"WHATEVER IT MEANS TO YOU": ETHNICITY, LANGUAGE, AND THE SURVEY RESPONSE IN TELEPHONE-ADMINISTERED HEALTH SURVEYS OF AFRICAN AMERICANS

by

Rachel Ellen Davis

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Health Behavior and Health Education)
in The University of Michigan
2008

Doctoral Committee:

Professor Kenneth Resnicow, Chair
Professor Nancy K. Janz
Associate Professor Cleopatra Howard Caldwell
Research Professor Michael Couper

© Rachel Ellen Davis
2008

ACKNOWLEDGEMENTS

This dissertation would not have been accomplished without the support and guiding wisdom of my committee. In particular, I would like to thank Ken Resnicow for his willingness to support this research, both intellectually and financially, and for helping me to navigate around the many bumps in the road. I am also indebted to Mick Couper, who provided key insight from a survey methodology perspective and always had an open office door. I am grateful to Nancy Janz for sharing her strong research skills and acumen about the dissertation process. And, last but certainly not least, I am obliged to Cleopatra Caldwell, who has challenged me to see things from different perspectives and to grow in new ways.

I would also like to acknowledge the wonderful support I received from the *Eat for Life* study team. I am very grateful to Ken Resnicow and the rest of the team for allowing me to add my own research activities to the project and for access to *Eat for Life* study data. Despite some concern about the sensitivity of exploring race and ethnicity, the members of the study team provided support and feedback throughout the process. I feel very lucky to have spent five years in such excellent company.

This work would also not have been possible without the expert consultation of Jen Nguyen. Thank you for sharing your knowledge of linguistics and for contributing invaluable ideas to the research design.

This research was funded by support from the Charles Cannell Fund, the Robert Wood Johnson Foundation Health and Society Scholars Program Small Grant Program, the University of Michigan Center for Health Communications Research, Rackham Graduate School, the University of Michigan Department of Health Behavior and Health Education, and the National Science Foundation's Directorate for Social, Behavioral & Economic Sciences, Methodology, Measurement & Statistics Program. I deeply appreciate this financial support, without which this project would not have been possible. I would also like to express my gratitude for the Richard Janz Memorial

Award, which funded my tuition and enabled me to focus on this project and finish in a timely manner.

I am indebted to all of the other people who made this project happen – most especially, the study participants and linguistic coders. Their willingness to participate in this research and to code these data are both deeply appreciated.

I want to acknowledge the sources of inspiration for this work. The uniquely interdisciplinary culture of the University of Michigan inspired me to think beyond my pre-existing boundaries and to bravely explore new disciplines. This experience has taught me that there are many ways from which to view a common issue and that more heads are truly better than one. I am also deeply grateful to the examples set by the many, hard-working survey and public health professionals who aspire to improve our methods and create a better, more healthful life for all. But, above all, it is the people who remain underserved by our efforts whom have most inspired my attempts. I have been blessed by opportunities that I know many may not have shared. I hope that I am able to use the skills and experience that I have acquired through this dissertation in ways that will touch lives, open doors, and foster the health and well-being that we all so much deserve.

Lastly, I am also extremely lucky to have so many wonderful people in my personal life who supported me through this process. My friends and family have heard more than they ever wished to know about measurement error, yet they tirelessly contributed intellectual feedback and emotional support. I have learned so much from you all. In particular, I want to thank my husband, Shannon, without whom I would have been lost. You and Riley bring joy to my every day. I am extremely excited about proceeding into a dissertation-free life with you two at my side. Here's to many new and, for you, more enjoyable adventures!

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ……..…**ii**

LIST OF TABLES ……**vi**

ABSTRACT ……..**vii**

CHAPTER 1 – INTRODUCTION …….**1**

CHAPTER 2 – THE IMPORTANCE OF INTERVIEWER ERROR IN PUBLIC HEALTH SURVEYS ……...**5**

Introduction ……....**5**

What is Interviewer Error? ……..**6**

Why Should Public Health Professionals Care? ……..…**8**

Interpreting Research on Interviewer Effects ……... **11**

Interviewer Race and Ethnicity Effects …… **15**

Interviewer Gender Effects ……... **23**

Interviewer Age Effects …… **27**

Other Interviewer Characteristics …… **29**

Future Directions …….. **31**

CHAPTER 3 – DO AFRICAN AMERICAN TELEPHONE SURVEY RESPONDENTS PREFER AFRICAN AMERICAN INTERVIEWERS? ……**44**

Introduction ……...**44**

Methods …….…... **47**

Results …….……..**55**

Discussion ……**59**

CHAPTER 4 – LINGUISTIC EXPRESSIONS OF IDENTITY IN THE SURVEY INTERACTION: THE USE OF AFRICAN AMERICAN ENGLISH IN TELEPHONE SURVEYS ……**70**

Introduction ……...**70**

Background …….….. **71**
Methods …… **81**
Results …….……..**98**
Discussion ……**105**
CHAPTER 5 – CONCLUSION ……**122**
APPENDIX A …….**125**
BIBLIOGRAPHY ……..**128**

LIST OF TABLES

Table

2.1 Interviewer Race and Ethnicity Effects Studies ……..…... 35

2.2 Interviewer Gender Effects Studies ……... 39

2.3 Interviewer Age Effects Studies …… 42

3.1 Descriptive Statistics for the Study Sample ……... 66

3.2 Results of Linear Mixed Models Predicting Preferences for an Interviewer of One's Same Race and Ethnicity and Comfort with a White Interviewer by Ethnic Identity and Racial Salience …….. 67

3.3 Results of Linear Mixed Models Predicting Preferences for an Interviewer of One's Same Race and Ethnicity and Comfort with a White Interviewer by Type of Survey ……. 68

3.4 Mean Importance of Having an Interviewer of One's Same Race and Ethnicity and Comfort with a White Interviewer by Ethnic Identity Component ……... 69

4.1 Study Participants Included in Hypothesis Testing ……..……. 115

4.2 Descriptive Statistics for Study Participants …….. 116

4.3 Percents of Use of African American English Variants of Seven Phonological Features by African American Interviewers from Washington and Michigan ……. 117

4.4 Tests of Association Between Interviewers' Ethnic Identity Type Components and Their Use of African American English Variants of Seven Phonological Features ……. 118

4.5 Number of Tokens of African American English Variants of Seven Phonological Features Occurring in Recruitment, Racially Topical, and Sensitive Survey Sections …….…….. 119

4.6. Changes in the Proportion of African American Interviewers' Use of African American English Phonological Features During Recruitment, Racially Topical, and Sensitive Sections of a Survey Script ……. 120

4.7. Results of Models Testing Associations Between Interviewers' Ethnic Identity Type and Use of African American English Phonological Features and Ethnic Identity Subscale Means for *Eat for Life* Participants ……. 121

ABSTRACT

"WHATEVER IT MEANS TO YOU":
ETHNICITY, LANGUAGE, AND THE SURVEY RESPONSE IN TELEPHONE-ADMINISTERED HEALTH SURVEYS OF AFRICAN AMERICANS

by

Rachel Ellen Davis

Chair: Kenneth Resnicow

Many public health surveys match the race of telephone interviewers to the anticipated race of respondents. This practice may be particularly prevalent when surveying populations with a possible mistrust of health research, such as African Americans. Interviewer race effects have received little scientific evaluation in telephone-administered health surveys with African Americans, and no empirical studies have explored respondents' preferences for interviewer race or the role of ethnic identity in survey interactions. These dynamics may have important implications for respondents' experiences with health surveys and, ultimately, survey data. This dissertation reviewed the interviewer effects literature and considered the importance of interviewer-associated measurement error in public health surveys. It also described the results of two empirical studies. The first study explored African American telephone survey respondents'

preferences for interviewer race. This study found that respondents with Afrocentric, Black American, and Cultural Mistrust ethnic identity components preferred to interact with African American interviewers and that interviewer race was less important to respondents with Assimilated, Bicultural, or Multicultural identity components. Respondents' preferences for being surveyed by an African American interviewer were stronger when a survey contained more racial content. The second study explored the influence of African American telephone interviewers' ethnic identity types and the use of African American English (AAE) on the survey interaction. Interviewers with a Bicultural ethnic identity component may have been more prone to using AAE features during telephone surveys with African American respondents than interviewers with other identity components. Interviewers used AAE features less when engaged in recruitment tasks and appeared to have used them more when administering nonracial, sensitive survey items. No pattern emerged for use of AAE during racially topical survey sections. Whether interviewers had a Black American ethnic identity component and their use of AAE features had no impact on respondents' answers to racially topical survey items. The findings from this dissertation suggest that public health professionals should measure and control for interviewer effects on health survey data, be cognizant of the potential role of identity expression during survey interactions, and resist a general policy of matching African American interviewers and respondents by racial characteristics alone.

CHAPTER 1

INTRODUCTION

Despite increased public health attention, health disparities continue to plague the lives of many American citizens. These disparities are often most evident when comparing birth outcomes, life expectancies, and chronic disease mortality rates between African Americans and Whites.[1] In order to remedy these disparities, public health professionals must have access to the most effective and culturally appropriate means of communicating with African American populations. Health educators typically envision this communication as the means by which they design and deliver health messages to targeted recipients. However, health educators are also reliant upon communication in the opposite direction – namely, the receipt of health-related information from targeted populations. Public health professionals often rely upon baseline and follow-up survey data to evaluate the impact of public health programs. However, it may be less obvious that survey data typically drive the design of health education programs, as these data are often drawn from pre-existing sources. These data define the scope of a health problem, whom the problem affects, and the health-related beliefs and behaviors of the population. Such information is essential in determining what public health problems exists, which populations to target with health education programs, how to approach the target population, and what health communication messages to convey. As a result, the degree to which any health education program is effective is at least partially dependent on the quality of data that was used to inform the program design. If the communication methods employed to collect the data used to create a health program are flawed, then the capacity of the program itself to succeed is inherently weakened.

---

[1] There are many sources of information on racial and ethnic health disparities. For a good overview and list of useful links, see the website for the Centers for Disease Control and Prevention's Office of Minority Health and Health Disparities (http://www.cdc.gov/omhd/Highlights/2007/HFeb07.htm).

In an effort to improve public health information flow, this dissertation explores a singular aspect of data collection: measurement error in interviewer-administered health surveys. Like any science, the science of survey methodology is an evolving field. Much remains to be learned about the processes of survey data creation, reporting, recording, tracking, and analysis. These processes are impacted by many factors, but a prominent influence is the mode of survey administration. Whether a survey is self-administered or interviewer-administered is a primary consideration in survey design, because the presence or absence of an interviewer can impact several aspects of the survey process such as how comfortable a respondent feels about the confidentiality of his responses, how well the respondent understands the survey task, and the willingness of a respondent to disclose the information being requested.

Many public health surveys are administered by interviewers. As such, this dissertation seeks to specifically explore how the involvement of an interviewer might impact the collection of survey data from African American respondents. Many empirical investigations of the impact of interviewer race on survey data have been reported; however, these inquiries have focused on comparing data collected by White versus African American interviewers. In contrast, no work has been published that explores how within-group cultural variability among African Americans might play a role in the survey interaction. In an effort to address this knowledge gap, this dissertation focuses on how culture and the expression of culturally related social identity might impact survey data produced by African American interviewers and respondents. By better understanding the role of culture in survey interactions among African Americans, this research hopes to inform and improve the information base available to public health professionals.

This dissertation was designed to accomplish four primary objectives. The first objective, which is the focus of Chapter 2, is to summarize the state of the literature on the impact of interviewer characteristics such as race, gender, and age on survey data. This chapter is written for public health researchers and practitioners who may or may not have a strong background in survey research. Thus, Chapter 2 defines key terms, provides information on how to interpret research on interviewer characteristics, and

argues for the importance of assessing and reporting interviewer effects in public health survey data. Chapter 2 also highlights many avenues for future research.

The second objective of this dissertation is to explore the preferences of African American respondents for interacting with African American interviewers. These issues are examined in Chapter 3. Chapter 3 also introduces the larger study from which this dissertation was derived: the *Eat for Life* trial. *Eat for Life* was a public health intervention trial that sought to test whether tailoring health communication materials on ethnic identity or motivational predisposition would increase fruit and vegetable consumption among African American adults (Resnicow et al., 2008; Resnicow et al., under review). In designing this study, our research team had many interesting conversations about whether or not to employ an all-African American interviewing staff to administer the telephone surveys that were utilized to collect pre- and post-intervention data. On the one hand, many members of our team and their respective Institutional Review Boards felt that it was important for African American interviewers to administer the *Eat for Life* surveys to an African American participant population, particularly since the surveys contained racial attitude questions. However, because it is illegal to hire employees based on race or ethnicity, the members of the research team in charge of administering the surveys could not guarantee the exclusive staffing of African American interviewers on our project. Our discussions of this topic would have been greatly enhanced by data on the importance of interviewer race to African American survey respondents. However, almost no literature existed that could be used to inform our decision. Chapter 3 is a first step in this direction; however, much empirical work remains to be conducted on this important issue, as it may have implications for health survey data as well as the experience of survey participation for African American respondents.

The third objective of this dissertation is to explore how racial and ethnic identity might be conveyed by African American interviewers during telephone-administered surveys. It is known from Chapter 2 that interviewer race and ethnicity can and sometimes do have an impact on telephone-administered survey data. These effects can only be possible if interviewers have a means of conveying their race and ethnicity in a telephone setting. But, how is this accomplished during a scripted, standardized

telephone interview? When might interviewers strive to convey or not convey their racial and ethnic affiliations to African American respondents? And, do such conveyances impact the answers that respondents provide to different types of survey items? In an attempt to answer these questions, Chapter 3 presents the results of a small study designed to explore African American interviewers' use of culturally associated linguistic features in telephone interviews with African American respondents.

It is hoped that this dissertation will provide practical information for public health professionals in their use of telephone surveys with African American populations. However, the present investigations are but small steps in an attempt to fill a much larger void in our collective knowledge about the survey interaction. If this dissertation succeeds in inspiring further study of the role of within-group cultural variability and communication in the survey process, this work will have achieved its fourth and final objective.

CHAPTER 2

THE IMPORTANCE OF INTERVIEWER ERROR IN PUBLIC HEALTH SURVEYS

INTRODUCTION

Survey data often provide the rationale for the initiation, continuation, and termination of programs charged with the heady responsibility of promoting the public's health. Prevalence estimates, incidence rates, and information about health-related attitudes, behaviors, and beliefs are routinely assessed through surveys. Survey data are also predominantly used to evaluate the effects of public health interventions. With the responsibility of the public's health on their shoulders, many public health professionals are committed to grounding their work in the most valid and reliable data that can be obtained with the resources that they have available. Yet, survey methodology is an evolving science, and much remains unknown about how to extract health-related survey data from a complex social world. Public health professionals must therefore monitor advances in survey methodology in order to stay abreast of the most effective data collection procedures.

