis for the strategies they employed. To completely understand the forces that have moved surveys toward standardization, as well as those that have resisted, it is necessary to look at both works together. When this is done, new insights emerge regarding what issues have been left unresolved in the debate, and what common ground exists between the two schools of thought. Though no one can predict the future with certainty, the most logical directions that researchers should follow seem surprisingly clear.

2. The Early Development of Standardized Interviewing

Standardization as it exists today was developed to gain tighter control of the errors produced by interviewers. Researchers maintain this control through the precise scripting of questions and development of standardized behavior appropriate for interviewers to follow in all situations. The goal, as noted by Groves (1989, p. 358), is "nothing less than the elimination of the interviewer as a source of measurement error." This has not, however, always been a priority in survey research; in fact, this goal developed quite gradually.

A fundamental debate motivated changes in the degree of question scripting and interviewer standardization: should interviewers be "knowledgeable experts" who "conversationalized" respondents through interviews, using questionnaires as guidelines; or should interviewers be non-experts who obtained information on a mass-scale using standardized techniques? This debate goes back to the 1930s when interviewers were first used for academic research. Most researchers viewed interviewers as an escape from the restrictions of closed, self-administered questionnaires and were anxious to use this potential freedom (Converse 1987). The U.S. Department of Agriculture's Division of Program Surveys, headed by Rensis Likert, pioneered the methodology. This method was a scientific extension of an earlier program within the Department, in which mobile observers discussed an informal set of topics with arbitrarily selected farmers. With the new method, the questions were standardized, probing was neutral, and probability sampling was used. However, the interviews retained a conversational tone, consisting entirely of open-ended questions, and interviewers necessarily had wide latitude in determining when respondents had adequately answered the questions. To compensate for that, interviewers were highly trained, generally holding graduate degrees. Following the field period, the open-ended data demanded meticulous attention for complex coding. The process had high validity, but was expensive and time consuming. There were reliability concerns as well (Hyman 1991; C.F. Cannell, personal communication, May 1993).

This method was initially viewed as being more sound than traditional "polling" methods, which relied on unskilled interviewers asking easily-coded "yes/no" type questions. While the polling method generated faster answers at significantly lower cost, more academically-minded researchers viewed poll questions as "incomplete, artificial, and rigid, almost bound to distort people's attitudes" (Converse 1984, p. 272). Then Lazarsfeld (1944), in a widely known article, forcefully argued that a compromise position was viable – relying on "interlocking poll questions" with some open questions providing elaboration. With that, a move toward greater standardization was begun (Hyman 1991).

Still, while questions became more closed and interviewers became part time and less-skilled (C.F. Cannell, personal communication, May 1993; Hyman 1991), standardization did not develop to anywhere near the extent we see today. By 1945, the National Opinion Research Center (NORC) had made major strides toward standardizing interviews and reducing the amount of interpretation required by interviewers, but some open questions were still used. At the Bureau for Applied Social Research, Lazarsfeld had encouraged compromise, but always believed that the meanings of questions should be more fixed than actual wordings. Consequently, they still used looser forms of interviewing. The University of Michigan's Survey Research Center (SRC) also compromised. They moved toward standardization, but continued a considerable commitment to asking open questions, even though they had no experimental evidence supporting the utility of the method (Converse 1987).

Standardization continued to grow because it was cheaper, faster, and seemingly effective. Furthermore, Hyman (1954) began to rigorously explore the effect of interviewer variation resulting from non-standardized behavior, which paved the way for further methodological work. It is logical, then, to wonder why resistance to standardization continued. SRC's *Manual for Interviewers* (1954, pp. 29–31, 33) provides some insight into the continuing compromise position there. The manual balances such statements as "the questionnaire helps to standardize the interview," with instructions such as "use the questionnaire, but use it informally;" another set of instructions reads "ask the questions exactly as worded on the questionnaire," while a few pages later, the interviewer is empowered to "reword the question slightly ... as a last resort" to clear up misinterpretations. The manual instructs interviewers to work in a dual role, as a "technician who applies standard techniques to each interview," while maintaining the role of "human being who builds up a permissive and warm relationship with each respondent." On the one hand, we see a genuine desire to embrace standardization – even though interviewers are given some leeway, they are instructed to use it with great caution. There are many mentions throughout the manual that standardization is a key component of surveys' usefulness. "Obviously," states the manual, "if a question is differently worded for different respondents, it will not yield comparable results." Given a statement such as that, it is surprising that interviewers maintained as much discretion as they did.

