Cultural Variability in Event History Calendar and Conventional Questionnaire Interviews: A Verbal Behavior Analysis

Mario Callegaro, Robert F. Belli, Emilio Serrano[1], David R. Palmer[2]
University of Nebraska, Lincoln. Program in Survey Research and Methodology[1]
MarketRx Inc.[2]

Keywords: Behavioral Coding, Event History Calendar, Standardized Interviewing, Interview Conversational Rapport, Interview Cognitive Difficulty.

1. Introduction

Behavioral coding has been generally used to assess interviewer performance, to find cognitive problems with questionnaires, and as a tool to assess the quality of the data collection in interviewer administered surveys. In this study we apply a verbal behavior coding scheme to study cross cultural variability between African American and European American respondents in a split ballot experiment where half of the respondents were assigned to a standardized conventional questionnaire (CQ) interviewing methodology and the other half to a more flexible Event History Calendar (EHC) interviewing methodology. We use race/ethnicity as a proxy for respondent culture to study two groups of verbal behaviors that were named "cognitive difficulty" and "conversational rapport" in both interviewing conditions. The previous findings on cultural variability in response comprehension, interpretation, in response styles and in conversational rapport all come from conventional questionnaire data. In this paper we want to test the performance of Event History Calendar in a culturally diverse population focusing on the concepts of cognitive difficulty and conversational rapport.

2. Cultural Variability in Cognitive Difficulty and Conversational Rapport in Survey Interviews

Cognitive difficulty and conversational rapport has been studied in the context of survey interviews as an indicator of data quality. Cognitive interviews, behavioral coding, response latencies and experiments are usually employed to determine cognitive problems in survey interviews.

In the context of behavioral coding, Fowler and Cannell (1996) offer a detailed summary of the common problems affecting the respondent behavior. For example, respondents' requests for clarification are an indication of unclear or unfamiliar terms, unclear respondent task, and poor question order. When respondents give inadequate answers they are indicating an unclear meaning of the question, unclear response task or a task that is difficult or does not fit the answers. Belli, Lepkowski and Kabeto (2001), and Belli and Lepkowski (1996) studied the relationship between cognitive difficulty and response quality. They found a direct relationship between cognitive difficulty and the accuracy of retrospective reports for health reports measured with validation data from hospital records. Cognitive difficulty was measured with interviewer verbal behaviors such as the exact repeating of a question and probing, and with respondent behaviors such as seeking clarification, giving an inadequate answer, and correcting a previous answer.

Cognitive difficulty has also been studied in the field of cross cultural research with similar concepts of comprehension and interpretation of survey questions. With race/ethnicity serving as an indicator of culture, group differences in the comprehension and interpretation of conventional questionnaires between African Americans and Whites have been identified with African Americans showing higher scores of comprehension difficulties than Whites (Holbrook, Cho, & Johnson, forthcoming; Johnson et al., 2006; Johnson et al., 1997). The authors explain these differences using the acculturation concept. Questions are written from the prospective of the dominant culture (white) (Holbrook et al., forthcoming) and for this reason members of minority groups [in this case African Americans -- but the findings are in the same direction for Latinos as well (Johnson et al, 2006)] show some comprehension difficulties. Acculturation has also been associated with acquiescence and extreme response styles of African American respondents (Bachman & O'Malley, 1984a, 1984b; Clarke, 2000).

The debate surrounding conversational rapport hinges on the broader topic of standardized interviewing (Beatty, 1995) and since the beginning of survey research methodology the concept has been defined in different ways (Goudy & Potter, 1976). Authors such as Henson, Cannel and Lawson (1976) and Dijkstra (1987) define rapport in the context of interviewing style, contrasting a formal style to a

more "interpersonal" style. Authors starting from Hill and Hall (1963), Goudy and Potter (1976), and Davis and Silver (2003) for example, developed rapport scales that were compiled by the interviewers with items such as how often did the respondent and the interviewer feel at ease, how favorable was the respondent to the interview, or how cooperative and interested was the respondent (Davis and Silver, 2003). Other researchers define rapport with a single behavior, for example laughter (Lavin & Maynard, 2001).