This paper reviews the current state of knowledge about an aspect of survey methodology that may be a source of confusion and concern for many public health professionals – interviewer-related measurement error. Herbert Hyman argued:

> (T)he demonstration of error marks an advanced stage of a science. All scientific inquiry is subject to error, and it is far better to be aware of this, to study the sources in an attempt to reduce it, and to estimate the magnitude of such errors in our findings, than to be ignorant of the errors concealed in the data. (Hyman, Cobb, Feldman, Hart, & Stember, 1954/1975, p. 4)

Only a handful of articles are devoted to interviewer error in the public health literature, and many of these are affected by severe methodological problems. Yet, due to a growing emphasis on cultural sensitivity in public health practice and research, health professionals are often confronted with the question of whether or not to match interviewers with the anticipated sociodemographic characteristics of survey respondents. As a consequence, there is a strong need for guidance about whether and how interviewer characteristics impact survey data.

This paper reviews the state of knowledge of interviewer error in an attempt to assist public health professionals in executing best practices for survey data collection and interpreting survey-based literature. This review encompasses those interviewer-associated sociodemographic characteristics that are most often used as interviewer selection criteria and for which some empirical interviewer error research exists. Since much remains to be learned about interviewer error, this paper also aspires to equip the reader with tools to evaluate future interviewer error research. This review generally excludes studies conducted with children and in cultural settings outside the United States. However, it is hoped that this paper will prepare the reader to evaluate research with other populations and settings. To these ends, this review defines interviewer error, describes the potential effects of interviewer error on health related survey data, outlines aspects to consider in evaluating interviewer error research, reviews existing interviewer error research, and discusses select implications of the extant literature for public health research and practice.

WHAT IS INTERVIEWER ERROR?

Error is the difference between the observed value of a phenomenon, which is measured using survey data, and the true value of that phenomenon, which is often impossible to measure (Groves, 2004). According to the total survey error model put forth by Groves (2004), the total amount of error associated with a particular survey statistic is the mean square error. The mean square error represents the sum of both bias and variable errors impacting that survey statistic. Both bias and variable errors describe

what would be expected to occur over hypothetical, repeated administrations of the same questionnaire with the same sampling design, recruitment protocol, and data collection procedures, regardless of whether or not the survey is actually re-administered. A specific administration of the survey is regarded as a sampling of what would happen over these repeated survey administrations. However, whereas bias describes errors that are expected to occur on every administration of the same survey design, variable errors describe errors that are expected to vary across survey administrations. Both bias and variable errors may derive from sources of error relating to coverage, nonresponse, sampling, interviewer, respondent, instrument, and mode (Groves, 2004).

Interviewer error refers to variance in survey data that arises from the fact that data collected by either a specific individual interviewer or a specific set of interviewers may be different than data collected by another individual or set of interviewers administering the same questionnaire to the same population of respondents. Interviewer error can result from role-dependent effects such as differences in the way interviewers administer the survey, pronounce questions, probe, or provide off-script instructions to respondents (Brenner, 1982; Groves, 2004). Interviewer error can also result from role-independent features such as how respondents react to perceived interviewer sociodemographic characteristics such as race, gender, or age (Hagenaars & Heinen, 1982). Interviewer errors are usually variable errors related to the particular interviewers selected; however, bias may result if the source of error is rooted in an aspect of the survey design, such as the use of only female interviewers or identical but flawed interviewer training procedures across repeated administrations of the survey (Groves, 2004). The term "interviewer effects" refers to measurement error attributable to a specific interviewer characteristic such as race or gender (Dijkstra, 1983). As with interviewer error, interviewer effects incorporate both bias and variable errors but typically refer to variable errors.

WHY SHOULD PUBLIC HEALTH PROFESSIONALS CARE?

The presence of an interviewer can affect how a respondent forms an answer to a survey question and whether and how a respondent edits his answer before communicating it (Davis & Silver, 2003; Krysan & Couper, 2003; Schuman & Converse, 1971; Tourangeau, Rips, & Rasinski, 2000). Williams proposed that respondents may be motivated to edit their answers to both increase the potential rewards and decrease the negative ramifications of a survey interaction (Dohrenwend, Williams, & Weiss, 1969; Williams, 1968). If this is true, then public health surveys may be particularly vulnerable to response editing, as a respondent's answers to a health questionnaire can determine whether he will gain access to health-related medical and educational services, be eligible to participate in research offering desirable incentives, expose himself to social harm through the revealing of stigmatized behaviors or health conditions, or invite an unwelcome medical diagnosis. In such situations, respondents may not only seek to provide a positive self-image, but also to scrutinize their interviewer for clues as to what responses might result in the best outcomes from their participation in the survey. These clues – sociodemographic characteristics such as race and gender, personality traits, speech styles, etc. – can thereby result in interviewer effects.

In accordance with standard survey methodology practice, project managers typically try to minimize role-dependent interviewer effects by training interviewers to convey as little personal information as possible to respondents. Interviewers are generally discouraged from expressing their own opinions, reacting affectively to respondents' statements, or wearing clothing or accessories that suggest certain behavioral or attitudinal dispositions. However, role-independent interviewer effects are more difficult to avoid, since interviewer characteristics such as gender, age, and race are often hard to conceal. These qualities may be more obvious in face-to-face encounters than in telephone administrations. But, even when no visual cues are accessible, such clues as an interviewer's name, speech characteristics, intonation, off-script word choices, and voice pitch may enable respondents to form opinions about their interviewer's sociodemographic profile.

Interviewer effects can have significant consequences for public health survey data. These consequences arise from the fact that interviewer effects are generally associated with data from numerous survey respondents. Since the number of respondents in a survey is usually greater than the number of interviewers engaged, multiple respondents tend to be interviewed by each interviewer, resulting in the clustering of respondents by interviewers. Any effects associated with that interviewer may then apply to data obtained from the entire group of respondents that were surveyed by that interviewer. For example, if a survey employed one brunette interviewer and four blonde interviewers to administer a face-to-face survey to 100 respondents about attitudes toward people with brown hair, the 20 respondents interviewed by the brunette interviewer might be expected to voice more positive attitudes toward brown-haired people than either they or the 80 respondents interviewed by the blonde interviewers really hold. The differences between the 20 respondents' true attitudes about people with brown hair and the answers they reported are interviewer errors. But, because there was a group of respondents who adjusted their responses to the brunette interviewer, the errors associated with this group of respondents are correlated with one another. As a result, the overall mean responses for items querying attitudes toward brown-haired people from all 100 respondents will be artificially high, the errors associated with these means will not be independent, and the overall variance, and, by association, standard deviations and standard errors of the means, will be inflated by interviewer error (Groves, 2004). Interviewer error can weaken the stability of survey statistics, increase or decrease the magnitude of estimates, and influence the relationships observed among variables (Davis, 1997b).

Interviewer effects are typically quite small, but even small effects can have a substantial impact on survey findings. For instance, the Kish intraclass correlation coefficient, ρ^*_{int}, is often used to measure interviewer variance associated with a survey statistic. Possible values for ρ^*_{int} generally range from a low of zero, although negative values are sometimes reported, to a high of 1.0, which indicates perfect correlation. Calculations of ρ^*_{int} values for statistics measured across ten face-to-face surveys (Groves, 2004) and nine telephone surveys (Groves & Magilavy, 1986) suggest average ρ^*_{int} values of 0.031 and 0.009, respectively. However, significantly larger ρ^*_{int} values

have been reported (Freeman & Butler, 1976), and Groves and Magilavy (1986) have found that ρ^*_{int} estimates are themselves highly unstable and associated with large variances. Thus, different compositions of interviewing staff could result in considerably higher and lower ρ^*_{int} values across replications of the same survey (Groves & Magilavy, 1986). Although these numbers appear small, their impact on survey statistics is multiplied by the numbers of interviews conducted by each interviewer. The general impact of ρ^*_{int} on the variance of a survey statistic, $deff_{int}$, is estimated by $deff_{int} = 1 + p_{int}(m - 1)$, where m is the average number of interviews completed per interviewer (Groves & Magilavy, 1986). (Note: Alternate formulas can be used to calculate $deff_{int}$ using the exact size of each interviewer's workload.) For instance, inserting the small ρ^*_{int} of 0.009 and a small average interviewer workload of 25 into the $deff_{int}$ equation results in a 122% increase in the variance of a sample mean. Holding ρ^*_{int} constant at 0.009 and increasing the average interviewer workload to 75 yields an even higher increase in the variance of a sample mean of 167%. Similarly, holding the average interviewer workload constant at 25 and increasing ρ^*_{int} to 0.03 results in a 172% increase in the variance of a sample mean. Thus, either a large interviewer effect or a large interviewer workload can result in a sizeable impact on survey estimates.

Interviewer effects can be explored using multi-level statistical models capable of accounting for the clustering of respondents by interviewers. Such models define interviewers as a separate level of data apart from respondent data and account for the nesting of respondent data within interviewers. By structuring the data into levels, the amount of variance derived from differences across interviewers can be measured. If this between-interviewer variance is substantial, it can be further evaluated to determine the most significant sources of variance. However, if the between-interviewer variance is acceptably mild, it can be ignored, and simpler, single-level statistical models can be used. Several scholars of interviewer effects have documented how their study findings would have been different if they had not accounted for the clustering of respondents by interviewers (Dailey & Claus, 2001; Dijkstra, 1983; Fendrich, Johnson, Shaligram, & Wislar, 1999; Heeb & Gmel, 2001; Kane & Macaulay, 1993; Wilson, Brown, Mejia, & Lavori, 2002). In general, accounting for clustering appears to dampen the significance of direct effects and interactions among variables because unadjusted models

underestimate variance as a consequence of ignoring between-interviewer variance. Studies that do not account for interviewer clustering may, therefore, report stronger interviewer effects than those that would have been found using statistical techniques capable of adjusting for clustering effects (Dailey & Claus, 2001; Dijkstra, 1983; Fendrich et al., 1999), resulting in type 1 error.

Public health professionals should be aware that when significant interviewer effects are operant, changes in interviewer staffing may confound the findings of surveys that are repeatedly administered over time (Anderson, Silver, & Abramson, 1988a; Kish, 1962). In a study of five administrations of the face-to-face-administered National Election Study between 1964 and 1984, Anderson, Silver, and Abramson (1988a) found that the proportion of African American respondents interviewed by African American interviewers dropped from 43% in 1964 to 13% in 1984. As a result, Anderson and colleagues concluded that interviewer race effects and changes in the proportion of African American interviewing staff might have artificially reduced estimation of African Americans' reported closeness to other African Americans in longitudinal analyses (Anderson, Silver, & Abramson, 1988b).

Lastly, interviewer effects may impact subsequent respondent behavior. Very little research has been conducted on this topic, but Anderson, Silver, and Abramson (1988a) found that Black respondents living in certain geographical areas who were interviewed by Black interviewers in face-to-face pre-election surveys were about 12% more likely to vote in ensuing elections than Black respondents interviewed by White interviewers. It is not known whether interviewer characteristics influence health behaviors, but public health researchers have an ethical responsibility to ensure that surveys do not induce negative consequences to health.

INTERPRETING RESEARCH ON INTERVIEWER EFFECTS

Public health professionals should be mindful of several factors when interpreting studies of interviewer effects: the numbers of interviewers and respondents; interviewer assignment procedures; the clustering of respondents by interviewers; controlling for

extraneous interviewer characteristics; the number of statistical tests performed; mode of administration; and study timing. Rigorous studies of interviewer effects utilize large numbers of both interviewers and respondents (Groves & Fultz, 1985). Just as larger numbers of respondents lower sampling variance, so, too, will larger numbers of interviewers reduce interviewer-associated variance by diluting the effects of individual interviewers upon the data collection process (Stock & Hochstim, 1951). However, larger numbers of interviewers may increase role-dependent interviewer errors due to diluted training, supervision, and monitoring. Thus, while smaller interviewing teams are a disadvantage in role-independent interviewer effects research, smaller interviewing teams may be desirable when balancing multiple sources of error in practical applications.

Some studies of role-independent interviewer effects involve too few interviewers to isolate a single sociodemographic characteristic from other characteristics associated with the particular interviewers used (e.g., Becker, Feyisetan, & Makinwa-Adebusoye, 1995; Carpenter, Andrykowski, Cordova, Cunningham, & Studts, 1999; Fry, Rozewicz, & Crisp, 1996; Landis, Sullivan, & Sheley, 1973). Groves and Fultz (1985) have advocated for interpenetrated survey designs, which randomly assign interviewers to respondents to avoid confounding interviewer and respondent characteristics. However, survey designs are constrained by many practical factors, and these factors often preclude interpenetrated designs with large quantities of interviewers and respondents (Krysan & Couper, 2003). For example, due to geographical challenges, it is often necessary to employ a larger number of interviewers for face-to-face surveys. But, random assignment is easier to execute for telephone surveys. It is also true that many survey methods studies are conducted within the confines of studies designed for other purposes. As a result, much of the literature on interviewer effects consists of studies of either telephone surveys with small numbers of interviewers or face-to-face surveys where interviewers were either not or incompletely randomized to respondents (Krysan & Couper, 2003).

Rigorous studies should also make an attempt to measure and control for the clustering of respondents within interviewers. As previously noted, the failure to account for data clustering may result in artificially significant findings, as significant effects may

emanate from particularly influential qualities associated with individual interviewers rather than characteristics shared across interviewers such as gender or race (Dijkstra, 1983; Fendrich, Johnson, Shaligram, & Wislar, 1999). Most research on interviewer effects was conducted prior to the 1980s, when statistical software began to become available that was capable of accounting for clustering via multi-level regression models. Thus, most studies conducted prior to the 1980s and, unfortunately, many studies conducted since, have not accounted for data clustering. Before the advent of statistical software, many researchers used techniques such as analysis of variance to measure interviewer effects. Although there are some statistical limitations to analyses of variance that lead to underestimates of variance (Freeman & Butler, 1976; Groves & Magilavy, 1986; Hanson & Marks, 1958; Hox, de Leeuw, & Kreft, 1991; Kish, 1962), analyses of variance can be interpreted with more confidence than single-level techniques if the interviewer, as opposed to the respondent, is treated as the unit of analysis. However, most publications are limited in word space, and many descriptions of analytical procedures are insufficiently detailed to enable readers to determine how analyses were structured. Interviewer effects studies reporting the Kish ρ^*_{int} are useful because, unlike analysis of variance approaches, ρ^*_{int} is unit-free and can be compared across surveys with different designs (Groves, 2004; Groves & Magilavy, 1980). The Kish ρ^*_{int} is also appropriate for models that consider interviewer assignment as a random effect. It is ideal to treat the specific group of interviewers used to administer a survey as a random sample of the many potential samples of interviewers that could have been drawn from a larger, hypothetical pool of interviewers (Fendrich et al., 1999; Groves, 2004). This approach enables more confident generalizations of interviewer effects research findings to other survey designs employing different samples of interviewers (Dijkstra, 1983).

Even if a study has accounted for data clustering, it is important that researchers have isolated the characteristic under study by controlling for other influential interviewer characteristics (Groves & Fultz, 1985). For example, both older and female interviewers may tend to be more experienced; thus, an analysis of age or gender of interviewer effects that does not control for interviewer experience may be confounded by interviewer experience. Interviewer effects that are not the foci of a study can be

controlled through study design mechanisms, such as hiring interviewers who all possess similar characteristics, or statistical techniques applied in the data analysis phase.