3. The Importance of Rapport

It is also clear, however, that another issue is considered of great importance for the success of survey interviewing: rapport between the interviewer and respondent. Likert's (1947, pp. 199–200) interest in open interviewing can be partially explained by its perceived importance in generating rapport:

> Most attitude scale research was based on work with students. They, at times . . . insisted that none of the alternatives gave them an opportunity to express their attitudes correctly. These objections were usually ignored . . . When it was tried on the adult population . . . respondents, in their homes, were told that the fine points in their thinking did not matter, that they had to restrict their thinking on the problem to the dimensions seen by the experimenter . . . [and then] rapport went out the door and often the interviewer was close behind . . . The importance of respondent co-operation throughout the interview and the necessity of limiting the procedures used in polls and surveys to those which build and develop this rapport does not yet seem to be recognized generally.

Likert adds, in an otherwise diplomatic piece, "No wonder most polls restrict the interview to not more than ten minutes." For Likert, rapport and open-ended questions were closely linked – if the interview was too structured, rapport would necessarily suffer.

Clearly, Likert's definition of rapport involved a spirit of cooperation and respect between the interviewer and respondent. This type of task-directed "rapport" is still a goal of survey researchers today (C.F. Cannell, personal communication, May 1993). Nevertheless, the concept was never well-defined. Hyman (1954) was one of the first to recognize the difference between "total" involvement with the interviewer and more fundamental "task" involvement. Task involvement could help respondents assume their roles and increase response validity – but total involvement could lead to response bias, with respondents agreeing with perceived interviewer opinions. Later studies conceptualized rapport as involving both professional interest and personal warmth (Weiss 1968; Henson, Cannell, and Lawson 1976).

While researchers failed to clarify the ambiguity of the term, they simultaneously regarded rapport as essential – and rapport was often seen as contradictory to standardization. SRC's *Manual for Interviewers* (1954, pp. 25 and 34) makes this quite clear. The word "rapport" appears dozens of times, most prominently in this statement of purpose: "Your goal is 'rapport' with the respondent . . . [This] designates the *personal relationship* of confidence and understanding between the interviewer and respondent which provides the *foundation for good interviewing*" (italics mine). Elsewhere, whole sections are designed to help maintain rapport ("the interviewer recognizes that good rapport must be maintained throughout the interview to insure [sic] full and valid information") that explain its many benefits (such as "good rapport stimulates discussion"). In short, the 1954 manual stresses the importance of rapport above virtually anything else. The manual also indicates that too much standardization was at odds with rapport building. Consider this excerpt:

> You should keep the questionnaire in sight during the interview, glancing at it

> before asking each question. Put each question to the respondent in a natural and conversational tone of voice, not obviously reading it. Try to avoid drawing too much attention to the questionnaire since your goal is to set up a friendly relationship between *yourself* and the respondent; too much obvious attention to the questionnaire makes for an atmosphere of interrogation, which is something you want to avoid. Each question should be asked in a manner implying that it presents an interesting topic, and that you are extremely interested in having the respondent's ideas on it (p. 31).

Although standardization seemed more desirable over time, the higher importance of rapport stood as a barrier to its development. However, the original "task-oriented" conceptualization of rapport no longer fit, with several references to personal relationships and personal interest in respondents' answers. A clean definition of the concept remained elusive in spite of its deemed importance.

Weiss's (1968) study of response validity delivered the first serious blow to rapport. Interviewers in this study rated respondents on a scale from "confiding" to "hostile," taken as a measure of "total involvement." Respondents' answers were verified through record checks. Surprisingly, the most "confiding" respondents provided the least valid information, indicating that this type of rapport could actually be harmful.