The results from the conversational rapport studies are conflicting and they depend on the different definitions just mentioned. Hill and Hall (1963) found that an increase of rapport corresponds to decrease in item nonresponse. Dijkstra (1987) found how rapport increases respondent motivation and the quality of the responses while Henson and colleagues (1976) did not observe differences in accuracy of answers regarding automotive accidents based on validation data. Belli, Lepkowski and Kabeto (2001) found that the verbal expression indicating conversational rapport were not associated with respondent accuracy, but Davis and Silver (2003) observed that respondents with high level of rapport reported more correct answers to political knowledge questions than respondents with lower level of rapport.

All the findings reviewed so far come from interviews collected with conventional questionnaires. In a recent study Belli, Lee, Stafford and Chou (2004) analyzed verbal behaviors in both CQ and EHC interviews. The data come from an elaborate coding scheme that after a factor analysis converged to 4 factors, one of them representing cognitive difficulty and another representing conversational rapport. The behaviors that loaded on the two factors are similar to those found in a previous study by Belli, Lepkowski and Kabeto (2001). The more recent results indicate that cognitive difficulty is correlated with poorer quality of retrospective reports in more strongly in CQ interviews than in EHC interviews. Similar patterns were observed with rapport behaviors – they were more detrimental to retrospective reports in CQ interviews in comparison to EHC ones. The authors offer the explanation that the conversational, flexible style of EHC interviews can offset the potentially deleterious consequences of cognitive difficulty and rapport.

In one pioneering study relevant for our research, Marquis (1974) explored the effects of the respondent's race on the interviewer's behavior. Using a very detailed behavioral coding scheme, he looked at the correlation between interviewer behaviors and respondent behaviors. More specifically, two indexes of rapport were constructed a priori, one for the interviewers and one for the respondents. The correla-

tion between the two indexes was $\underline{r}$ =.42. When analyzing the four combinations of interviewer and respondent race (White and African American) the mean respondent rapport behavior index score was higher when the race of interviewer and respondent matched than for the mismatch. Davis and Silver (2003) found no statistically significant differences between the level of rapport rated by the interviewers by the race of the respondent. Rapport was however measured differently than the previous study as discussed previously.

3. Methods

The data were collected via telephone during a 6 week span from May through June 1998 with a random nationwide subset of the regular 1997 Panel Study of Income Dynamics (PSID). The respondents and 20 interviewers were randomly assigned to the CQ condition (N=307, 84.1% AAPOR Cooperation Rate 1) or to the EHC condition (N=309, 84.45% AAPOR Cooperation Rate 1). In the CQ condition, the interview was conducted using a 25 page paper and pencil questionnaire with standardized questions replicated from the regular 1987 PSID study. Topics were places of residence, household composition, jobs held, earned income from employment, entitlement income from ADC/AFDC, and food stamps. Years working, weeks unemployed, weeks out of the labor force, and weeks missing work because of vacation, personal illness or illness of another, were also measured. A parallel EHC instrument was developed to measure the same variables as the CQ survey. A 18x28-inch paper and pencil calendar contained seven domains of inquiry or time lines (landmarks, residence, household composition, employment, no employer, time away from work, and entitlements). Interviewers were specifically trained to use the calendar and were given scripted questions and probes to introduce each domain. The length of the two interviews did not significantly differ between conditions.

Both CQ and EHC interviewers were trained for 15 hours over the course of three days with an instruction program focused on maximizing reporting accuracy that was appropriate for each condition. This is important information that the reader should keep in mind when evaluating the verbal behaviors results. Interviewers were trained to use conversational probing (Schober & Conrad, 2002). For example, if a respondent is unsure how to answer a particular question, the interviewers are instructed not to say "whatever it means to you" but instead to refer to the definitions provided in the "Question by question objective manual" that accompanies the training. The

complete details of the data collection have been provided by Belli, Shay, and Stafford (2001).

With the permission of the respondents, 95% of the interviews were tape recorded and transcribed. A coding scheme was developed resulting in 56 verbal behaviors. Because of refusals, poor transcriptions (Poland, 2003), inaudible tapes and attrition of the coders, 217 of the 309 EHC interviews (70.2%), and 161 of the 307 CQ interviews (52.4%) were usable in this study of verbal behaviors. Of these 378 interviews, 38 (10.1%) interviews (17 EHC and 21 CQ) were independently coded by both coders (blindly) in order to measure the reliability of the coding scheme. The inter coder reliability was judged to be adequate. Forty one of the 56 behaviors had a Pearson $r \geq .40$, many of them with $r \geq .60$. Within each domain (e.g., residence, employment, etc.) codes were assigned in the order of their occurrence[1]. The behaviors are organized in two classes: interviewers' verbal behaviors and respondents' verbal behaviors. The full list behavioral coding scheme is described in details by Belli, Lee, Stafford & Chou (2004).