Another methodological issue is the vulnerability of findings in the literature due to chance. Some studies evaluate large numbers of individual survey items using independent statistical tests for evidence of interviewer effects. This approach increases one's odds of encountering significant findings due to chance alone (Feldman, Hyman, & Hart, 1951-1952). Many published studies do not disclose the number of tests conducted; thus, it is impossible for the reader to evaluate the odds of finding the reported results (Dijkstra, 1983). Some studies have addressed this challenge by selecting either a small number of items to test or testing subscale means in lieu of individual items.

Studies of interviewer effects may also be impacted by the mode of survey administration. Respondents in face-to-face surveys may have access to a range of auditory and visual clues to an interviewer's sociodemographic or attitudinal characteristics. But, telephone survey respondents only have access to auditory clues, which may or may not lead to accurate assumptions about an interviewer's age, gender, race, ethnicity, social class, attitudinal, or other personal characteristics. Thus, studies of face-to-face surveys may be more prone to finding statistically significant interviewer effects than studies of telephone-administered surveys.

Survey interviews reflect prevailing race, gender, and other social relations from particular moments in time (Davis, 1997b; Krysan & Couper, 2003; Schuman & Converse, 1971). As these social relations evolve, the impact of particular interviewer effects is likely to covary. Limited support for this argument can be drawn from literature on interviewer race effects. In a study of interviewer race effects in the National Election Studies, Anderson, Silver, and Abramson (1988a) compared reports of voting to actual voting as documented in official voting records. In the first two years under analysis, 1964 and 1976, voting records indicate that Black nonvoters were more likely to report voting to White interviewers than to Black interviewers. However, Black nonvoters in 1978, 1980, and 1984 were more likely to report voting to Black interviewers than to White interviewers. While these discrepancies between voting records and reporting may result from vote tampering, these discrepancies may also be an artifact of the changing political context in which the survey waves took place (Anderson

et al., 1988a). Webster suggests that findings from surveys conducted in majority Anglo social settings may differ from those conducted in social environments where Hispanics are the majority (Webster, 1996). Thus, interviewer effects should be expected not only to change over time, but to also vary across social environments.

INTERVIEWER RACE AND ETHNICITY EFFECTS

Matching interviewer race or ethnicity to the anticipated race or ethnicity of survey respondents is a common practice (Groves, 2004). Intuition suggests that survey respondents will feel more comfortable and, therefore, be more open and honest, to interviewers of their same race and ethnicity. However, respondents may or may not provide more accurate data to racially and ethnically homophilous interviewers. For example, respondents may provide consistently similar answers to interviewers who do not match their own racial profiles in an effort to dispel race-based stereotypes. Conversely, respondents might provide consistently different responses to subgroups of same-race interviewers, as same-race social pairings may induce salient ethnic differences between members of the same race. In these ways, the systematic matching of interviewers and respondents by race or ethnicity across studies can result in bias (Groves, 2004). Thus, it is important to evaluate whether available data provide support for or advocate against the matching of interviewers and respondents on certain racial or ethnic characteristics. The following sections summarize research to date on interviewer race and ethnicity effects in face-to-face and telephone surveys.

Interviewer Race and Ethnicity Studies – Face-to-Face Surveys

As detailed in Table 2.1, at least four studies assessed interviewer race effects in face-to-face surveys in the late 1950s and early 1960s. Overall, this era of studies suggests that interviewer race effects may be associated with data from both Black and White respondents and be invoked merely by the presence of a survey administrator (Summers & Hammonds, 1966), regardless of the level of interpersonal interaction.

Findings from these studies indicate that interviewer race effects may be particularly strong when interviewers are low on objectivity (Williams, 1968), when interviewers and respondents are discordant on several sociodemographic characteristics (Lenski & Leggett, 1960; Williams, 1964; Williams, 1968), and when race-related questions are involved (Summers & Hammonds, 1966; Williams, 1964; Williams, 1968). The direction of the effect appears to be in the direction of deference to the interviewer's race, meaning that respondents tend to respond in ways that would generally be expected to be perceived more positively by a member of the interviewer's race (Summers & Hammonds, 1966; Williams, 1964; Williams, 1968). However, while these early findings are intriguing, the general lack of statistical adjustment for data clustering limits the conclusions that can be drawn from these early findings.

Studies from the late 1960s to early 1980s indicate that interviewer race and ethnicity effects are evident in data obtained from African American (Anderson et al., 1988b; Campbell, 1981; Schaeffer, 1980; Schuman & Converse, 1971), White (Campbell, 1981; Freitag & Barry, 1974; Hatchett & Schuman, 1975-1976; Schaeffer, 1980), and Mexican American (Welch, Comer, & Steinman, 1973) respondents. With the exception of the Carr study (Carr, 1971), findings from this group of studies are generally quite similar. Many, if not most, survey questions from this era appear not to be affected by interviewer race or ethnicity (Anderson et al., 1988b; Campbell, 1981; Carr, 1971; Schaeffer, 1980; Welch et al., 1973), including demographics and health status questions (Schaeffer, 1980; Schuman & Converse, 1971). When interviewer effects were present, they mostly occurred for race-related items, and respondents' answers were in the direction of deferring to the interviewers' race (Anderson et al., 1988b; Campbell, 1981; Hatchett & Schuman, 1975-1976; Schaeffer, 1980; Schuman & Converse, 1971). Schuman and Converse (1971) observed that among racial items demonstrating interviewer race effects, opinion items and items with no obvious best response option were most likely to be affected. The Carr (1971) study analyzed nonracial survey items but found a general acquiescence bias, or tendency to agree, among the survey sample. Findings from the Campbell (1981) study indicate that interviewer race effects occur in the presence of interviewers, regardless of whether a survey question is interviewer- or self-administered. The Freitag and Barry (1974) study

suggests that item nonresponse might be higher for White respondents when surveyed by Black interviewers. However, several studies from this era do not report their interviewer assignment procedures (Carr, 1971; Freitag & Barry, 1974; Welch et al., 1973), and only the Schuman and Converse (1971) and Hatchett and Schuman (1975-1976) studies appear to have randomly assigned interviewers to respondents. The Hatchett and Schuman (1975-1976) and Schaeffer (1980) studies are the only studies that attempted to adjust for the clustering of respondents by interviewers. Thus, many of these studies may have overestimated the significance of true interviewer race effects.

Findings from two 1990s studies suggest that Black and White respondents may be more likely to report substance use to White or Hispanic interviewers than Black interviewers (Fendrich et al., 1999; Johnson & Parsons, 1994). Black and White respondents may also be more likely to report physical abuse to White interviewers than to Black interviewers; however, reporting of sexual abuse may not be affected by interviewer race (Dailey & Claus, 2001). The Webster (1996) study provides further support for the potential influence of interviewer ethnicity on racial and ethnic survey items, both for Anglo and Hispanic respondents. This study also suggests that ethnic homophily may improve survey response rates and that ethnicity and culture may result in interactions between interviewer race or ethnicity and other interviewer characteristics such as gender. Despite some significant findings from these studies, most nonracial survey items were not impacted by interviewer race or ethnicity (Johnson & Parsons, 1994; Webster, 1996; Zimmerman, Caldwell, & Bernat, 2002). Thus, it appears that interviewer race effects do not affect all survey topics equally. Unlike studies from the 1950s-1980s, however, only one of the 1990s surveys (Webster, 1996) contained explicitly racial or ethnic items. But, although none of the 1990s studies may have employed random interviewer assignment, the fact that most of these 1990s studies adjusted for interviewer clustering gives more weight to these findings than studies from prior decades.

The most recent face-to-face interviewer race effects study is a laboratory experiment in which Krysan and Couper (2003) used live and virtual interviewers in a survey of racial attitudes. Videotapes of two Black and two White interviewers were used to guide Black and White respondents through a computer assisted self-administered

version of the same survey in the virtual interviewer condition. Respondents were assigned to either a live Black, live White, virtual Black, or virtual White interviewer. The data suggest that Black respondents reported more racially deferential attitudes to both live and virtual White interviewers than to Black interviewers. White respondents expressed more positive attitudes about Blacks for three items to live and virtual Black interviewers but varied their attitudes toward Whites only when interacting with live interviewers. For other racial items, White respondents demonstrated opposing response patterns to live and virtual interviewers or demonstrated no interviewer race effects. This study indicates that respondents of different races may interact differently with live and virtual interviews; however, this line of research requires further investigation with larger samples, as the small sample size of this study may have limited the number of significant findings.

Interviewer Race and Ethnicity Studies – Telephone Surveys

The first telephone-administered surveys to examine interviewer race and ethnicity effects were published in the 1980s. Taken together, these 1980s studies provide evidence that interviewer race and ethnicity can impact telephone-administered surveys among White (Cotter, Cohen, & Coulter, 1982; Finkel, Guterbock, & Borg, 1991; Reese, Danielson, Shoemaker, Chang, & Hsu, 1986), Black (Cotter, Cohen, & Coulter, 1982; Davis, 1997a, 1997b), and Hispanic (Reese, Danielson, Shoemaker, Chang, & Hsu, 1986) respondents despite the absence of visual cues. These studies suggest that interviewer race and ethnicity effects are most likely to occur for racial and ethnic survey items (Cotter, Cohen, & Coulter, 1982; Davis, 1997a, 1997b; Finkel, Guterbock, & Borg, 1991; Reese et al., 1986). Although some effects are noted for nonracial items (Cotter et al., 1982; Davis, 1997b), no effects are observed for most nonracial items (Cotter et al., 1982; Singer, Frankel, & Glassman, 1983). When present, interviewer effects tend to emerge when the items overtly query attitudes related to race or ethnicity (Cotter et al., 1982; Davis, 1997a, 1997b; Finkel et al., 1991; Reese et al., 1986). In all instances, respondents deferred to the interviewer's race by providing responses that would generally be perceived more positively by members of the

interviewer's race (Cotter et al., 1982; Davis, 1997a, 1997b; Finkel et al., 1991; Reese et al., 1986). The first Davis study (1997b) found that interviewer effects can affect many items throughout a questionnaire, thereby weakening or strengthening observed relationships among variables. Findings from this study also indicate that perceived interviewer race may have a greater impact on responses than actual interviewer race (Davis, 1997b). The second Davis study (1997a) suggests that African American respondents may adopt a general acquiescence stance when interacting with White interviewers. Respondents surveyed by White interviewers were significantly more likely to express support for both the Democratic and Republican parties, Jesse Jackson and Ronald Reagan, and Black politicians seen as supportive of Jesse Jackson and Ronald Reagan than respondents surveyed by Black interviewers. Davis (1997a) proposed that agreement with cognitively dissonant responses resulted from efforts on the part of Black respondents to appease White interviewers by neutralizing their expressed political opinions. Davis (1997a) argued that African American respondents may "take more accommodating and powerless positions in response to White interviewers" as a self-defense strategy (p. 319). An alternate explanation is that these patterns evince a general distrust of White interviewers. Both theories warrant substantiating research. Findings from the Singer, Frankel, and Glassman study (1983) indicate that interviewer race does not appear to impact survey or item nonresponse over the telephone; however, this study did not contain racial items. All but one (Singer et al., 1983) of the 1980s telephone studies used random assignment of interviewers. However, the Singer, Frankel, and Glassman study was the only study to statistically control for the clustering of respondents by interviewers. Thus, despite many methodological strengths, these 1980s studies may represent overestimates of statistically significant findings.

Three studies report on investigations of interviewer race and ethnicity effects in telephone surveys from the 1990s and early 2000s. These most recent telephone studies suggest that perceived race can be a stronger predictor of interviewer race effects (Davis & Silver, 2003), that racial items are most affected (Wolford, Brown, Marsden, Jackson, & Harrison, 1995), that nonracial items tend not to be affected (Davis & Silver, 2003; Livert, Kadushin, Schulman, & Weiss, 1998), and that African American respondents will defer to the race of the interviewer (Wolford et al., 1995). Building on the work of

Claude Steele (Steele, 1997), Davis and Silver (Davis & Silver, 2003) hypothesized that stereotype threat activation, in which "the pressure to disconfirm … negative stereotype(s) produces anxiety that interferes with the processing of information" (p. 35), might be more responsible for interviewer race effects than social desirability bias, as is commonly assumed. Their study findings suggest that African American respondents may fare poorer on knowledge tests when they believe that their interviewer is White (Davis & Silver, 2003), but whether these findings are due to the stereotype threat hypothesis requires further examination. In the single study of the 1990s group that accounted for respondent clustering, Livert, Kadushin, Schulman, and Weiss (1998) concluded that when interviewer race and ethnicity effects occurred in reporting of substance use, interviewer race and ethnicity interacted with respondent race and ethnicity. In the Livert et al. (1998) study, White respondents were generally unaffected by interviewer race or ethnicity. But, Black respondents reported more alcohol and marijuana use, approval of alcohol and marijuana use, and poorer overall health status to Black interviewers, and Hispanic respondents reported greater perceived harm from alcohol use, marijuana use, and approval of marijuana use to Black interviewers. The Livert et al. (1998) study indicates that interviewer race and ethnicity has no bearing on reporting of tobacco, cocaine, barbiturate, crack, or heroin use, but this finding may be an artifact of the fact that use of these latter drugs is relatively rare. Of these most recent interviewer race effects telephone studies, only the Livert study utilized both random assignment and the statistical methods to control for the clustering of data by interviewers.

Interviewer Race and Ethnicity Studies – Summary and Future Directions

The accumulated research on interviewer race and ethnicity effects results in several broad conclusions. For one, evidence indicates that interviewer race and ethnicity can have a significant impact on certain types of survey data by influencing nonresponse as well as the magnitude, accuracy, and relations among study findings. Significant interviewer race and ethnicity effects appear to be the exception rather than the rule. But, when interviewer race and ethnicity effects occur, they tend to occur for overtly racial or

ethnic attitude questions and not to occur for more subtle racial/ethnic or nonracial/nonethnic questions where a connection to race or ethnicity is merely implied. Interviewer race and ethnicity appears not to impact sociodemographic items. However, few studies report analyses of such items, and more research is warranted. Reporting of overall health status has had mixed findings. Studies of reporting of physical abuse and substance use indicate that significant interviewer race and ethnicity effects may be operant. More interviewer race and ethnicity effects studies are needed on surveys of health-related topics. Interviewer race and ethnicity effects can occur for face-to-face surveys, telephone surveys, and even self-administered surveys when live or virtual interviewers are present. However, live interviewers may have different implications for the response process than virtual interviewers. When interviewer race and ethnicity effects occur, respondents defer to the race of the interviewer by reporting in ways that would appear to be perceived more positively by members of the interviewer's race or ethnicity.