Further attacks on rapport developed from the fact that it had no consistently applied definition. Goudy and Potter (1975, p. 541), citing various definitions used over the years, noted that no matter how rapport was defined, there was no evidence that it improved interviewing productiveness; in fact, citing an observation by Gordon (1969), they questioned whether rapport might actually be destructive: "Often the neophyte thinks he has conducted an excellent interview because 'rapport was perfect' and the respondent was 'completely at ease, talked spontaneously, and documented that she had enjoyed the interview.' Yet when the interview is analyzed for the amount and clarity of relevant data, it is found to be incomplete, superficial, and ambiguous." Goudy and Potter (1975, p. 543) therefore concluded:

> When we realize that problems exist in reading questions correctly, recording information, and coding that information [they cite multiple sources here] then studying rapport seems less important. Unless agreement on a conceptual and operational definition can be reached and unless elements in the interviewing situation purporting to constitute rapport can be isolated and tested, further empirical studies of rapport may be useless . . . [Researchers should] abandon the concept [of rapport], admit that interviewing is an art that may contain certain scientifically controllable elements, and work to attain reliability and validity in interview data through interview effects currently measurable.

Thus, by the mid-1970s, survey researchers were questioning the wisdom of the rapport-building approach, turning their attention to "interview effects currently measurable." Research on controlling interviewer variation through standardization, previously hindered by concerns for rapport, had new opportunities to flourish.

4. Reduction of Interviewer Effects

The demise of rapport has made a major impression on interviewing methodology as practiced today. Until this point, relatively flexible interviewing could be defended on several grounds: first, it facilitated rapport-building, providing a basis for the interviewer-respondent relationship. This was regarded as essential for validity and generating discussion. Also, but less noticeably, it provided a diagnostic capability for communication problems. Of the two justifications, rapport arguably attracted more attention; with that gone, there was little standing in the way of more standardized interviews. Meanwhile, researchers had become increasingly curious about the errors introduced by interviewers. Kahn and Cannell (1957) developed theoretical frameworks to describe interviewer/respondent interactions in psychological terms, while Hyman led a series of NORC projects on interviewer effects, culminating in his 1954 book. These works, however, were primarily descriptive. For example, Hyman explored whether interviewers' own opinions and ideologies influenced respondents. His goal, though, was *understanding* where these biases were, not formulating a plan to reduce them. *Reducing* interviewer error was a different matter altogether, as this would require exercising a great deal of control over interviewers – not an option considering the importance of rapport. The interest in controlling interviewer effects had clearly grown as rapport's importance faded. Cannell, Marquis, and Laurent (1977, pp. 1 and 77) noted that in spite of increasingly complex demands from survey research, "little has been done toward perfecting the interview procedure as a method of data collection." Their summary of interviewing research led to specific recommendations for interviewer controls, which would "improve communication between the participants in the interview but will not direct the responses." As more time passed, methodological work was created with the goal of stamping out interviewing error altogether. Such efforts focused on the development of training techniques and precise prescriptions for what interviewers are told to do (Groves 1989).

Fowler and Mangione (1990) provide a thorough overview of the philosophy and practice of standardized interviewing as it has become. They recommend reading questions exactly as worded, using only nondirective probing, and maintaining a neutral interviewing stance with no interviewer discretion in recording answers. They also call for practice-oriented training, systematic interview monitoring, and close supervision, in further attempts to control interviewer variation.

But perhaps the result of this movement may be examined most clearly in its practical application: the interviewing manual from SRC (1983), *General Interviewing Techniques* (GIT). The GIT manual differs from the 1954 manual chiefly in its emphasis on examples designed to develop precise, standardized behavior from interviewers. Gone is the 1954 manual's friendly prose describing the ideal relationship between the respondent and interviewer, motivated by explanations about how the interviewer fits into the big picture. GIT is an ambitious technical document, three times the size of its 1954 counterpart, filled with expanded sections on question asking, clarification techniques, probing, feedback, and so on. Its purpose is not general description; it is intensely specific, containing dozens of examples and seven lengthier exercises for trainees.

Consider the case of probing: the 1954 manual devotes eight pages to the topic, offering a list of possible probes to use when rapport alone was insufficient to elicit complete answers. The 1983 GIT devotes 30 pages to the topic, providing a complete list of acceptable probes, with practice exercises to illustrate proper applications. Training materials such as this clearly document the attempted suppression of any interviewer variation.