The final dataset used for the analysis contained 207 European American respondents, 149 African American respondents and 22 respondents who classify themselves as other race. In the EHC condition we had 7 European American interviewers and 2 African Americans, in the CQ we had 8 European American interviewers and 2 African Americans. Table I summarizes the number of interviews by race of interviewer and race of the respondent. For this analysis we dropped the 22 respondents who classified themselves as another race.

The race of the interviewer was obtained by the interviewer's background information from the data collection facility. Since race of the interviewer will be used in the analysis and the questionnaire was administered by telephone, it is important to point out that respondents correctly determine the race of the interviewer on average 75% of the time[2] (Callegaro, De Keulenaer, Krosnick, & Daves, 2006). In the current experiment the respondents were not asked to determine the race of the interviewer. For this reason, taking the actual race of the interviewer instead of the perceived race of the interviewer introduces some degree of error in the analysis and interpretation of the results (Davis, 1997; Wilson, 2006; Wilson & Olesen, 2002).

The initial behavioral coding scheme was developed to include behaviors that have been identified in previous research as being relevant to standardized

interviews (Fowler & Cannell, 1996; Oksenberg, Cannel, & Kalton, 1991). One example of such behaviors is *significant change in question wording*. In addition, the scheme was supplemented with codes that can capture retrieval and conversational behaviors potentially important for both standardized interviewing and EHC (Belli, Lepkowski, & Kabeto, 2001; Houtkoop-Steenstra, 2000). Examples of these behaviors are *interviewer and respondent digression* and *respondent sequential retrieval*. The initial coding scheme was improved and refined after weekly meetings and listening sessions among the research group and the coders. The coders were trained in two steps. During the first step, which lasted ten weeks, the coder became familiar with the coding scheme. In the second step, which lasted twelve weeks, the coders independently coded 38 randomly selected transcripts. On average each transcript required 4-5 hours of coding. The complete details of the behavioral coding scheme and analysis have been provided by Belli and colleagues (2004).

3.1. Variable definitions and statistical analysis

The first step of the data analysis was to compute the mean behavior per domain for each respondent. This was done for each behavior by summing the behavior counts across domains and then dividing that total by the applicable number of domains. Afterward, to reduce the large number of behaviors to a more manageable number the factor analysis results described in Belli et al. (2004) were used. Specifically, the 41 behaviors were reduced to four factors: retrieval cues, detailed interviewing, cognitive difficulty and conversational rapport.

As mentioned earlier, only cognitive difficulty and conversational rapport are treated in this paper. Table II shows the verbal behaviors that loaded on the factors cognitive difficulty and conversational rapport. The factor scores were computed "as a sum of the standard scores for each of the behaviors that loaded on the factor" (Belli et al. 2004, p. 207).

In the factor analysis, the factor variables are assumed to be orthogonal, and thus each can be modeled separately without a Type I error correction. A constant was then added to each factor score[3] so that a square root transformation could be performed; thereby making the factor score distributions more like the normal distribution. As previously stated, the factors were modeled with separate regressions, yet the same independent variables were included in each of the models. The effects (main effects and interaction effects) of respondent race, interviewer race, and

---

[1] Order was however not analyzed.
[2] The percentage is referred to an average within studies reviewed by the authors and it refers to the correct prediction of African American and White races only.

[3] To bring the smallest value to at least zero.

questionnaire condition on each factor were the focus of the study.

In addition, a number of other control variables were also included in the model: respondent education (in years), respondent and interviewer gender, respondent and interviewer age, and interview complexity. Interview complexity was measured with the following variables:

- Receipt of aid to dependent children (ADC) in 1996 and 1997
- Receipt of food stamps in 1996 and 1997
- Jobs during 1996 and 1997
- Weeks being unemployed or temporarily laid off in 1996, 1997
- Weeks away from work due to illness (respondent or someone else), vacation, or strike in 1996, 1997
- Total wages, bonuses and professional practice income from experimental reports of 1996 and 1997 incomes
- Interviewing minutes during 1998 experimental interviews, excluding time spent on landmarks domain
- Number of respondent moves between 1996 and 1997 interview dates
- Number of family members moved in or out of the household

Forty-five cases had missing values on at most 3 of the control variables. In total, 78 values were imputed, representing roughly 1% of all independent variable values used in the model[4].