The literature suggests that interviewer race and ethnicity effects can and do occur in surveys of Black, White, and Hispanic respondents. It is therefore likely that interviewer race and ethnicity can impact survey data from additional racial and ethnic groups. One study of juvenile respondents indicates that interviewer race and ethnicity effects do occur among other racial and ethnic populations in the United States (Weeks & Moore, 1981); however, this hypothesis requires further exploration. In addition, no studies have published findings on the effects of variability in racial and ethnic identity within racial and ethnic groups. Research indicates that African Americans, for example, vary in their feelings about race and ethnicity (Cross, 1991; Sellers, Smith, Shelton, Rowley, & Chavous, 1998). As a consequence, African Americans who view being Black as one of the most important aspects of their personal identities (i.e., racial centrality) may interact differently with African American interviewers who place varying importance on being Black in construing their own personal identities. Research is needed on whether survey respondents' views about their own race and ethnicity impact their interactions with both interviewers of various races and ethnicities and interviewers with strong or weak ties to their own racial and ethnic groups.

The few telephone studies that report analyses of actual and perceived interviewer race suggest that perceived interviewer race and ethnicity might be a better predictor of interviewer effects for telephone-administered surveys. Studies suggest that survey respondents' abilities to correctly judge the race of a telephone interviewer vary widely, with estimates of correct interviewer race identification ranging from 14%-82% (Davis, 1997b; Wilson, 2007; Wolford et al., 1995). Because perceptions of race over the telephone are often inaccurate, Wilson (2007) argues that analyses utilizing only variables for actual interviewer race in telephone studies incur measurement error resulting from differences between perceived and actual race, since respondents are more likely to be affected by perceived interviewer race than actual interviewer race. These differences between actual and perceived interviewer race may therefore undermine the validity of study findings based only on actual interviewer race (Wilson).

Lastly, but perhaps most importantly, the extant literature on interviewer race and ethnicity effects fails to conclude whether survey respondents feel more comfortable with, trust, prefer, or provide more accurate data to interviewers of their same race and ethnicity. Many existing studies assume that racially and ethnically homophilous interviewer-respondent dyads produce more valid survey data. However, as other authors suggest, it is not known whether survey data obtained by racially and ethnically matched interviewers are more accurate because most studies focus on attitude items for which there are no accessible validity checks. Whether a Black respondent provides more accurate data to a Black interviewer likely depends on the survey topic, the era and social environment in which the survey was administered, and the respondent's feelings about his own racial and ethnic identity. Yet, these factors are assumed to be unimportant when interviewers and respondents are intentionally matched by race or ethnicity alone. Additional research on the impact of matching respondents to interviewers by racial and ethnic characteristics is sorely needed.

INTERVIEWER GENDER EFFECTS

Whether a survey is conducted in person or over the telephone, gender may be one of the most easily identifiable interviewer characteristics. Benney, Riesman, and Star observed that "'(N)ice girls' don't know how to shoot craps" (p. 144) in 1956, and, although they are continually changing over time, many gender-specific social expectations persist for males and females today. The strength of such gender roles fuels the possibility that interviewer gender impacts survey data, particularly when surveys query issues related to gender norms. In such situations, respondents may invoke gender-based stereotypes when editing their survey responses to increase their odds of obtaining a desirable result from the survey interaction, be it a positive impression made on the interviewer or access to health information or services. For example, male respondents who enjoy flirting with women may be more inclined to provide survey responses that are supportive of or flattering to women's causes. Women, on the other hand, may be more likely to provide feminist responses to female interviewers to demonstrate in-group solidarity. Since interviewers in the United States tend to be female, the exploration of interviewer gender effects is an important component of endeavors to better understand and cope with the effects of interviewer error.

As displayed in Table 2.2, research on interviewer gender effects has been almost as extensive as research on interviewer race. In one of the earliest studies of interviewer gender, Benney, Riesman, and Star (1956) concluded that female interviewers were more likely than male interviewers to rate respondents as being frank and honest. A slightly later study (Colombotos, Elinson, & Loewenstein, 1969) found no evidence for interviewer gender effects. However, studies from the 1950s-1960s are too sparse to draw any conclusions from this era with confidence.

The 1970s spawned several investigations of interviewer gender research. Among the earliest of these explorations was reported by Landis (1973), who reported that female respondents expressed significantly more feminist responses to a male interviewer than to a female interviewer. However, since only two interviewers were employed, the findings of this study may be more attributable to non-gendered characteristics of the two interviewers than to gender alone. Other studies from this era

report a lack of consistent, interpretable interviewer gender effects (DeLamater, 1974; Groves & Magilavy, 1986; Johnson & Delamater, 1976). Freeman and Butler (1976) associated higher ρ^*_{int} values with male interviewers, but the number of male interviewers in their study was too small to explore the interaction of interviewer and respondent gender. Given the small number of interviewers used in the Landis (1973) study and the fact that the Freeman and Butler (1976) study utilized few male interviewers, the 1970s studies provide little overall evidence upon which to make inferences about interviewer gender effects.

Whereas the prior studies examined face-to-face surveys, the five studies on interviewer gender effects during the 1980s were all conducted over the telephone. Overall, the 1980s studies provide little evidence for the influence of interviewer gender effects on survey findings (Groves & Fultz, 1985; Hutchinson & Wegge, 1991; Johnson & Moore, 1993) or nonresponse (Groves & Fultz, 1985). When they did occur, interviewer gender effects were most likely to occur when respondents were answering gender attitudes items (Grimes & Hansen, 1984; Lueptow, Moser, & Pendleton, 1990), questions about the economy (Groves & Fultz, 1985), or non-gender-related political attitude items that may have invoked a gender-based stereotype threat (Groves & Fultz, 1985; Hutchinson & Wegge, 1991). Effects were associated with both male (Groves & Fultz, 1985; Hutchinson & Wegge, 1991; Lueptow et al., 1990) and female (Grimes & Hansen, 1984; Groves & Fultz, 1985; Lueptow et al., 1990) respondents, both of whom provided more socially progressive responses or responses that might be perceived more favorably by members of the interviewer's gender. The Groves and Fultz (1985) study was the only 1980s study to account for the clustering of respondents by interviewers.

The 1990s produced nine studies on interviewer gender effects, all but one of which (Pollner, 1998) applied techniques to account for interviewer clustering. Although most surveys found some evidence of interviewer gender effects, these effects were generally limited to only a few items in each survey. In those instances where effects emerged, it is difficult to see consistent patterns across studies. Interviewer gender effects were associated with factual (Dailey & Claus, 2001; Pollner, 1998), behavioral (Catania et al., 1996; Fendrich et al., 1999; Johnson & Parsons, 1994; Wilson, Brown, Mejia, & Lavori, 2002), and attitudinal (Huddy et al., 1997; Kane & Macaulay, 1993)

questions. Interviewer gender effects existed as both direct effects (Dailey & Claus, 2001; Huddy et al., 1997; Johnson & Parsons, 1994; Kane & Macaulay, 1993; Pollner, 1998) and interactions with respondent gender (Catania et al., 1996; Johnson & Parsons, 1994). And, interviewer effects occurred in both face-to-face (Dailey & Claus, 2001; Fendrich et al., 1999; Johnson & Parsons, 1994; Pollner, 1998; Wilson, Brown, Mejia, & Lavori, 2002) and telephone (Catania et al., 1996; Huddy et al., 1997; Kane & Macaulay, 1993) surveys.

Interestingly, all of the 1990s studies focused on either a gender-related topic or a sensitive topic such as substance use or mental health (see Table 2.2). The survey topic was explicitly gender-focused for two of the most rigorous of the 1990s studies, and, for each, both male and female respondents primarily provided responses that might be interpreted favorably by a person of their interviewer's gender (Huddy et al., 1997; Kane & Macaulay, 1993). These findings may indicate that respondents consciously edit their answer to an interviewer's gender for questions that overtly draw upon gender attitudes. In contrast, the impulse to edit one's responses may be contingent upon the interviewer-respondent gender combination involved for other types of topics. For instance, interviewer gender sometimes interacted with respondent gender when sexual behaviors were surveyed (Catania et al., 1996; Wilson et al., 2002). Although more research is needed, it appears that both male and female respondents may provide higher reports of some sexual behaviors to same-gender interviewers (Catania et al., 1996; Wilson et al., 2002). Female interviewers tended to obtain higher reports of psychiatric symptoms (Pollner, 1998) and sexual abuse (Dailey & Claus, 2001) from both male and female respondents. It may be that respondents feel more comfortable disclosing information to female interviewers about personal issues that much of the public might view as victimizing. In contrast, disclosure of behaviors that much of the public might view as behaviors of choice, such as illicit drug use, was higher among male interviewers (Fendrich et al., 1999; Johnson & Parsons, 1994) when conducted face-to-face but unaffected by interviewer gender when conducted over the telephone (Livert et al., 1998). In one of the few studies of respondent preferences, Catania and colleagues (1996) found that 82% of women and 72% of men stated a preference for interviewer gender at the beginning of a sexual behavior telephone survey. Among those respondents who stated a

preference, 94% of female respondents and 55% of male respondents preferred a female interviewer. However, these analyses did not control for the gender of the interviewer who offered the gender choice.

At least one study has explored the potential for gender to impact self-administered surveys. Tourangeau, Couper, and Steiger (2003) conducted an experiment in which the image accompanying a Web survey varied between a photo of a female investigator, a photo of a male investigator, or a logo with no investigator photo. In general, no significant effects were found for sensitive items such as cocaine use, alcohol consumption, sexual behavior, and church attendance. However, respondents viewing the female investigator's image were less likely to say they had voted in the last election and more likely to report pro-feminist attitudes than respondents viewing the male investigator. In a related experiment, Tourangeau and colleagues (2003) compared responses to many of the same sensitive behavior and gender attitude items from the first experiment among telephone respondents using interactive voice response (IVR). Whether the IVR voice was male or female had no significant impact on reporting of sensitive behaviors or gender attitudes. These studies suggest that the presence or stimulus of a male or female image or administrator may affect survey responses. However, the context of when and why these effects emerge requires further investigation with studies using multiple male and female personae.

Given the wide range of study designs, temporal contexts, survey topics, statistical approaches, and findings represented in the current body of interviewer gender research, it is difficult to draw any firm conclusions about the effect of interviewer gender upon survey data. It is unclear whether data collected by male or female interviewers is more accurate and if or when interviewers should be matched to respondent gender. However, the extant literature suggests that interviewer gender effects do occur and, as a result, inspires many avenues for future research that are of specific interest to public health professionals. For one, more research is warranted on the types of topics that may be affected by interviewer gender. For instance, public health professionals frequently engage in the measurement of gender-related and sensitive behaviors, yet many questions remain: Which types of gender-related or sensitive behavior reporting are affected by interviewer gender? Are different types of

sensitive behaviors differentially affected by interviewer gender? Do some sensitive topics invoke direct effects while others are conditional upon the gender make-up of the interviewer-respondent dyad? And, are attitudes and beliefs regarding sensitive behaviors also impacted by interviewer gender? Public health professionals are also often engaged in measuring behaviors, attitudes, and beliefs that may be impacted not only by the strictly binary concept of male versus female, but also by variations in gendered identities. For instance, do respondents provide the same answers to gender- or sex-related questions to interviewers who are heterosexual, lesbian, bisexual, gay, or transgendered? Does an interviewer's gender identity influence other types of survey questions? How do interviewer and respondent gender identities interact during the survey process? These and other aspects of interviewer gender effects warrant additional investigation.

INTERVIEWER AGE EFFECTS

It is conceivable that the perceived age of an interviewer might affect a respondent's answers to survey questions, particularly if the questions invoke reporting about attitudes or behaviors associated with aging or issues that might be generational. Unfortunately, few published studies have evaluated interviewer age effects in the United States. This area of interviewer error research may be understudied; however, the sparsity of published research may also emanate from a lack of statistically significant findings.

To date, interviewer age effects research suggests that older interviewers are associated with higher ratings of respondent frankness and honesty (Benney et al., 1956), higher indicators of adolescent independence (Ehrlich & Riesman, 1961), higher ρ^*_{int} values (Freeman & Butler, 1976), lower survey nonresponse (Singer et al., 1983), both higher (Johnson & Parsons, 1994) and lower (Fendrich et al., 1999) reports of marijuana use, less knowledge about sex among Hispanic female youth (Ford & Norris, 1997), higher and lower reporting of specific sexual behaviors (Wilson et al., 2002), and no

effects for a variety of items (Dailey & Claus, 2001; Fendrich et al., 1999; Ford & Norris, 1997; Freitag & Barry, 1974; Johnson & Parsons, 1994; Wilson et al., 2002).

However, interpretation of the interviewer age effects literature is obscured by several limitations. As indicated in Table 2.3, only two studies (Fendrich et al., 1999; Johnson & Parsons, 1994) employed the same age cutoffs to define their interviewer age groups for analysis, making it difficult to compare results across studies. Further, many studies are limited by the dichotomization of interviewer age into only two age groups (Benney et al., 1956; Fendrich et al., 1999; Freeman & Butler, 1976; Johnson & Parsons, 1994; Wilson et al., 2002). This dichotomization renders interviewers of vastly different generations to be statistically equivalent, yet it is hard to believe that interviewer age effects would be the same, for example, for 35-year-old and 60-year-old interviewers. These studies also vary widely in the age, racial, cultural, and gender compositions of their samples, as well as in the survey topics studied, modes of administration, and outcomes of interest. Some studies focus on nonresponse, while others measure variance or reporting levels of sensitive behaviors. These differences may become strengths as the corpus of research on interviewer age effects grows. But, with so few studies to date, identifying patterns across such disparate studies is chancy at best. Like other areas of interviewer effects research, interviewer age research is also limited by the lack of accounting for data clustering and random interviewer assignment.

Given these many limitations, no interviewer age effects trends may be confidently proposed at this juncture. Much additional research is needed, and the reporting of findings, whether significant or not, would be instrumental in advancing research. Future research should employ random interviewer assignment and rigorous statistical approaches to adjust for data clustering and possible confounders such as interviewer experience, gender, race, and interpersonal skills. Although difficult to obtain in practice, future research should also aspire to involve large numbers of respondents and interviewers representing wide age ranges. Additional research should consider the possibility of both direct and indirect interviewer age effects. Interviewer age may well interact with respondent characteristics such as age, gender, race, and ethnicity. If possible, analyses should define age as either a continuous variable or as a categorical variable measuring respondent-interviewer age distance following the lead of

Dailey and Claus (2001). Additional telephone-administered studies should also be considered, as only one study has reported on interviewer age effects in a telephone survey (Singer et al., 1983). Future telephone studies should assess perceived age as a possibly more reliable predictor, as interviewer age may be more difficult for a respondent to accurately determine over the telephone than other characteristics such as gender and race. Lastly, interviewer effects research as a whole would benefit from studies on interviewer age effects in surveys that explicitly query respondents about age-related topics. No existing study has focused on age-related questions. Yet, research on interviewer race, ethnicity, and gender effects suggest that whether a survey item overtly queries these characteristics may be related to increased interviewer error. In sum, the only conclusion that may be formed at present is that there is insufficient evidence to either support or advocate against the matching of interviewers and respondents by age.