The increase in survey standardization clearly brought many advantages. First, the survey process became much faster. This is particularly important in an age when surveys are the basis of forecasting elections and economic outcomes, developing business strategies, and making other time-sensitive decisions. Without standardization, it is difficult to conceive the amount of processing that would be necessary before data would be available for large academic research programs. Statistical precision is another benefit; increasingly elaborate statistical methods may be applied to data collected through standardized methods. Stouffer et al.'s (1949) landmark analyses relied on simple tabulations; analyses today draw upon such techniques as multiple regression and log-linear models, which require much stronger assumptions about the data. Also, as we have seen, standardization gave survey researchers new power to study and attack the most easily measurable sources of survey error, including interviewer error.

There are, however, important sources of survey errors that are more difficult to measure. Less structured interviewing, as mentioned previously, had two advantages. The most noticeable of them, rapport, had been discredited. Yet the other, diagnosing communication problems, had not been addressed directly as interviewer style changed. This shortcoming, I argue, is the root of the current standardized/non-standardized interviewing controversy.

5. The Current Controversy

The basis of current attacks on the usefulness of standardized interviewing is that it strangles communication with respondents. Interviewers have very little discretion to exercise in resolving communication problems. A possible alternate type of interviewing, as described by Groves (1989, p. 404), dictates that "interviewers should be trained in the concepts inherent in the questions and be allowed to probe, rephrase, and adapt the questionnaire to individual respondent needs." Mishler (1986) and Briggs (1986) provide detailed justifications for non-standardized interviewing and suggestions for alternate strategies. Mishler calls for interviewers and respondents to work together to "jointly construct" meanings of questions and answers, while Briggs lays out a plan for "acquiring metacommunicative competence" to interpret interview data.

More recently, Suchman and Jordan (1990, pp. 232 and 241) claim that the conflict between interviews as conversations and interviews as data-collection interactions has never been adequately resolved. According to their argument, surveys rely on conversational norms to succeed (i.e., asking and answering questions) while suppressing "interactional resources that routinely mediate uncertainties of relevance and interpretation [in conversations]". They conclude that "the standardized interview has become such a fragile, technical object that it is no longer viable in the real world of interaction" – a strong indictment. However, the entire debate, based on a resurgent interest in diagnosing communication problems, is quite understandable when one considers the evolution of interviewing in theory and practice. These problems have been set aside in the ambitious effort to wipe out interviewer error, and Suchman and Jordan are simply bringing them back to the forefront.

Critics of Suchman and Jordan might counter that though survey interviews do have certain problems, their representations are highly overstated. They surely present worst-case scenarios: rambling, undirected interactions between interviewers and respondents in which the interviewer follows no observable conventions for clarification of misunderstandings (such as repeat of question, providing definitions or guidelines, and neutral probes) and keeping the respondent focused (through feedback to emphasize the proper role of the respondent). Training interviewers to perform these tasks is essential for successful standardized interviewing (Fowler and Mangione 1990); furthermore, Billiet and Loosveldt (1988) have shown that better-trained interviewers were more likely to obtain adequate answers to questions demanding more respondent instructions, probing, and feedback. Suchman and Jordan have not provided examples of the failure of proper interviewing to yield useable survey data; they have merely provided examples that interviewing, when performed poorly, can yield unusable data, which is not the same thing.

Yet, though their specific examples are easily criticized, their general notions cannot be easily dismissed. The strict rules of GIT de-emphasize any intelligent contribution from interviewers; thus, it is possible that meaning may not be standardized even though the literal wording of the question has been properly conveyed. For example, Suchman and Jordan document one case of a respondent's confusion about what constitutes an "alcoholic beverage" (the respondent failed to treat wine as an alcoholic beverage, which seemingly contradicted previous answers). The interviewer clearly exacerbates the situation by following procedures that most survey researchers would consider inadequate. Nevertheless, it is not clear that an interviewer could have eradicated the problem using GIT techniques – at least given the questionnaire as written. Standardized interviewing did not address the problem of communicating correct meanings to the respondent. Furthermore, one could argue that even if interviewers did have adequate discretion to address ambiguities and misunderstandings, they are unable to pay attention to such matters because these tasks have lower priority than following highly detailed rules.