4. Results

The results of the regression on the cognitive difficulty factor indicate that there is no interaction between race and interview type. None of the race interactions were significant. In fact, the only significant variable of interest was respondent race with a p-value less than .01. In other words, the EHC interviewing style does not seem to be more cognitively taxing than the CQ, an important finding for a new methodology such as the EHC.

Table III shows that African American respondents had an adjusted cognitive difficulty factor score mean of 3.11 whereas the adjusted factor score mean for European Americans was 2.77. These means are significantly different from each other with a p-value of .008. Though these factor scores are difficult to interpret in terms of raw frequency of behaviors, they

do make it easy to see that European Americans were less prone than African Americans to demonstrate verbal behaviors indicative of cognitive difficulty.

Two concurrent explanations can account for the above differences in means: a true difference in the amount of cognitive difficulty actually experienced, and different communication or response styles. In the first case African Americans, as a minority group, show more cognitive difficulty in answering questions because they are written from predominant culture point of view. In this sense our results concur with the studies done by the researchers of the University of Illinois at Chicago (Holbrook, Cho, & Johnson, forthcoming; Johnson, Holbrook, & Cho, 2006; Johnson et al., 1997). Our results also extend the findings to the Event History Calendar methodology. The second concurrent explanation is the realm of different communication and response style. In fact, the behaviors that were labeled cognitive difficulty can be also seen as behaviors that deviate from standardized interviewing (see Table II). Race differences in response styles have been documented as early as 1984 (Bachman & O'Malley, 1984a, 1984b; Johnson & Bowman, 2003; Warnecke et al., 1997). There is some evidence that African Americans are likely to have more spontaneous, flexible, and open minded communications than European Americans who were found to be more regulated and structured (Shade, 1986). According to these findings a higher factor mean for African American respondents can be indicating a different communication style[5].

A significant three-way interaction of respondent race, interviewer race and questionnaire condition is observed when looking at conversational rapport. To interpret a complex 3-way interaction such as this it may be useful to look at the data in smaller pieces before combining it all together. In Table IV the adjusted means for rapport are broken down by interviewer and respondent race. The factor means for EHC fall in a small range between 2.4 and 2.68. The one exception (3.4) occurs when an African American respondent is paired with an African American interviewer. The higher factor mean indicates that rapport behaviors in the matched race with African Americans occurred more frequently than in any other case. When the respondent race and interviewer race are matched there tends to be more rapport building behavior (means = 2.68 and 3.4 for EHC) than when they are not (means = 2.65 and 2.4).

Just as in the EHC condition, when the respondent race and interviewer race are mismatched it doesn't make much difference in the frequency of rapport building behaviors if the interviewer is African American and respondent is European American

---

[4] These values were imputed using the EM algorithm employed by "proc mi" in SAS

[5] The results are controlled for respondent's level of education.

or visa versa. Unlike the EHC condition however, when the race of respondent and interviewer are matched they are not necessarily higher than the mismatched case. Indeed, in the CQ when European Americans are matched the amount of rapport related behaviors increases (mean = 2.84), whereas for African Americans it decreases substantially (mean = 1.58). For any particular case the conversational rapport factor means for EHC and CQ are approximately equal with one glaring exception. When an African American interviews an African American respondent the difference in conversational rapport behaviors between the EHC and CQ are dramatic. Specifically, in EHC the conversational rapport factor mean jumps to 3.4, the highest of the rapport factor means; in CQ the mean drops to 1.58, the lowest of rapport factor means.

It is not surprising that in the EHC condition the overall mean of conversational rapport is higher than for the CQ. This is due to the more flexible interviewing style of the EHC methodology. What is interesting is how conversational rapport is the highest when matching African American interviewer and African American respondent. In other words, when the interviewing style is less constrained, the amount of conversational rapport between African American interviewers and respondents is unleashed.