OTHER INTERVIEWER CHARACTERISTICS

Although interviewer race and ethnicity, gender, and age have been the targets of most empirical investigations of interviewer effects, several other sociodemographic characteristics have been examined. These characteristics tend to be less perceptible to respondents and have generally been investigated in only a few studies. For instance, the two published studies on interviewer education effects indicate that interviewer education has no impact on survey response rates (Singer et al., 1983) or reporting of physical and sexual abuse (Dailey & Claus, 2001). Additional research on interviewer education is warranted, particularly across samples of respondents with varied educational distributions.

More experienced interviewers have been associated with fewer field coding errors (Stember & Hyman, 1949-1950), more valid data (Feldman et al., 1951-1952), more numerous open-ended item responses (Feldman et al., 1951-1952), higher consent rates for obtaining medical records (Cleary, Mechanic, & Weiss, 1981), higher reporting of psychological symptoms (Cleary et al., 1981; McGlone, Aronson, & Kobrynowicz, 2006), lower survey response rates (Singer et al., 1983), and lower survey administration

times (Olson & Peytchev, 2007). As with interviewer age effects studies, however, interviewer experience studies are difficult to compare because such different definitions of "high" or "low" experience exist across studies, and the ranges of experience levels within the categories used for analysis may be too broad. Interviewer experience effects are also difficult to study, as field administrators often assign more experienced interviewers to those cases where they expect to encounter less cooperation. Further, the very nature of what constitutes experience is in question. For example, does market research or polling work count as interviewing experience, or must prior experience be strictly classified as research survey work? And, does face-to-face interviewing experience translate to telephone interviewing experience and vice versa? Lastly, as both Cleary (1981) and McGlone (2006) and colleagues have noted, interviewing experience may be related not only to interviewers' skill levels but also to interviewers' enjoyment and interest in their work. Thus, experience may be confounded with other interviewer characteristics. Additional investigation of interviewer experience effects is clearly needed.

Several studies have investigated the effects of interviewers' social status on survey responses (Benney & Geiss, 1950; Dohrenwend, Colombotos, & Dohrenwend, 1968; Feldman et al., 1951-1952; Freeman & Butler, 1976; Katz, 1942; Williams, 1964). However, social status is not included in this review because it is an elusive characteristic to define and an even more evasive characteristic to measure. Interviewer social status has typically been measured as a combination of other characteristics such as education, income, and race. It would be better to measure these characteristics independently in lieu of introducing the entropy that invariably accompanies the creation of a social status variable. In contrast, constructed indices of social distance, or the degree of sociodemographic similarity between interviewers and respondents, can be useful, as they have the potential to advance interviewer effects theory by exploring why interviewer effects occur.

FUTURE DIRECTIONS

Public health professionals have a strong need to be knowledgeable about the impact of interviewer effects on health-related survey data. Although interviewer effects appear not to affect most survey items, research to date indicates that interviewer effects such as race, ethnicity, gender, and age do sometimes occur in all interviewer-administered survey modes and can significantly alter survey findings. If these effects are replicated across surveys due to the repeated application of a flawed survey design, such as the consistent use of an all-Black or all-female interviewing staff when interviewer race or gender effects are present, then bias will result. This bias, in turn, may impact findings across studies and the general understanding of a health issue.

More is unknown about interviewer effects than is known. Existing research suggests that interviewer race and ethnicity effects are most likely to occur when survey items query race- or ethnicity-related attitudes and that interviewer gender effects may be more likely to occur for questions about gender attitudes. However, even these conclusions may be moot for future studies, as race, ethnicity, and gender relations in the United States are continuously changing over time. No patterns can be confirmed for interviewer age effects. Additional studies are required that query age-related topics and enable comparisons of age effects across studies.

In addition to further investigation of interviewer race, ethnicity, gender and age effects, future research should consider other, more unexplored interviewer characteristics. For instance, interviewer income level may be apparent face-to-face surveys, but this issue has not been explored in isolation of social class. Research is also needed to examine whether racial and ethnic identity or sexual orientation affect survey data. Telephone survey researchers should compare perceived race, ethnicity, gender, and age to actual interviewer characteristics. When possible, interactions among multiple interviewer qualities should also be engaged to explore the additive effects of sociodemographic characteristics such as race and gender, age and gender, race and gender and age, etc. Additional research is warranted to investigate all types of interviewer effects in health topics, including the exploration of interviewer gender effects among non-racial sensitive topics such as sexual behavior, substance use, sexual

abuse, and mental health. Researchers should also be cognizant of geographical context. For example, race relations are likely to be different in the Southern U.S. than in the Northern U.S., and gender attitudes may be different in urban versus rural areas. Further research on data validity is also of the utmost importance, as evidence to date is lacking about whether respondents provide more valid data to sociodemographically similar or dissimilar interviewers. For some topics, such as racial attitudes, it may be appropriate to adopt the perspective of Rhodes (1994) that "truth" is not static, but, rather, varies according to context. Research is likewise needed on which respondents feel more comfortable with, trust, and prefer interacting with homophilous or heterophilous interviewers, as these factors may influence the data that respondents provide.

Given the current emphasis on increasing cultural sensitivity in public health research and practice, public health professionals need more empirical guidance on whether to match interviewers and respondents on varied sociodemographic characteristics. As per the recommendations of Groves (2004), future research will ideally include large sample sizes, large corps of interviewer, interpenetrated survey designs, and statistical techniques capable of accounting for the clustering of respondents by interviewers. These ideal study qualities may be difficult to achieve in many public health survey settings. However, even if lacking perfect design, the repeated investigation and reporting of interviewer effects, whether significant or null, will contribute to a significantly enhanced understanding of the magnitude and frequency of interviewer effects in public health research and practice over time.

Health practitioners should be cognizant that they may find themselves working in very specific cultural contexts, whether in the United States or abroad, that may bear upon the presence of interviewer effects. For example, Becker and colleagues (1995) found significant interviewer gender effects for a family planning survey of women in Nigeria. Interviewer gender effects were strongest in the Islamic, more culturally conservative state of Kano, where a male partner or chaperone was present for 79% of interviews administered by male interviewers but in only 12% of interviews administered by female interviewers. In contrast, a male was present for 10% of interviews conducted by males in the more liberal state of Imo but for 19% of interviews conducted by females. Differences in item refusals and variance associated with study parameters also varied by

gender across states. Although interpretation of these results must be tempered by the small number of interviewers utilized and the lack of adjusting for data clustering, these findings underscore the need for health professionals to consider the cultural context of survey data collection. For instance, in some cultures, certain types of interviewers or survey topics may necessitate the presence of a male chaperone to gain access to interviewing women, and the presence of an observer may moderate the survey responses obtained. In addition, what is considered a sensitive topic in one culture may not be considered sensitive, or sensitive to the same degree, in another culture. In a study in rural Nepal (Axinn, 1991), male interviewers reported that they would be perceived as making a sexual advances if they asked female respondents if they were pregnant and voiced fears of being attacked by respondents' husbands if they administered a pregnancy survey item. As a result, many male interviewers in the study skipped asking this question. These data indicate that as sex roles vary across cultures, interviewer gender effects are also likely to vary. In other cultural contexts, the socioeconomic status of the interviewer may be particularly influential. For example, the effect of interviewer social status may be more pronounced in cultures with more observable social class hierarchies. Cultures also vary in their valuation of age, resulting in the potential for varying interviewer age effects in cultures where elders are more or less esteemed. Race and ethnicity are influential social characteristics in most cultures, but their exact effects are likely to vary by the race or ethnicity in question from culture to culture. Cultures also vary in their mistrust of and attitudes toward strangers (Weinreb, 2006). Research on cross-cultural interviewer effects is still in its infancy, but much progress is anticipated in the next decade.

Public health professionals have an interest in evaluating interviewer effects because interviewer characteristics may interact with those characteristics most often used to target populations for public health research and practice. Public health programs frequently define their target populations by race, ethnicity, gender, age, and other sociodemographic characteristics. As a consequence, public health surveys frequently include topics related to these same sociodemographic qualities. Examples include the measurement of racial identity, culturally related health attitudes, sexual behaviors, and aging. If interviewer characteristics impact survey data, then interviewer effects may

have a particularly strong role in public health data collection. Further, at least one recent study indicates that African Americans, women, the aged, and persons with lower income and educational levels are likely to believe that their ability to project a positive self-image in health care settings is an important factor in securing quality medical care (Malat, van Ryn, & Purcell, 2006). This finding emphasizes an aspect of the survey context that may be unique to public health: surveys often provide access to appropriately targeted health education programs that may have real consequences for respondents' future well-being and mortality. Even with the advent of new technologies, interviewer-administered surveys are likely to persist. However, without fully understanding what types of interviewer-respondent pairings result in the most accurate data for different survey designs, public health research and practice may be adversely affected by undetected measurement error. With these risks in mind, public health professionals should seize opportunities to mitigate, measure, and control for interviewer error. In so doing, they will rise to Hyman's challenge and contribute to the advancement of public health science.

Table 2.1: Interviewer Race and Ethnicity Effects Studies

Authors	Date*	Mode	Number of Interviewers	Racial/ Ethnic Composition of Interviewers	Number of Respondents	Race or Ethnicity of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Race or Ethnicity Attitude Items	Survey Topic
Lenski & Leggett	1958	FTF	NR	White	624	Black and White	NR	No	No	Anomie
Summers & Hammonds	1960	FTF	NR	Black and White	NR	White	NR	No	Yes	Racial attitudes
Williams	Pub. 1964	FTF	22	13 Black; 9 White	840	Black	Yes	No	Yes	Racial attitudes, political participation, social group participation
Williams	Pub. 1968	FTF	21	12 Black; 9 White	452	Black	Yes	Yes	Yes	Racial attitudes, media use
Carr	1968	FTF	6	4 Black; 2 White	151	Black	NR	No	No	Anomie
Schuman & Converse	1968	FTF	42	25 Black; 17 White	619	Black	Yes	No	Yes	Racial attitudes
Hatchett & Schuman	1971	FTF	16	9 Black; 7 White	106	White	Yes	Yes	Yes	Racial attitudes
Welch, Comer & Steinman	Pub. 1973	FTF	6	3 Mexican American; 3 Anglo	178	Mexican American	NR	No	No	Health care attitudes, health behaviors, political attitudes, political behavior
Freitag & Barry	Pub. 1974	FTF	54	Black and White	724	White	NR	No	No	Life satisfaction, interviewer attitudes

* If known, the year that the survey data were collected is listed. If unknown, the date that the study was published (Pub.) is listed.

** Note: There is likely a small degree of error in these classifications, as the space limitations of publications often prohibit sufficient description of study methodologies. Any misrepresentation of these studies is regretted.

Abbreviations used in table: NR = not reported; FTF = face-to-face; CASI = computer assisted self-interviewing

Table 2.1: Interviewer Race and Ethnicity Effects Studies, Continued

Authors	Date*	Mode	Number of Interviewers	Racial/ Ethnic Composition of Interviewers	Number of Respondents	Race or Ethnicity of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Race or Ethnicity Attitude Items	Survey Topic
Campbell	1974	FTF	12	4 Black; 8 White	944	Black and White	No	No	Yes	Political attitudes, racial attitudes, parent attitudes
Schaeffer	1972-1977	FTF	NR	Black and White	Varies	Black and White	No	Yes	Yes	Racial attitudes
Cotter, Cohen, & Coulter	Pub. 1982	Phone	12	3 Black; 9 White	542	Black and White	Yes	No	Yes	Racial attitudes, nonracial questions
Singer, Frankel, & Glassman	Pub. 1983	Phone	35	28 Black; 7 White	1,014	NR	No	Yes	No	Substance use, sex, leisure activities, mental health, opinions of surveys
Anderson, Silver, & Abramson (1988b)	1964, 1976, 1978, 1980, 1984	FTF	NR	Black and White	1,389	Black	No	No	No	Voting behavior, political attitudes
Reese, Danielson, Shoemaker, Chang, & Hsu	1984	Phone	15	11 Hispanic; 4 Anglo	1,004	Hispanic and Anglo (includes White and Black)	Yes	No	Yes	Ethnicity, ethnic/cultural attitudes, media us, ethnic social interaction
Davis (1997b)	1984	Phone	76	27 Black; 49 White	1,150	Black	Yes	No	Yes	Political attitudes and party affiliations; racial attitudes
Davis (1997a)	1984	Phone	76	27 Black; 49 White	1,150	Black	Yes	No	Yes	Political attitudes and party affiliations

Table 2.1: Interviewer Race and Ethnicity Effects Studies, Continued

Authors	Date*	Mode	Number of Interviewers	Racial/ Ethnic Composition of Interviewers	Number of Respondents	Race or Ethnicity of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Race or Ethnicity Attitude Items	Survey Topic
Finkel, Guterbock, & Borg	1989	Phone	NR	33% Black; 66% White	252	White	Yes	No	Yes	Political attitudes
Johnson & Parsons	1990	FTF	14	9 Black; 5 White	481	Black and White	No	Yes	No	Substance use
Fendrich et al., 1999	1992	FTF	127	36 Black; 74 White or "other"; 18 Hispanic	3,978	NR	No	Yes	No	Drug use
Wolford, Brown, Marsden, Jackson, & Harrison	1993	Phone	NR	NR	1,206	Black	NR	No	Yes	Political attitudes, racial attitudes
Zimmerman, Caldwell, & Bernat	1995-1996	FTF	15	Black and White	591	Black	NR	No	No	Health, social relationships, school attitudes, psychological distress
Dailey & Claus	1995-1997	FTF	22	Black and White	8,276	Black and White	NR	Yes	No	Physical and sexual abuse
Webster	Pub. 1996	FTF	79	40 Hispanic; 39 Anglo	NR	Hispanic and Anglo	No	Yes	Yes	Sociodemographics, culture, consumer product information
Livert, Kadushin, Schulman, & Weiss	1997	Phone	343	Black and White (numbers not reported)	12,872	Black, White, and Hispanic	Yes	Yes	No	Substance use and attitudes about substance use

Table 2.1: Interviewer Race and Ethnicity Effects Studies, Continued

Authors	Date*	Mode	Number of Interviewers	Racial/ Ethnic Composition of Interviewers	Number of Respondents	Race or Ethnicity of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Race or Ethnicity Attitude Items	Survey Topic
Davis & Silver	2001	Phone	NR	Black, White, Black and White, Other	855	Black and White	Yes	No	No	Political knowledge, political attitudes
Krysan & Couper	Pub. 2003	FTF and video CASI	4	2 Black; 2 White	160	Black and White	NR	No	Yes	Racial attitudes

Table 2.2: Interviewer Gender Effects Studies

Authors	Date	Mode	Number of Interviewers	Gender Composition of Interviewers	Number of Respondents	Gender of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Gender-Attitudes	Survey Topic
Benney, Riesman & Star	1950s	FTF	NR	Male and Female	4,708	Male and Female	NR	No	No	Political attitudes, mental health
Colombotos, Elinson, & Loewenstein	1960-1961	FTF	31	Male and Female	1,479	Male and Female	No	No	No	Psychiatric symptoms
Landis	Pub. 1973	FTF	2	Male and Female	90	Female	Yes	No	Yes	Feminist attitudes
DeLamater	Pub. 1974	FTF	NR	Male and Female	238	Male and Female	No	No	No	Sexual attitudes and behavior
Johnson & DeLamater	Pub. 1976	FTF	24	5 Male; 19 Female	1,361	Male and Female	Yes	No	No	Sexual attitudes and behavior
Freeman & Butler	Pub. 1976	FTF	33	Male and Female	2,600	Male and Female	Yes	Yes	No	Mental retardation
Groves & Magilavy	1979 (Pub. 1986)	Phone	30	Male and Female	954	Male and Female	Yes	Yes	No	Health, media use
Grimes & Hansen	1980	Phone	NR	Male and Female	240	Male and Female	NR	No	Yes	Gender role attitudes
Groves & Fultz	1981-1983	Phone	120	40 Male; 80 Female	7,300	Male and Female	Yes	Yes	No	Consumer attitudes
Hutchinson & Wegge	1984	Phone	26	11 Male; 15 Female	795	Male and Female	NR	No	No	Political attitudes
Johnson & Moore	1989	Phone	22	7 Male; 15 Female	449	Male and Female	Yes	No	No	Attitudes about pornography

* If known, the year that the survey data were collected is listed. If unknown, the date that the study was published (Pub.) is listed.