Theoretically, if interviewers were allowed more leeway to diagnose communication problems and clarify the meanings of questions, the goals of current survey research would be more successfully met: to standardize the *meanings* of questions so that we may infer something about the population in question. Yet, we have seen that standardization of interviews addresses important concerns that Mishler, Briggs, and Suchman and Jordan ignore, and as Schaeffer (1991, p. 368) notes, "reforms that ignore the justification for standardization run the risk of repeating old mistakes." The question then becomes: what level of standardization maximizes the usefulness of survey research? Arguments have been raised on extreme positions for and against standardization, which may seem incompatible. Is there a compromise that addresses the concerns of both sides?

6. Moving Toward Sensible Compromise

Before we can move forward, it is important to remember the needs that standardization fulfills. The benefits of practical simplicity, timeliness, and greater statistical power have cemented standardization into the core of modern surveying. To deny this is to ignore the most practical facts of the matter.

This does not mean, however, that compromise is impossible and will not occur. We are, after all, interested in reducing *total* error in surveys. If attacking the slightest interviewer deviation brings about modest reduction of interviewer error – but simultaneously causes a *greater increase* in error from the respondent, who is unable to draw on the communicative resources of an informed, intelligent interviewer – then the strategy is self-defeating. Technologically, we could eliminate interviewer error altogether by using computers to conduct interviews, either by phone or through computerized self-administered methods (U.S. Office of Management and Budget 1990). Arguably, one reason that researchers do not widely do this is because human interaction still has powerful advantages. In fact, survey researchers first used interviewers to capitalize on these advantages. The benefits necessarily go hand-in-hand with some drawbacks. It is essential for researchers desiring the "total elimination" of interviewer error to accept that this is impossible. Following precise interviewing rules remains critical to collecting the rewards standardization brings, but standardization advocates must recognize that tightly prescribed interviewer behavior does have an ultimate point of diminishing returns.

On the other hand, "interviewer liberation" offers little hope for the future of survey research. Liberating interviewers merely trades one set of problems for another – and "may mask an unwillingness to face similar problems posed by other styles of research" (Schaeffer 1991, p. 370). Briggs's (1986), for example, criticizes standardized interviews because interviewers may misunderstand norms of conversation and respondent frames of reference in different respondent groups. While researchers should be concerned about these issues, Briggs' proposed solutions return to previous problems of low reliability, high interviewer variance, inefficiency with large samples, and high cost. In addition, the fact that even slight wording changes can dramatically alter response distributions has been well documented (Schuman and Presser 1981). Given these drawbacks, the difficulties of measurement without standardization are apparent. The real question that remains is: how can researchers solve communication problems while harnessing the full benefits of standardization?

Survey questions themselves are of central importance. Successful standardized interviewing depends upon questions that can be clearly and consistently understood by respondents (Fowler and Mangione 1990). This is a crucial point, because detractors of standardization such as Suchman and Jordan use poor questions as examples of standardization's failure (Schaeffer 1991). Yet, standardization advocates open themselves to attack because "many questions ... in major surveys with large samples carried out by very professional organizations, could be readily identified as not meeting [basic] standards." This "creates [problems] difficult to solve ... in a standardized way" (Fowler and Mangione 1990, p. 137).

The importance of pretesting, which allows one to assess questionnaire communication difficulties through dozens of fresh perspectives, must therefore continue to grow. Too often, researchers regard pretests as a convenience for "when time allows" (Converse and Presser 1986). Casual attitudes toward pretesting must be continually discouraged if standardized surveys are to be successful. In addition, Fowler (1992) has explored systematic pretest methods, using them in one case to identify and correct unclear terms that were likely biasing survey estimates. Another more systematic method is cognitive interviewing, designed to help researchers understand how respondents interpret questions. The volume of research recently published in this area (Willis, Royston, and Bercini 1991; Jobe and Mingay 1991; Tanur 1992) indicates that enthusiasm for the collaboration between cognitive psychology and survey methodology continues to rise. Interestingly, the use of cognitive interviewing to improve the viability of standardized interviewing is similar to what Lazarsfeld suggested in 1944 – the use of closed questions with more open-ended techniques used for development and elaboration. More quantitative procedures such as behavior coding may also grow (at least to the extent that such labor-intense activities can be carried out in a timely manner). Some researchers who design and analyze large-scale surveys have only cursory knowledge of these procedures, so there is hope that their effect on survey quality will continue to expand.