These findings can be linked to the previous findings of Table III with the intervening variable, race of interviewer. According to the social distance model, the more social distance is perceived by the respondent, the more interviewer effect is found (Williams, 1968). Looking at the social distance model a different way, it follows that a reduction of social distance, achieved by matching the characteristics of the interviewer with the characteristics of the respondent, should lead to small or non-existent interviewer effect. This idea is suggested by some scholars such as Schaeffer (1980) and Wolford et al (1995). The social distance model should work symmetrically, i.e. the magnitude of the effect should be the same in African American – European American respondent pair interviewer, as in the opposite combination. Nevertheless some studies found asymmetric effects, with European Americans being less sensitive to the effect (Davis & Silver, 2003; Hyman, Coob, Feldman, Hart, & Stember, 1954; Krysan & Couper, 2003). In our case, the conversational rapport can be a proxy measure of social distance meaning that African American respondents feels more free to talk to an African American interviewer and when they are allowed to do so, like in the EHC condition, the conversational rapport behaviors score higher. This should be the case for European American respondent as well. The higher mean for the interviewer-respondent race matching (European American – European Ameri-

can) do not however differ from the non matching mean at a statistically significant level. This can be further evidence that European Americans respondents are less sensitive to social distance as showed by some of the previous mentioned studies.

5. Discussion

The goal of this paper was not to measure the association of interviewing techniques and data quality but to explore how the two interviewing methodologies interact with different cultures using the race/ethnicity as a proxy for respondent culture. We concentrated our analysis on two groups of verbal behaviors named cognitive difficulty and conversational rapport. To our knowledge there are no studies that attempt to measure race of interviewer effect and race differences in response style for the Event History Calendar. We found higher scores for the cognitive difficulty factor for African American respondents when compared to European American respondents. This higher score in both conditions (CQ and EHC) indicates more difficulty in comprehending the questions asked during the interview and also a response style of African Americans. The same behaviors are however an indication of a less standardized interviewing communication between respondent and interviewer. Our explanation is supported by findings indicating a less structured communication style of African Americans (Shade, 1986) in comparison to European Americans.

When focusing on "conversational rapport" we notice higher scores in the EHC condition. This finding is not surprising and derives from the more flexible style of interviewing that is allowed in EHC condition. We want, however, to point out that the EHC condition is still a structured communication style and that the interviewers are not allowed to behave in a manner that can in any way influence the respondent's answers. The matching of African American interviewers and respondents elicited the higher score of conversational rapport factor for the EHC and the lowest for the CQ condition. We explained this results by referring to the social distance model (Williams, 1968). African American respondents feel less constrained to express themselves African American interviewers and this tendency is accentuated in the EHC because of its less contained and more flexible communication style.

There are two limitations of our study. The first one deals with the problematic nature of verbal behavior data. Utterances are not always easy to classify and the entire process of coding is very demanding and prone to errors. Second, our sample size of African Americans was obtained from national sample. Unless oversampling or another procedure is

applied, African Americans turn up proportionally. In this case, no such sampling scheme was employed and the final N is relatively small[6] -- which yields a power smaller than one might like (Smith, 1993).

6. Conclusions

Our results point in the direction of different communicative styles between European Americans and African Americans in the survey interview, and in the direction of race of interviewer effects, specifically towards the social distance explanation. The conversational style of the EHC interview, however, unleashes the amount of rapport between an African American interviewer and an African American respondent. Although the literature on the relationship between rapport and survey quality provides mixed results, we lean to support the idea that an increase of rapport increases respondent motivation, decreases anxiety, and ultimately increases the quality of the responses.

Finally, panel respondents are trained to be good respondents, and due to survey learning may tend to minimize behaviors that were under investigation (e.g. digressions and request for clarification). For this reason we believe our results would be stronger in the general population where the amount of behaviors indicating cognitive difficulty and conversational rapport is expected to be higher.

References

Bachman, J. G., & O'Malley, P. M. (1984a). Black-White differences in self-esteem: Are they affected by response styles? *American Journal of Sociology, 90*, 624-639.

Bachman, J. G., & O'Malley, P. M. (1984b). Yea-saying, nay-saying, and going to extremes: Black-White differences in response styles. *Public Opinion Quarterly, 48*, 491-509.

Beatty, P. (1995). Understanding the standardized/non-standardized interviewing controversy. *Journal of Official Statistics, 11*, 147-160.