** Note: There is likely a small degree of error in these classifications, as the space limitations of publications often prohibit sufficient description of study methodologies. Any misrepresentation of these studies is regretted.

Abbreviations used in table: NR = not reported; FTF = face-to-face; IVR = interactive voice recognition

Table 2.2: Interviewer Gender Effects Studies, Continued

Authors	Date	Mode	Number of Interviewers	Gender Composition of Interviewers	Number of Respondents	Gender of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Gender-Attitudes	Survey Topic
Lueptow, Moser, & Pendleton	Pub. 1990	Phone	NR	Male and Female	432	Male and Female	NR	No	Yes	Gender role attitudes
Johnson & Parsons	1990	FTF	14	10 Male; 4 Female	481	Male and Female	No	Yes	No	Substance use
Kane & Macaulay	1990-1991	Phone	29	14 Male; 15 Female	1,749	Male and Female	No	Yes	Yes	Gender role attitudes
Catania, Binson, Canchola, Pollack, Hauck, & Coates	1992	Phone	40	15 Male; 25 Female	2,030	Male and Female	Yes	Yes	No	Sexual behavior
Fendrich et al.	1992	FTF	127	56 Male; 71 Female	3,978	Male	No	Yes	No	Drug use
Huddy et al.	1991 & 1993	Phone	52	22 Male; 30 Female	658	Male and Female	Yes	Yes	Yes	Feminist attitudes
Dailey & Claus	1995-1997	FTF	22	4 Male; 18 Female	8,276	Male and Female	NR	Yes	No	Physical and sexual abuse
Livert, Kadushin, Schulman, and Weiss	1997	Phone	343	42% Male; 58% Female	12,872	Male and Female	Yes	Yes	No	Substance use
Wilson, Brown, Mejia, & Lavori	1998-1999	FTF	66	24 Male; 42 Female	1,146	Male and Female	No	Yes	No	Sexual behavior
Pollner	Pub. 1998	FTF	112	62 Male; 50 Female	3,131	Male and Female	NR	No	No	Mental health

Table 2.2: Interviewer Gender Effects Studies, Continued

Authors	Date	Mode	Number of Interviewers	Gender Composition of Interviewers	Number of Respondents	Gender of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Gender-Attitudes	Survey Topic
Tourangeau, Couper, & Steiger	Pub. 2003	Web	2 (photos)	1 Male; 1 Female	3,047	Male and Female	Yes	NA	Yes	Sensitive behaviors, gender role attitudes
Tourangeau, Couper, & Steiger	Pub. 2003	Phone (IVR)	2	1 Male; 1 Female	1,022	Male and Female	Yes	NA	Yes	Sensitive behaviors, gender role attitudes

Table 2.3: Interviewer Age Effects Studies

Authors	Date	Mode	Number of Interviewers	Age Categories of Interviewers Used for Analysis	Number of Respondents	Age of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Age-Related Items	Survey Topic
Benney, Riesman & Star	1950s	FTF	NR	Under 40; 40+	4,708	Under 30 to over 40	NR	No	No	Political attitudes, mental health
Ehrlich & Riesman	1956	FTF	97	24-40; 41-52; 53+	NR	18 and under	Yes	No	No	Independence
Freeman & Butler	Pub. 1976	FTF	33 (29 with known ages)	Under 32; 32+	2,600	The "aged"	Yes	Yes	No	Mental retardation
Freitag & Barry	Pub. 1974	FTF	54	NR	724	NR	NR	No	No	Life satisfaction, interviewer attitudes
Singer, Frankel, & Glassman	Pub. 1983	Phone	35	18-21; 22-34; 35+	1,014	NR	Yes	Yes	No	Substance use, sex, leisure activities, mental health, opinions of surveys
Johnson & Parsons	1990	FTF	14	Under 35; 35+	481	NR	No	Yes	No	Substance use
Ford & Norris	1991	FTF	60	21-30; 31-40; 41+	1,435	15-24	No	Yes	No	Sexual behavior
Fendrich et al.	1992	FTF	127	Under 35; 35+	3,978	8-20	No	Yes	No	Drug use

* If known, the year that the survey data were collected is listed. If unknown, the date that the study was published (Pub.) is listed.

** Note: There is likely a small degree of error in these classifications, as the space limitations of publications often prohibit sufficient description of study methodologies. Any misrepresentation of these studies is regretted.

Abbreviations used in table: NR = not reported; FTF = face-to-face

Table 2.3: Interviewer Age Effects Studies, Continued

Authors	Date	Mode	Number of Interviewers	Age Categories of Interviewers Used for Analysis	Number of Respondents	Age of Respondents	Random** Assignment of Interviewers	Effort** to Account for Interviewer Clustering	Queried Age-Related Items	Survey Topic
Dailey & Claus	1995-1997	FTF	22	4 categories representing closeness to respondents' ages	8,276	Average age = 33.2 years	NR	Yes	No	Physical and sexual abuse
Wilson, Brown, Mejia, & Lavori	1998-1999	FTF	66	Under 30; 30+	1,146	Men: 18-40; Women: under 18-40	No	Yes	No	Sexual behavior

CHAPTER 3

DO AFRICAN AMERICAN TELEPHONE SURVEY RESPONDENTS PREFER AFRICAN AMERICAN INTERVIEWERS?

INTRODUCTION

Public health survey designers often assign African American interviewers to African American respondents. This practice is particularly prevalent when surveys query race-related topics. The rationale for matching is likely based on three related factors: (1) mistrust of health research among African Americans (Achter, Parrott, & Silk, 2005; Corbie-Smith, Thomas, Williams, & Moody-Ayers, 1999; Gamble, 1997); (2) widespread emphasis on the importance of cultural sensitivity in health research and programming; and (3) evidence of race of interviewer effects on survey data collected from African American respondents (see Chapter 2). Many researchers assume that race matching will reduce mistrust, put respondents at ease, and, overall, yield more valid data. Yet, there is surprisingly little evidence regarding the validity of these assumptions. As a consequence, several survey researchers have questioned the appropriateness of race matching (Anderson et al., 1988b; Aspinall, 2001; Groves, 2004; Heeb & Gmel, 2001; Rhodes, 1994).

One question that has not been asked is: Do African American respondents themselves have preferences regarding interviewer race? No studies querying African American respondents' preferences for interviewer race could be found in searches of major research databases such as PubMed, ISI Web of Science, or PsychInfo in Spring 2008. However, one study (Warnecke et al., 1997) explored respondents' reporting of whether they thought others from their same cultural group would be comfortable answering questions about alcohol consumption with an interviewer from a different cultural group. Over 90% of White respondents reported that other Whites would be

comfortable with an interviewer from a different cultural group whereas only 60% of African Americans reported that other African Americans would be comfortable with non-African American interviewer. These findings suggest that African Americans may have stronger preferences for interviewer race than White respondents. But, additional research is clearly needed.

It is possible that African Americans' interviewer race preferences correlate with their ethnic identity orientations. Ethnic identity is defined by Cokley (2007) as "the subjective sense of ethnic group membership that involves self-labeling, sense of belonging, preference for the group, positive evaluation of the ethnic group, ethnic knowledge, and involvement in ethnic group activities". African Americans' feelings about being African American are theorized to vary widely (Cross, 1991; Sellers et al., 1998). No studies have evaluated ethnic identity as a correlate of interviewer race preferences among African American survey respondents. But, findings from related research in counseling indicates that African Americans with stronger ties to Afrocentric or Black American culture may be more likely to express a preference for same-race counselors than African Americans with weaker ties to Afrocentric or Black American culture (Atkinson, Furlong, & Poston, 1986; Morten & Atkinson, 1983; Parham & Helms, 1981). In a similar vein, African Americans with Afrocentric or Black American identity types may prefer to be surveyed by African American interviewers. Individuals with these identity orientations may feel more comfortable with and trusting of interviewers of their same race. According to Nigrescence Theory (Cross, 1991), some African Americans may be considerably angry toward and mistrusting of Whites. These attitudes are embraced in the concept of cultural mistrust, which, as defined by Terrell and Terrell (1981), is a generalized mistrust of Whites. Therefore, although cultural mistrust is not an ethnic identity orientation, it may be likely to correlate with survey respondents' self-identification with African American people and culture, and persons with high cultural mistrust may be wary and non-disclosing to White interviewers. Matching survey interviewers and respondents by race alone may be contraindicated for other African American survey respondents. At least one study of counselor preferences indicates that African Americans with more bicultural or multicultural identity orientations may be more likely to report no preference for counselor race than to endorse

a preference for an African American counselor (Morten & Atkinson, 1983). African Americans with more bicultural or multicultural identity orientations may likewise have no interviewer race preferences. People with a bicultural outlook tend to equally value their ability to interact with African Americans and White Americans alike, while individuals with a multicultural perspective may view themselves as being equally comfortable with persons from various racial and ethnic groups. Cross (1991) contends that some African Americans place little value on racial identity and, therefore, their own identity as African Americans. These persons have low racial salience and construct their social identities using other labels such as "American", "wife", "father", "engineer", "student", "community member", etc. According to Cross, some individuals in this group may even have adopted negative stereotypes and opinions about African Americans as a group. These individuals have low racial salience and high endorsement of White, majority culture. It is possible that these African American survey respondents prefer White interviewers, since they may view White interviewers as more professional, competent, or trustworthy. Yet, despite the possibility that interviewer race preferences may vary among African American respondents in many ways, no research has been published on this topic.

This study explores the influence of ethnic identity on African American respondents' preferences for interviewers of their same race and ethnicity in telephone-administered surveys. The telephone context may be a more conservative test of respondent preferences, as it is likely that interviewer characteristics are more influential in face-to-face survey administrations. This study tests the premise that African Americans with more Afrocentric, Black American, and culturally mistrustful identity orientations will be more likely to say that it is important to them to have a telephone interviewer of their same racial and ethnic background than African Americans with weaker or less exclusive ties to African American people and culture. This study also examines whether African Americans with Afrocentric, Black American, and culturally mistrustful identity types report lower hypothesized comfort with a White interviewer. Lastly, this study explores whether African American respondents' preferences for interviewer race vary according to whether a survey contains questions that specifically query racial or ethnic topics. Identifying whether interviewer race matching is important

to respondents, to whom interviewer race matching matters, and when interviewer race preferences are strongest may have considerable implications for public health research and practice.

METHODS

Study Sample

This study was conducted as part of a larger intervention trial called *Eat for Life*, which explored the efficacy of using tailored health education materials to increase fruit and vegetable consumption among African American adults. The *Eat for Life* trial was comprised of two related studies. The first study tested the tailoring of health materials on motivational constructs derived from Self-Determination Theory and Motivational Interviewing. This research is described elsewhere (Resnicow et al., 2008). The second study compared increases in fruit and vegetable intake between participants who received health materials tailored to their ethnic identity type versus participants who received health materials targeting a general African American audience. Details about the design and results for this study are also found elsewhere (Resnicow et al., under review).

Participants in the current study consisted of baseline telephone survey respondents in the *Eat for Life* motivation and ethnic identity studies. Since the experimental and control groups within each study completed the sample baseline survey, experimental and control group participants are combined within each study for the purposes of the present analyses.

The *Eat for Life* sample was recruited from the memberships of two integrated healthcare delivery systems serving predominantly urban residents in Detroit and Atlanta, respectively. In Detroit, the sample consisted of randomly selected healthcare system members whose medical records identified them as African American. The Atlanta healthcare system did not include racial or ethnic information in its medical records. Therefore, the Atlanta sample was randomly selected from healthcare system members with home addresses in Census blocks in which 80% or more of the residents were African American as of the 2000 Census. Recruitment letters containing pre-incentives

in the form of $2 bills were mailed to potential participants in one-week waves until enrollment targets were achieved. These letters were followed by recruitment telephone calls conducted between May 2006 and July 2007 for the motivation study and September 2006 and July 2007 for the ethnic identity study. During the eligibility portion of these calls, interviewers ensured that participants were between the ages of 21 and 70, self-identified as Black or African American and were not Hispanic or multiracial, ate fewer than ten servings of fruit and vegetables per day, were not currently hospitalized or living in skilled care facilities, had lived in the U.S. for more than half of their lives, and had no mental or physical conditions that would inhibit or be endangered by their participation in *Eat for Life*.

A total of 1,650 recruitment letters were mailed for the motivation study, yielding 533 completed baseline surveys by eligible respondents [American Association for Public Opinion Research (AAPOR) Response Rate 1 = 36.5%](American Association for Public Opinion Research, 2003). However, only the 272 motivation study respondents who were surveyed by African American interviewers were retained in the analyses. In the ethnic identity study, 2,018 recruitment letters were mailed, and 625 eligible participants completed the baseline telephone survey (AAPOR Response Rate 1 = 34.5%). Of these 625 participants, 617 respondents had sufficiently complete data to be included in the present analyses.

All *Eat for Life* participants who completed a baseline survey received a thank-you letter containing a $5 bill within a couple of weeks of completing the survey and three tailored newsletters over the next three months. Participants were subsequently contacted to complete a follow-up survey and receive an additional incentive of a $15 gift card to a retail store.

Eat for Life respondents were randomly assigned to interviewers in both studies. All of the ethnic identity study interviewers were African American, and all ethnic identity study respondents were told that their interviewer was African American. About half of the motivation study interviewers were African American. Only those motivation study respondents who had an African American interviewer were told that their interviewer was African American. All respondents who were informed that their interviewer was African American heard the following language near the beginning of

the introductory recruitment script: "I am calling as part of a team of African American interviewers …" Motivation study respondents whose interviewers were not African American received the following language as a substitute: "I am calling to see if you qualify …" All ethnic identity study respondents and those motivation study respondents with African American interviewers were further cued to their interviewer's race by the interviewers use of "our community" near the end of the introductory recruitment script. Ethnic identity study respondents heard two more references to "our community" during the introduction to the section of the survey in which ethnic identity was assessed. The ethnic identity section was located near the end of the survey.