Also, researchers need to reevaluate the sort of behavior they identify as "good interviewing" and reward interviewers for accordingly. They should keep in mind that the goal of standardized survey interviewing is to obtain complete answers to questions uniformly administered and clearly understood. Often, however, interviewer success is measured by how well the interviewer conforms to specifically prescribed behaviors (Was that an acceptable probe? Was that acceptable feedback given to the respondent?) While these are important – directive probes and inappropriate feedback can damage data quality – this approach is incomplete. It is possible that communication problems arise because interviewers spend too much time making sure that they conform to highly specific criteria (which they are rewarded for), and not enough time evaluating whether respondents understand and answer questions correctly (which they *should* be rewarded for). This weakness can be overcome *within* the constraints of standardized methodology – not by allowing much "interviewer leeway," but by avoiding overzealous enforcement of the smallest letter of the law when clear answers are obtained through generally acceptable means.

Researchers also need to spend more time teaching interviewers the concepts inherent in the questions they ask – not so that interviewers can create conversations, but so interviewers can judge the completeness of respondents' answers. Recent studies have shown that non-scripted interviewer behavior – perhaps one-sixth of interviewer activity – usually centers on clarifying inadequate answers, such as through probing (Sykes and Collins 1992). Also, questions requiring the most probing are most likely to create interviewer effects (Mangione, Fowler, and Louis 1992). These studies confirm that greater emphasis on questionnaire development, which will reduce the amount of probing necessary during interviews, is time well spent. Yet, some probing is inevitable in interactions of this sort. Researchers can do much more to ensure that non-directive probing is based on a sound understanding of survey objectives, rather than interviewers' own speculation of what constitutes a complete answer. Question-by-question specifications (Q-x-Q's) should be written consistently, and more time should be spent to ensure that interviewers understand them. Interviewers often read the Q-x-Q's for the first time when a respondent indicates that he or she does not understand a question. Standardized practice essentially teaches interviewers to follow algorithms in order to elicit responses; this practice may have led to unwarranted optimism that interviewers can do their job successfully with very little understanding of the questions they ask.

Another issue – interviewer style – has received comparatively little attention recently, but the appropriateness of personal (as opposed to formal) styles of interviewing may not be a closed issue. A study by van der Zouwen, Dijkstra, and Smit (1991) suggests that personal interviewing styles – such as allowing interviewers to act in an "understanding manner" – may generate more complete or accurate answers from respondents under some circumstances. Unfortunately, interviewers employing this style tend to use more suggestive probes and accept incomplete answers from respondents. Thus, respondent performance may improve through a more personal relationship, but extra steps must be taken to guard against interviewer error. Traditionally, maximizing the interviewer-respondent relationship and ensuring standardized measurement have been viewed as contradictory goals – as we saw in the debate surrounding the term "rapport." Yet, researchers may find choices other than "conversation," with all the unstructuredness it implies, and a completely scripted, austere interaction. Exploring the consequences of various interviewer styles, while maintaining the general rules of standardization, is an important topic for research.

These are the areas where the most energy should be spent in reducing survey communication error. Many of them take place during questionnaire design, before interviewers are involved; others involve adapting the role of the interviewer. Advocates of both standardization and non-standardization raise important points – issues that have been debated since the earliest beginnings of survey research. But it is important to note that while the issues raised by non-standardization advocates can be addressed within the context of standardization, the reverse is not true. That is, precision of measurement, timeliness, and other issues previously discussed are extremely difficult to achieve without reasonable control of variation introduced by interviewers. Thus, the most logical course of action is to address communication difficulties within the context of standardized interviewing, and to empirically investigate the point where standardization starts to substantially interfere with communication.