Belli, R. F., Lee, E. H., Stafford, F. P., & Chou, C.-H. (2004). Calendar and question list survey methods: Association between interviewers behaviors and data quality. *Journal of Official Statistics, 20*, 185-218.

Belli, R. F., Lepkowski, J., M., & Kabeto, M. U. (2001). The respective roles of cognitive processing difficulty and conversational rapport on the accuracy of retrospective reports of doctor's office visits. In M. Cynamon, L. & R. A. Kulka (Eds.), *Seventh conference on health survey research methods. DHHS Publication No. (PHS) 01-1013* (pp. 197-203). Hyattsville, MD: U.S. Government Printing Office.

Belli, R. F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly, 65*, 45-74.

Callegaro, M., De Keulenaer, F., Krosnick, J. A., & Daves, R. (2006). Interviewer effect in a RDD telephone preelection poll in Minneapolis 2001. An analysis of race and gender of the interviewer effect. In *2005 Proceedings of the American Statistical Association, 60th Annual conference of the American Association for Public Opinion Research [CD-ROM]* (pp. 3815-3821). Alexandria, VA: American Statistical Association.

Clarke, I. I. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior and Personality, 15*, 137-152.

Davis, D. W. (1997). Nonrandom measurement error and race of the interviewer effects among African Americans. *Public Opinion Quarterly, 61*, 183-207.

Davis, D. W., & Silver, B. D. (2003). Stereotype threat and race of interviewer effects in a survey on political knowledge. *American Journal of Political Science, 47*, 33-45.

Dijkstra, W. (1987). Interviewing style and respondent behavior. An experimental study of the survey-interview. *Sociological Methods and Research, 16*, 309-334.

Fowler, F. J., & Cannell, C. F. (1996). Using behavioral coding to identify cognitive problems with survey questions. In N. Schwarz & S. Sudman (Eds.), *Answering questions. Methodology for determining cognitive and communicative processes in survey research* (pp. 15-36). San Francisco: Jossey-Bass.

Goudy, W. J., & Potter, H. R. (1976). Interview rapport: Demise of a concept. *Public Opinion Quarterly, 39*, 529-543.

Henson, R., Cannell, C. F., & Lawson, S. (1976). Effects of interviewing style on quality of reporting in a survey interview. *The Journal of Psychology, 93*, 221-227.

Hill, R. J., & Hall, N. E. (1963). A note on rapport and the quality of interview data. *Southwestern Social Science Quarterly, 44*, 247-255.

---

[6] Our Table I is an example of the above concept.

Holbrook, A. L., Cho, Y. I., & Johnson, T. P. (forthcoming). The impact of question and respondent characteristics on comprehension and mapping difficulties. *Public Opinion Quarterly*.

Houtkoop-Steenstra, H. (2000). *Interaction and the standardized survey interview. The living questionnaire*. Cambridge: Cambridge University Press.

Hyman, H. H., Coob, W. J., Feldman, J. J., Hart, C. W., & Stember, C. H. (1954). *Interviewing in social research*. Chicago: University of Chicago Press.

Johnson, T. P., & Bowman, P. J. (2003). Cross-cultural sources of measurement error in substance use surveys. *Substance Use & Misuse, 30*, 1447-1490.

Johnson, T. P., Cho, Y. I., Holbrook, A. L., O'Rourke, D. P., Warnecke, R. B., & Chavez, N. (2006). Cultural variability in the effects of questions designs features on respondent comprehension o health surveys. *Annals of Epidemiology, 16*, 661-668.

Johnson, T. P., Holbrook, A. L., & Cho, Y. I. (2006, August 6-10). *The effects of acculturation on survey question comprehension among Latino respondents in the U.S.* Paper presented at the Joint Statistical Meeting, Seattle, WA.

Johnson, T. P., O'Rourke, D. P., Chavez, N., Sudman, S., Warnecke, R. B., Lacey, L., et al. (1997). Social cognition and responses to survey questions among culturally diverse populations. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 87-113). New York: Wiley.

Krysan, M., & Couper, M. P. (2003). Race in the live and the virtual interview: Racial deference, social desirability, and activation effects in attitude surveys. *Social Psychology Quarterly, 66*, 364-383.

Lavin, D., & Maynard, D. W. (2001). Standardization vs. rapport: Respondent laughter and interviewer reaction during telephone interviews. *American Sociological Review, 66*, 453-479.