This study was approved by human subjects review committees at the University of Michigan and the participating integrated healthcare delivery systems in Detroit and Atlanta.

Measures

Ethnic Identity – Ethnic identity was only measured in the *Eat for Life* ethnic identity study. Participants in the ethnic identity study were classified into one of 16 ethnic identity types using the Black Identity Classification Scale (BICS) (Davis et al., under review). The BICS contains 32, 7-point "Strongly Disagree"/"Strongly Agree" items, a single-item racial salience question, and two health material preference questions. The 7-point questions include 14 items created by the *Eat for Life* study team, 9 items adapted from the Survey of Black Life (Resnicow, Soler, Braithwaite, Selassie, & Smith, 1999), 5 items from or adapted from the Multidimensional Inventory of Black Identity (Sellers, Rowley, Chavous, Shelton, & Smith, 1997), 3 items created by Black identity experts, and 1 item adapted from the Cultural Mistrust Inventory (Terrell & Terrell, 1981). These 32 items measure six core identity component subscales: Afrocentric, Black American, Bicultural, Multicultural, Racial Salience, and Cultural Mistrust. For the purposes of this paper, an ethnic identity "component" is defined as an aspect of one's ethnic identity orientation. The single racial salience question, which is separate from the Racial Salience subscale, asks "How important is being Black to your overall identity?" Responses range from zero ("Not at All Important") to ten ("Very Important"). This item

was used to assess racial centrality (Sellers, Rowley, Chavous, Shelton, & Smith, 1997), and, thereby, to classify respondents as Assimilated or non-Assimilated at an early stage in the algorithm. The two health material preference questions ask respondents to select the type of ethnic orientation they would most prefer for their health education program (e.g., a focus on Americans vs. Black Americans vs. peoples from around the world, etc.). When needed, responses to these items were used as tiebreakers in determining ethnic identity types. Together, the 32 BICS items classify respondents into one of 16 potential identity types:

- Assimilated
- Black American
- Black American with Cultural Mistrust
- Bicultural
- Multicultural
- Bicultural/Multicultural
- Black American/Bicultural
- Black American/Bicultural with Cultural Mistrust
- Black American/Multicultural
- Black American/Multicultural with Cultural Mistrust
- Afrocentric/Black American
- Afrocentric/Black American with Cultural Mistrust
- Afrocentric/Bicultural
- Afrocentric/Bicultural with Cultural Mistrust
- Afrocentric/Multicultural
- Afrocentric/Multicultural with Cultural Mistrust

A person with an Assimilated identity type is defined as having low racial salience and placing little importance on being a member of a racial or ethnic group. In contrast, being Black is a valued aspect of personal identity for respondents with Black American, Afrocentric, Bicultural, and Multicultural identity components. A person with a Black American component is connected to Black American people and culture while an Afrocentric person feels a strong connection to Africa. A Bicultural person tends to perceive the world as a Black/White dichotomy whereas a Multicultural person views the world as a collection of many valued cultures. Persons with Black American and Afrocentric identity components can have an additional Cultural Mistrust component, which is defined as a generalized mistrust of Whites and White society (Terrell & Terrell,

1981). Details about the development and performance of the BICS can be obtained elsewhere (Davis et al., under review). A listing of BICS items is included in this dissertation as Appendix A.

Interviewer Preferences – At the end of the *Eat for Life* baseline survey, ethnic identity study respondents were asked two questions about their preferences for interviewer race for a survey like the one they had just completed. The first item queried the importance of having an interviewer with a similar racial and ethnic background: "How important is it to you to be interviewed by an interviewer of your same race and ethnicity for a survey like this?" Response options ranged from one ("Not at All Important") to ten ("Very Important"). The second item explored predicted comfort if the interviewer had been White: "How comfortable would you have felt if this interview had been done by a White interviewer?" Response options ranged from one ("Not at All Comfortable") to ten ("Very Comfortable"). Motivation study respondents were asked the same two questions at the end of the motivation study baseline survey. However, the second question about comfort with a White interviewer was only asked if the motivation study interviewer was not White.

Racial Survey Content – As a result of the differing goals of the two studies, the two baseline surveys contained different proportions of items that queried respondents' racial or ethnic identification or attitudes. Excluding eligibility items and the two interviewer race questions at the very end of the survey, the ethnic identity survey contained 40 out of a total of 101 items that explicitly queried racial attitudes, preferences for an ethnically oriented health program, and preferred terminology to describe one's racial and ethnic affiliation (e.g., Black American, Black, African American, American, or another term). The motivation survey contained only one explicit racial attitude item out of a total of 109 survey items in the main body of the survey.

Other Measures – Several demographic characteristics were measured and used as control variables in the analysis. Control variables included *Eat for Life* respondents' gender ("Are you male or female?"), age ("How old are you?"), income ("Approximately

what was the total income of your household last year before taxes?”), and education (“What is the highest grade or degree you have completed?”).

Analysis Plan

Several researchers have noted the importance of measuring and controlling for the clustering of respondents by interviewers in survey data (Dijkstra, 1983; Fendrich et al., 1999). Failing to account for data clustering may result in type I error, because unadjusted models ignore between-interviewer variance and thereby underestimate measurement error (Dijkstra, 1983; Fendrich et al., 1999). All models presented in this paper adjust for the clustering of respondents by interviewers by using multilevel modeling to treat interviewers as a random sample of the potential samples of interviewers that could have been drawn from a larger, hypothetical pool of interviewers (Fendrich et al., 1999; Groves, 2004). This approach enables more confident generalizations of findings to hypothetical replications of surveys involving different samples of interviewers (Dijkstra, 1983).

In order to control for the clustering of data by interviewers, these analyses use the linear mixed modeling approached outlined by West, Welch, and Gałecki (2007). This approach permits the estimation of both fixed effects associated with data obtained from *Eat for Life* study respondents and random effects resulting from the assignment of respondents to interviewers in a single model. All analyses were conducted using SAS 9.1.3 for Windows (SAS Institute, 2002-2003). Linear mixed models were estimated using the SAS proc mixed procedure.

Because cell sizes did not permit separate analyses of the 16 BICS identity types, the 16 types were represented by six variables reflecting whether or not a respondent had each of the following identity components: Assimilated, Afrocentric, Black American, Bicultural, Multicultural, or Cultural Mistrust. A respondent was coded as “1” for each identity component that he or she had and “0” if that component was not part of his or her ethnic identity classification. Since the 16 BICS identity types allowed up to three identity components, each respondent could have up to three “1” codes for the six identity variables. No other variables were collapsed.

Linear mixed models were run to test six hypotheses. The six hypotheses and accompanying models are detailed as follows:

Hypothesis 1: Ethnic identity study respondents with Afrocentric, Black American, and Cultural Mistrust identity components will be more likely to state a preference for an interviewer of their same race and ethnicity than respondents without these identity components, while no preferences will be indicated between respondents with and without the Assimilated, Bicultural, or Multicultural identity components. *[**Model 1:** Dependent variable – importance of having an interviewer of one's same race and ethnicity (scored from 1-10); Independent variables – Assimilated, Afrocentric, Black American, Bicultural, Multicultural, and Cultural Mistrust identity components (each scored 0/1); Control variables – respondent gender (male/female), age (continuous), education (scored from 1-8), and income (scored from 1-7)]*

Hypothesis 2: Ethnic identity respondents with Afrocentric, Black American, and Cultural Mistrust identity components will be more likely to report that they would have felt less comfortable if their baseline survey had been conducted by a White interviewer than respondents without these components, while no preferences will be indicated between respondents with and without the Assimilated, Bicultural, or Multicultural identity components. *[**Model 2:** Dependent variable – comfort if the interviewer had been White (scored from 1-10); Independent variables –Assimilated, Afrocentric, Black American, Bicultural, Multicultural, and Cultural Mistrust identity components (0/1); Control variables – respondent gender (male/female), age (continuous), education (1-8), and income (1-7)]*

Hypothesis 3: Ethnic identity respondents with higher levels of racial salience will report stronger preferences for an interviewer of their same race and ethnicity than respondents with lower levels of racial salience. *[**Model 3:** Dependent variable – importance of having an interviewer of one's same race and ethnicity (1-10); Independent variable – single racial salience item (scored from 0-10); Control variables – respondent gender (male/female), age (continuous), education (1-8), and income (1-7)]*

Hypothesis 4: Ethnic identity respondents with higher levels of racial salience will report less hypothetical comfort if their interviewer had been White than respondents with lower levels of racial salience. *[**Model 4:** Dependent variable – comfort if the interviewer had been White (1-10); Independent variable – single racial salience item (0-10); Control variables – respondent gender (male/female), age (continuous), education (1-8), and income (1-7)]*

Hypothesis 5: Respondents who participated in the ethnic identity baseline survey will be more likely to say that they preferred an interviewer of their same race and

ethnicity than respondents who participated in the motivation baseline survey. *[**Model 5:** Dependent variable – importance of having an interviewer of one's same race and ethnicity (1-10); Independent variable – study affiliation (ethnic identity/motivation); Control variables – respondent gender (male/female), age (continuous), education (1-8), and income (1-7)]*

Hypothesis 6: Respondents who participated in the ethnic identity baseline survey will be less likely to say that they would have been comfortable with a White interviewer than respondents who participated in the motivation baseline survey. *[**Model 6:** Dependent variable – comfort if the interviewer had been White (1-10); Independent variable – study affiliation (ethnic identity/motivation); Control variables – respondent gender (male/female), age (continuous), education (1-8), and income (1-7)]*

An individual who gravitates toward one or two BICS identity types is likely to have weaker associations with the remaining identity types. As a consequence, there is some collinearity among the ethnic identity variables examined in Hypotheses 1 and 2. The ethnic identity variables modeled for Hypotheses 1 and 2 are binary; thus, the collinearity is better examined by measuring the bivariate correlations between respondents' scores on the BICS subscales for the core identity type components: Afrocentric, Black American, Bicultural, Multicultural, and Cultural Mistrust. The BICS does not contain an Assimilated subscale. But, since the Assimilated type is the only ethnic identity type defined as having low racial salience, the Racial Salience subscale score may be used as a proxy for Assimilated in computing these correlations. The correlations for the following pairings are all below 0.35 and considered weak associations: Afrocentric/Bicultural (0.07), Afrocentric/Cultural Mistrust (0.14), Black American/Bicultural (0.10), Black American/Multicultural (0.27), Black American/Cultural Mistrust (0.27), Bicultural/Multicultural (0.23), Bicultural/Racial Salience (0.08), Bicultural/Cultural Mistrust (-0.06), Multicultural/Cultural Mistrust (0.05), Multicultural/Racial Salience (0.32), and Racial Salience/Cultural Mistrust (24). The remaining four pairings have moderate correlations: Afrocentric/Black American (0.52), Afrocentric/Multicultural (0.47), Afrocentric/Racial Salience (0.57), and Black American/Racial Salience (0.65).

It should also be noted that the single racial salience item that was used as the main independent variable in Hypotheses 3 and 4 was also included in the algorithm used

to classify respondents into the original 16 BICS identity types. Thus, the independent variables for Hypotheses 1 and 2 slightly overlap with that for Hypotheses 3 and 4. The intent of Hypothesis 3 and 4 is to determine whether a single item could be substituted for the longer BICS scale in future survey administrations. The entire BICS scale may be richer in information about ethnic identity and, therefore, more desirable for certain survey applications. But, if significant, the single item may have more widescale applicability for survey designers wishing to briefly ascertain the likelihood that a survey respondent will have certain preferences for interviewer race and ethnicity.

In addition to the analyses described above, mean scores on the importance of having an interviewer of one's same race and ethnicity item and the predicted comfort level if the interviewer had been White item were calculated for ethnic identity respondents with each of the six identity components: Assimilated, Afrocentric, Black American, Bicultural, Multicultural, or Cultural Mistrust.

RESULTS

Description of Study Sample

Demographic characteristics of the *Eat for Life* ethnic identity and motivation study baseline survey respondents are listed in Table 3.1. Both samples were predominantly female with a mean age in the mid-to-upper 40s and an almost even split between the Detroit and Atlanta health care plans. Participants in both groups represented a range of educational attainment and income levels. Means on the single racial salience item were comparable at 8.0 and 8.2 for the ethnic identity and motivation groups, respectively. Ethnic identity type was not measured in the motivation survey. Among ethnic identity study respondents, the most prevalent ethnic identity component was Black American (54.8%), followed by Multicultural (45.5%), Bicultural (39.2%), Afrocentric (30.2%), Assimilated (13.0%), and Cultural Mistrust (11.7%), respectively (respondents may be classified as more than one type).

Ethnic Identity and Racial Salience as Correlates of Interviewer Preferences

The results of tests of significance for models associated with Hypotheses 1 through 4 are detailed in Table 3.2. All models controlled for respondent gender, age, education, and income.

The results support Hypothesis 1. Ethnic identity respondents were significantly more likely to say that they preferred an interviewer of their same race and ethnicity if they had an Afrocentric ($p = .02$) or Black American ($p = .0002$) identity component than if they did not. Respondents with a Cultural Mistrust ($p = .07$) identity component were borderline more likely to prefer an interviewer of their same race and ethnicity. Conversely, respondents with Assimilated, Bicultural, and Multicultural identity components were no more likely to express a preference for an interviewer of their same race and ethnicity for a survey like the ethnic identity baseline than those who did not. Respondent gender, age, education, and income had no significant impact on these results.

In contrast, there is no evidence to support Hypothesis 2. Ethnic identity type appeared to have no bearing on ethnic identity respondents' hypothesized comfort levels if their interviewers had been White. Respondent gender, age, and income were also nonsignificant. However, as educational level increased, respondents were significantly more likely to report that they would have been less comfortable if their interviewer had been White ($p = .02$).

Analysis of Hypothesis 3 indicates that as racial salience increased, ethnic identity respondents were significantly more likely to say they preferred an interviewer of their same race and ethnicity ($p = .001$). Gender, age, education, and income were not statistically significant. The model testing Hypothesis 4 suggests that racial salience did not correlate with respondents' expected comfort levels if their interviewer had been White. But, similar to the model testing Hypothesis 2, respondents with higher educational attainment were significantly more likely to report that they would have been less comfortable if their interviewer had been White ($p = .01$).

Although not shown, a model was also tested that included both the single racial salience variable and the six ethnic identity variables as correlates of the importance of

having an interviewer of one's same race and ethnicity. In this model, only the Afrocentric, Black American, and Cultural Mistrust variables were significant. As expected, the ethnic identity variables capture much of the same information that is contained in the single racial salience item.

Due to concerns regarding the assumptions of normality for the linear mixed models, the models for all four hypotheses were also run as generalized estimating equations (GEE) using the proc genmod procedure in SAS. These models permitted the treatment of the dependent variables as ordinal variables. The GEE analyses confirmed the pattern of results found in the linear mixed modeling approach.

All of the models tested included random effects to adjust for the clustering of respondents by interviewers. The inclusion of these random effects was significant for models testing Hypotheses 1 ($p = .01$) and 3 ($p = .003$) and borderline for the model testing Hypothesis 4 ($p = .10$). In these models, interviewer-level variance accounted for 2.2%, 3.1%, and 1.5% of the total variance, respectively. These numbers indicate that it was generally important to control for the clustering of respondent data by interviewers and that different results might have been obtained if interviewer-level variance had not been accounted for in the analyses.