Perhaps reaching a consensus has been difficult because researchers with very different aims have been involved in the debate. Most criticisms of current survey practice noted here come from anthropologists or sociolinguists, interested in exploring "the interview as a communicative event" (Briggs 1986, p.2). Briggs's (1986) discussion of "acquiring metacommunicative competence," focuses on a variety of anthropological and sociolinguistic studies in a Spanish-speaking community. Most of these are non-quantitative. Mishler's (1986) chapter on the "joint construction of meaning" between interviewers and respondents uses a physician/patient interview as an example. In both cases, the emphasis is on understanding details of individual behavior, mostly in a non-quantitative manner. However, as Fowler and Mangione (1990, p.

12) point out, the purpose of standardized survey research is to measure, or produce "quantitative or mathematical descriptions of the population . . . [There is] no intrinsic interest in the answers of these individuals per se." It is not surprising that researchers of the former type, looking to standardized survey researchers for a systematic interviewing framework, may find standardized, measurement-oriented interviewing to be inadequate. This is especially true since the success of surveys has been traditionally judged by sampling designs and response rates – measures of statistical accuracy rather than communicative accuracy. And, even though standardized survey researchers have become more interested in communication issues in the last decade, this interest has emerged primarily from psychology, not anthropology or sociolinguistics. Thus, the two schools of thought may have had little basis for communication themselves.

The lack of mutual understanding of advocates of standardization and non-standardization remains considerable and troubling. Literature advocating either side is common, but there is little acknowledgment of specific points of the others' arguments. That is one reason it is valuable to examine old interviewing manuals and old methodological pieces. It is easy for both sides of the current debate to write off the other as short-sighted, or ignorant of larger issues. When we look into the past, however, we see that elements of this debate have been present for a long time; we can also see what concerns have motivated changes in research methods. These are not new problems, but the debate over them continues. We have much to learn about tangible issues, such as where standardization breaks down (Schaeffer 1991), but these are clouded by fundamental misunderstandings such as this.

7. Conclusion

While arguments of the need for conversational interviews are overstated, they are also understandable. Over the last fifty years, survey researchers' views of the ideal interaction between interviewers and respondents have changed considerably. Interviewers were initially valued for the freedom they brought, though researchers increasingly realized that freedom went hand-in-hand with interviewer-introduced error. Paralyzed by rapport for some years, researchers finally confronted interviewer error head-on as rapport faded into the background. New awareness of the benefits of standardization continued this process. The resulting attempt to eliminate interviewer error also sharply reduced the ability of interviewers to resolve communication difficulties, due to the extremeness with which these changes were attempted. Only in the last few years have researchers started to understand this deficiency and compensate for it with increased attention to questionnaire design and evaluation prior to interviewing. The proper conclusion is not that standardized interviewing is unworkable, but that it is time for a greater understanding between its advocates and opponents. The elements of this debate have been present for over 50 years – whether "various non-directive means of stimulating full discussion in the interviewing situation" (Skott 1943) are preferable to "more objective methods of research" (Lazarsfeld 1944). Once again, it is time for negotiation. Rather than standing on opposite sides of the rift, all social researchers need to appreciate the legitimate concerns addressed by those on the other side, and work together toward fashioning strategies that will maximize the usefulness of survey research into the next century.

8. References

Billiet, J. and Loosveldt, G. (1988). Improvement of the Quality of Responses to Factual Survey Questions by Interviewer Training. Public Opinion Quarterly, 52, 190–211.

Briggs, C.L. (1986). Learning How to Ask: A Sociolinguistic Appraisal of the Role of the Interviewer in Social Science Research. Cambridge, MA: Cambridge University.

Cannell, C.F., Marquis, K.H., and Laurent, A. (1977). A Summary of Studies of Interviewing Methodology. Vital and Health Statistics, DHEW Publication no. (HRA) 77-1343, ser. 2, no. 69. Washington, D.C.: U.S. Government Printing Office.

Converse, J.M. (1984). Strong Arguments and Weak Evidence: The Open/Closed Questioning Controversy of the 1940s. Public Opinion Quarterly, 48, 267–282.

Converse, J.M. (1987). Survey Research in the United States: Roots and Emergence 1890–1960. Berkeley, CA: University of California.

Converse, J.M. and Presser, S. (1986). Survey Questions: Handcrafting the Standardized Questionnaire. Newbury Park, CA: Sage.

Fowler, F.J. and Mangione, T.W. (1990). Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Newbury Park, CA: Sage.