Marquis, K. H. (1974). Effect of race, residence, and selection of respondent on the conduct of the interview. In J. B. Lansing, S. B. Withey & A. C. Wolfe (Eds.), *Working papers on survey research in poverty areas* (pp. 292-318). Ann Arbor, Mi: Survey Research Center, Institute of Survey Research, University of Michigan, Ann Arbor.

Oksenberg, L., Cannel, C., F. , & Kalton, G. (1991). New strategies for pretesting survey questions. *Journal of Official Statistics, 7*, 349-365.

Poland, B. D. (2003). Transcription quality. In J. A. Holstein & J. F. Gubrium (Eds.), *Inside interviewing: New lenses, new concerns* (pp. 267-287). Thousand Oaks, CA: Sage.

Schaeffer, N. C. (1980). Evaluating race-of-interviewer effects in a national survey. *Sociological Methods and Research, 8*, 400-419.

Schober, M. F., & Conrad, F. G. (2002). A collaborative view of standardized survey interview. In D. W. Maynard, H. Houtkoop-Steenstra, N. C. Schaeffer & J. van der Zouwen (Eds.), *Standardization and tacit knowledge. Interaction and practice in the survey interviews* (pp. 67-94). New York: Wiley.

Shade, B. J. (1986). Is there an Afro-American cognitive style? An exploratory study. *Journal of Black Psychology, 13*, 13-16.

Smith, W. A. (1993). Survey research on African Americans. Methodological innovations. In J. H. Stanfield II & R. M. Dennis (Eds.), *Race and ethnicity in research methods* (pp. 217-229). Newbury Park, CA: Sage.

Warnecke, R. B., Johnson, T. P., Chavez, N., Sudman, S., O'Rourke, D. P., Lacey, L., et al. (1997). Improving question wording in surveys of culturally diverse populations. *Annals of Epidemiology, 7*, 334-342.

Williams, A. J., Jr. (1968). Interviewer role performance: A further note on bias in the information interview. *Public Opinion Quarterly, 32*, 287-294.

Wilson, D. C. (2006). Conversations with a stranger: Race of interviewer perceptibility. In AAPOR (Ed.), *Proceedings of the 61st annual conference of the American Association for Public Opinion Research [CD-ROM]*. Lenexa, KS: American Association for Public Opinion Research.

Wilson, D. C., & Olesen, E. P. (2002, May 17-19). *Perceived race of interviewer effects in telephone interviews*. Paper presented at the 57th AAPOR/WAPOR conference, St Pete Beach, FL.

Wolford, M. L., Brown, R. E., Marsden, A., Jackson, J. S., & Harrison, C. (1995). Bias in telephone surveys of African Americans. The impact of perceived race of interviewer on responses. *Proceedings of the Survey Research Methods Section, American Statistical Association*, 800-804.

Table I. Number of interviews by race of the interviewer and race of the respondent

	Interviewer's race	European American	African American	Total
Respondent's race				
European American		162	45	207
African American		122	27	149
Total		284	72	356

Table II. Behaviors that form the factors "cognitive difficulty" and "rapport"

Factors	Cognitive difficulty	Conversational rapport
Interviewer behaviors	Interviewer seek clarification	Unacceptable feedback
	Interviewer clarifies	Interviewer digression
	Acceptable feedback	Interviewer laughs
Respondent behaviors	Request for clarification	Respondent digression
	Answer does not meet question objective	Respondent laughs

From Table 4 in Belli et al, 2004, pp. 205-6.

Table III. Cognitive difficulty factor mean by respondent race

Respondent race	Factor mean
European American	2.77 (.20)
African American	3.11 (.21)**

Note. ** *p.* < 0.05. Standard errors in parentheses

Table IV. Conversational rapport factor means by condition

Interviewer race	African American	European American	African American	European American	Overall
Respondent race	European American	European American	African American	African American	Mean
EHC	2.65 (.33)	2.68 (.26)	3.40 (.38)	2.40 (.25)	2.78 (.25)
CQ	2.30 (.31)	2.84 (.27)	1.58 (.37)***	2.23 (.28)	2.24 (.24)***

Note: test of difference between conditions, ** *p.* < 0.05, ****p.* < 0.01. Standard errors in parentheses