Fowler, F.J. (1992). How Unclear Terms Affect Survey Data. Public Opinion Quarterly, 56, 218–231.

Gordon, R.L. (1969). Interviewing: Strategy, Techniques, and Tactics. Homewood, IL: Dorsey. Cited in Goudy, W.J. and Potter, H.R. (1975). Interviewer Rapport: Demise of a Concept. Public Opinion Quarterly, 39, 530–543.

Goudy, W.J. and Potter, H.R. (1975). Interviewer Rapport: Demise of a Concept. Public Opinion Quarterly, 39, 530–543.

Groves, R.M. (1989). Survey Errors and Survey Costs. New York: John Wiley.

Guenzel, P.J., Berckmans, T.R., and Cannell, C.F. (1983). General Interviewing Techniques. Ann Arbor, MI: The University of Michigan.

Henson, R., Cannell, C.F., and Lawson, S. (1976). Effects of Interviewer Style on Quality of Reporting in a Survey Interview. The Journal of Psychology, 93, 221–227.

Hyman, H.H. (1954). Interviewing in Social Research. Chicago: University of Chicago Press.

Hyman, H.H. (1991). Taking Society's Measure: A Personal History of Survey Research. New York: Russell Sage.

Jobe, J.B. and Mingay, D.J. (1991). Cognition and Survey Measurement: History and Overview. Applied Cognitive Psychology, 5, 175–192.

Kahn, R. and Cannell, C.F. (1957). The Dynamics of Interviewing. New York: John Wiley.

Lazarsfeld, P.F. (1944). The Controversy Over Detailed Interviews – An Offer for Negotiation. Public Opinion Quarterly, 8, 38–60.

Likert, R. (1947). The Sample Interview Survey: A Fundamental Research Tool of the Social Sciences. In Current Trends in Psychology, ed. W. Dennis, Pittsburgh, PA: University of Pittsburgh.

Mangione, T.W., Fowler, F.J., and Louis, T.A. (1992). Question Characteristics and Interviewer Effects. Journal of Official Statistics, 8, 293–307.

Mishler, E.G. (1986). Research Interviewing. Cambridge, MA: Harvard University Press.

Schaeffer, N.C. (1991). Conversation with a Purpose – or Conversation? Interaction in the Standardized Interview. In Measurement Errors in Surveys, eds. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, S. Sudman, New York: John Wiley.

Schuman, H. and Presser, S. (1981). Questions and Answers in Attitude Surveys. San Diego: Academic Press.

Skott, H.E. (1943). Attitude Research in the Department of Agriculture. Public Opinion Quarterly, 7, 280–292. Cited In Lazarsfeld, P.F. (1944). The Controversy Over Detailed Interviews – An Offer for Negotiation. Public Opinion Quarterly, 8, 38–60.

Stouffer, S.A., Suchman, E.A., DeVinney, L.C., Star, S.A., and Williams, R.M. (1949). The American Soldier, Volume I: Adjustment During Army Life. Princeton, NJ: Princeton University.

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

Survey Research Center, The University of Michigan (1954). Manual for Interviewers. Ann Arbor, MI: University of Michigan.

Sykes, W. and Collins, M. (1992). Anatomy of the Survey Interview. Journal of Official Statistics, 8, 277–291.

Tanur, J.T., ed. (1992). Questions About Questions: Inquiries into the Cognitive Bases of Surveys. New York: Russell Sage.

U.S. Office of Management and Budget (1990). Computer-Assisted Survey Information Collection. Statistical Policy Working Paper 19. Springfield, VA: National Technical Information Service.

van der Zouwen, J., Dijkstra, W., and Smit, J. (1991). Studying Respondent-Interviewer Interaction: The Relationship Between Interviewing Style, Interviewer Behavior, and Response Behavior. In Measurement Errors in Surveys, ed. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, S. Sudman, New York: John Wiley.

Weiss, C.H. (1968). Validity of Welfare Mothers' Interview Responses. Public Opinion Quarterly, 32, 622–633.

Willis, G.B., Royston, P., and Bercini, D. (1991). The Use of Verbal Report Methods in the Development and Testing of Survey Questionnaires. Applied Cognitive Psychology, 5, 1–17.

Received September 1993
Revised December 1994