Type of Survey as a Correlate of Interviewer Preferences

As shown in Table 3.3, results from the model testing Hypothesis 5 support the premise that ethnic identity study respondents were significantly more likely to state a preference for an interviewer of their same race and ethnicity than motivation study respondents ($p = .001$). Respondent gender, age, education, and income were nonsignificant.

Hypothesis 6 also appears to be supported by the data. Ethnic identity study respondents were more likely than motivation study respondents to say that they would have been less comfortable with a White interviewer ($p = .04$). Respondent gender, age, and income were nonsignificant in this model; however, education was borderline significant. As with Hypotheses 2 and 4, respondents were more likely to say that they

would have been less comfortable if their interviewer had been White as educational attainment increased (p = .07).

Findings for Hypotheses 5 were confirmed in analyses using GEE. However, GEE analysis for Hypothesis 6 did not yield significant findings for study type but did confirm a borderline result for education. Thus, the main finding from this final model may not be robust.

Both models included random effects to adjust for the clustering of respondents by interviewers. Tests for the inclusion of these random effects were significant for Hypothesis 5 (p = .006) and borderline for Hypothesis 6 (p = .06). These results indicate that it was important to adjust these models to account for interviewer-level variance and that different results might have been obtained if interviewer-level variance had not been estimated.

Mean Scores on Interviewer Race Preference Items by Ethnic Identity Component

Available responses to the importance of having an interviewer of one's same race and ethnicity survey item ranged from one to ten. Among ethnic identity study respondents, the mean for respondents with an Assimilated identity component was 4.4 (see Table 3.4). Respondents with a Bicultural identity component had a mean of 5.1. Multicultural respondents had a slightly higher mean of 5.2, followed by respondents with Black American and Afrocentric identity types, who had means of 6.5 and 6.6, respectively. Respondents with a Cultural Mistrust component had the highest mean at 7.2.

Mean scores on the survey item querying the predicted comfort level if the interviewer had been White followed a similar, but looser, pattern as those for the importance item. The means on the comfort item by ethnic identity component were as follows: 7.0 for Assimilated; 7.5 for Bicultural; 7.2 for Multicultural; 6.8 for Black American; 6.9 for Afrocentric; and 6.3 for Cultural Mistrust.

DISCUSSION

Findings from this study indicate that the ethnic identity types of African American telephone survey respondents are correlated with interviewer race preferences. Ethnic identity respondents who had an Afrocentric or Black American identity component were significantly more likely to express a preference for an interviewer who matched their race and ethnicity than respondents who did not have these identity components. Respondents with a Cultural Mistrust component were borderline more likely to prefer an African American interviewer. In contrast, no significant differences in interviewer race preferences were found between respondents with and without the Assimilated, Bicultural, or Multicultural identity components. These findings are consistent with prior research on Black racial and ethnic identity, which has demonstrated substantial within-group variability among African Americans in their feelings about personal race and ethnicity (Cross, 1991; Sellers et al., 1998). As with previous research on counselor race preferences, participants in this study with more pro-Black identity types had stronger preferences for interviewers who matched their race and ethnicity than identity types that, by definition, had weaker ties to African American people and culture.

Data from this study suggest that a single racial salience question can significantly correlate with respondents' preferences for an interviewer of their same race and ethnicity for a survey with racial content. Ethnic identity respondents with higher racial salience scores had higher preferences for an interviewer of their same race and ethnicity. The ethnic identity variables may provide richer information; however, the single racial salience item may have more practical applicability. If preferences are deemed important, one could ask respondents a single question to determine whether to match interviewers and respondents by race for an imminent or future survey interaction.

The two models that explored the role of ethnic identity and racial salience as correlates of respondents' estimated comfort levels if their interviewer had been White yielded no significant main effects. It is possible that this lack of main effects is attributable to social desirability. Interestingly, both models found that the educational level of respondents was a significant correlate with their comfort ratings. As educational level increased, respondents were more likely to report that they would have been less

comfortable if their interviewer had been White. Since all respondents included in these analyses had African American interviewers, their responses to the comfort question were based on the hypothetical case of having had a White interviewer. Asking respondents about a hypothetical situation may have yielded less valid data. However, it is also possible that asking respondents about their comfort levels with a White interviewer was a more sensitive question than asking them about their preferences for an African American interviewer. Whereas the latter question provides a respondent with an opportunity to voice affinity for members of one's racial and ethnic group, asking a respondent to comment on their feelings about interacting with Whites may be more likely to expose the respondent to expressing opinions that may be interpreted as racist. If true, then respondents may have been motivated to adjust their answers to provide more socially desirable responses. Some evidence for this notion may be derived from qualitative data that were collected from *Eat for Life* participants at the end of the baseline surveys. Respondents were asked to share their overall thoughts about the survey and to note whether there were any questions they found to be difficult or disturbing. Two respondents said they found the question about the White interviewer difficult or disturbing, and a third respondent commented: "I don't want to sound like a racist." Narayan and Krosnick (1996) examined the moderating effects of the educational level of respondents in a meta-analysis of response effects of data from over 130 survey experiments. In ten of the eleven experiments that explored acquiescence bias, Narayan and Krosnick found that respondents with lower education were more likely to acquiesce than respondents with higher education. If the survey item querying comfort with a White interviewer was a particularly sensitive item, then, in line with the findings of Narayan and Krosnick, respondents with more education may have been less prone to acquiesce to this item and more comfortable expressing a less favorable opinion about White interviewers.

This study also provides evidence that whether a survey contains explicitly racial content impacts respondents' opinions about interviewer race. The motivation survey contained almost no explicitly racial content, whereas 38% of the ethnic identity survey contained explicitly racial content. All of the ethnic identity and motivation respondents who were included in this paper were surveyed by African American interviewers. Yet,

the ethnic identity respondents were significantly more likely than the motivation respondents to say that they preferred an interviewer of their same race and ethnicity and that they would have been less comfortable if their interviewer had been White.

Overall, these findings lead to one overarching question: Do respondents' preferences about interviewer race matter? This question can be considered from two perspectives. From a measurement error perspective, it is important to know whether obliging respondents' interviewer race preferences decreases interviewer error. This question cannot be explored in this study because ethnic identity respondents were interviewed by African American interviewers in all *Eat for Life* data collection interactions. Interviewers of varying races and ethnicities were employed in the motivation study, but there were too few motivation respondents in each interviewer pairing category to explore differences in survey outcomes between respondents who did and did not have an interviewer at post-test who matched their stated interviewer race preference at baseline. This line of study merits further exploration. The second perspective considers respondent satisfaction with the survey experience. Catania et al. (1996) explored this issue in a telephone survey of sexual behavior by allocating respondents to three experimental conditions: (a) assignment of a gender-matched interviewer, (b) assignment of a gender-discordant interviewer, or (c) asking respondents to choose their preferred interviewer gender at the beginning of the survey. Respondents in the choice condition were significantly less likely to break off the interview than respondents in the pre-assigned gender conditions. Permitting survey respondents to have a say in establishing the context for a survey interaction may therefore increase their engagement and willingness to be involved. It is also reasonable to expect that respondents who hold opinions about whether or not they want to interact with interviewers of a particular race are more likely to have a positive survey experience if their interviewers match their preferences. A positive survey interaction is likely to increase the odds that a respondent will be willing to engage in a future survey interaction, whether issued by the same sponsor or not. This factor should be of particular importance in the conduct of surveys with populations who have a mistrust of research.

Among ethnic identity study participants, the mean importance of having an interviewer of one's same race and ethnicity for Assimilated respondents was 4.4, which was just below the midpoint and suggests that these respondents do not have a strong preference for an interviewer of their same race and ethnicity. Respondents with a Bicultural identity component had a mean of 5.1, which was almost exactly the midpoint of the scale and reflects a higher preference for African American interviewers than was found for the Assimilated respondents. Multicultural respondents had a slightly higher mean of 5.2. This response also fits their profile, as the Multicultural individual is interested in the struggles and cultures of many peoples around the world and may view an African American interviewer as more likely to understand their worldview. Respondents with Black American and Afrocentric identity types had means above the midpoint of the scale at 6.5 and 6.6, respectively, followed by respondents with a Cultural Mistrust component at a mean of 7.2. The means of these last three groups suggest that they all prefer to interact with an interviewer of their same race and ethnicity for a survey with racial content.

The mean scores on the predicted comfort if the interviewer had been White item followed a similar overall pattern to that found with the means of the importance item. As with the importance item, means on the comfort item were higher for ethnic identity respondents with the Assimilated (7.0), Bicultural (7.5), and Multicultural (7.2) identity components, as would be expected from persons who claim to be comfortable interacting with Whites (Assimilated), African Americans and Whites (Bicultural), and persons of varying races and ethnicities (Multicultural). Respondents with Black American and Afrocentric identity components expressed lower predicted comfort if the interviewer had been White at 6.8 and 6.9, respectively. Respondents with a Cultural Mistrust component had a mean comfort level of 6.3. The Cultural Mistrust mean was the lowest among the identity component types and corroborates an identity type that is characterized by a mistrust of White people. Overall ratings of comfort were high across the identity components, however, with no means indicating strong discomfort with a White interviewer.

This study has several limitations. For one, the data included in the present analyses only represent respondents who agreed to engage in a telephone survey with the

type of interviewer who called them. Thus, it is possible that the ethnic identity dataset, for example, only includes respondents with a certain threshold comfort level with African American interviewers. A floor effect is therefore possible. Future research will attempt to explore this issue by examining whether survey response rates varied by interviewer race in the motivation study dataset. Another limitation is that the available data did not enable this study to assess whether interviewer error increases if respondents who prefer to interact with interviewers of their same race and ethnicity are assigned to interviewers with dissimilar racial and ethnic characteristics. Further research with larger sample sizes is needed that can compare variance associated with data from respondents who prefer and are assigned to racially and ethnically similar interviewers to data from respondents who prefer but are not matched to racially and ethnically similar interviewers. This study was constrained by the study design for the parent study, the *Eat for Life* trial, which also limited the sample size for present analyses. A larger study might yield more respondents with Assimilated or Cultural Mistrust components. The study design also called for measuring ethnic identity only in the ethnic identity survey. Thus, the ethnic identity respondents receiving additional cueing about their interviewer's race via the use of "our community" in the introduction to the ethnic identity measure near the end of the survey. This additional cueing and its closer proximity to the interviewer race preference questions may have impacted the findings on the association between survey content and interviewer race preferences. It should also be noted that survey respondents were only asked about their preferences for interviewer race; thus, this study does not assess the relative importance of interviewer race and ethnicity in comparison with other interviewer characteristics such as gender, age, social status, voice qualities, etc. It is possible that factors other than race and ethnicity that were not included in this study serve as better correlates of respondents' preferences for interviewer race. An additional limitation of this study is that the BICS is a relatively new measure of ethnic identity and, as such, requires further development and refinement to achieve a rich understanding of each of the constructs that it measures. For example, the Cultural Mistrust scale is a very short, 3-item scale that, for reasons discussed elsewhere (Davis et al., under review), focuses on assessing respondents' opinions about the willingness of politicians and the government to help African Americans. Responses

to these items are believed to correlate with more sensitive and harder to measure attitudes about White people and White society, but further research is required to determine whether this short scale achieves this aim. Lastly, the data in this study are based on self-report, which may or may not reflect valid estimations of the variables under consideration.

Despite these limitations, this study provides a valuable contribution to the literature by documenting relationships between interviewer preferences and ethnic identity, racial salience, and questionnaire content among African American survey respondents. These findings provide evidence for substantial in-group ethnic variability among African Americans. This research also raises questions about the appropriateness of purposefully matching African American interviewers and respondents by race alone. Public health programs often target populations by race and ethnicity. When surveying a specific, racially defined population, it is not atypical for public health professionals to staff interviewing teams with interviewers who match the race of the target population. Yet, Aspinall (2001) argues that racial and ethnic matching are based on "an edict of epistemological privilege which holds that a single 'truth' or 'reality' exists for people of a particular ethnic group" (p. 840). If matching increases interviewer error, then our knowledge of social issues is biased to the degree to which such knowledge is based on surveys with matched designs. And, such matching, if based purely on race, will inherently demarcate differences among racial groups. On the other hand, this study indicates that many African Americans have preferences for interacting with racially and ethnically matched interviewers for a telephone survey with racial content. It is possible that these preferences would be even stronger for a face-to-face survey. Many African Americans may be more comfortable with African American interviewers, and this greater comfort may lead to reduced measurement error and a more positive survey experience. More research is needed to determine whether fulfilling respondents' preferences for interviewer race and ethnicity leads to the induction or reduction of measurement error. But, in the interim, this study provides evidence against a general policy of matching interviewers to respondents by race alone. It is clear that attitudes about race and ethnicity vary greatly among African Americans, and these attitudes are also likely to vary among other racial and ethnic groups. The importance of interviewer

race and other characteristics to respondents are similarly likely to vary within every population of survey respondents. Until these dynamics are better understood, public health professionals would be wise to measure, monitor, and control for interviewer effects in their collection, analyses, and interpretation of survey data.

Table 3.1. Descriptive Statistics for the Study Sample

	***Eat for Life* Ethnic Identity Arm Participants (n=617)**	***Eat for Life* Motivation Arm Participants (n=272)**
Female (%)	71.0	66.2
Mean Age in Years (SD)	48.6 (10.9)	46.3 (11.0)
Health Plan Affiliation (%):		
Detroit	48.0	48.9
Atlanta	52.0	51.1
Married or Living with Partner (%)	41.1	44.5
Educational Status (%):		
Less Than High School	2.7	3.9
High School or GED	24.1	27.6
Training Other Than College	6.2	4.3
Some College/2-Year College Graduate	31.8	35.8
College Graduate	20.8	21.3
Graduate School	14.5	7.1
Income (%):		
$20,000 or Less	8.1	8.0
$20,001 to $40,000	28.6	34.4
$40,001 to $60,000	27.2	31.3
$60,001 to $80,000	16.6	14.7
More Than $80,000	19.6	11.6
Mean on Single Racial Salience Item (SD)	8.0 (2.5)	8.2 (2.9)
Ethnic Identity (% with Component):[1]		
Assimilated	13.0	NA
Afrocentric	30.2	NA
Black American	54.8	NA
Bicultural	39.2	NA
Multicultural	45.5	NA
Cultural Mistrust	11.7	NA

[1] Participants may be classified as having more than one ethnic identity component.

Table 3.2. Results of Linear Mixed Models Predicting Preferences for an Interviewer of One's Same Race and Ethnicity (Hypotheses 1 and 3) and Comfort with a White Interviewer (Hypotheses 2 and 4) by Ethnic Identity and Racial Salience [1]

	Estimate	Standard Error	Pr > | t |
Hypothesis 1			
Assimilated	0.1416	0.8165	.86
Afrocentric	1.0572	0.4359	.02*
Black American	1.7578	0.4699	.0002*
Bicultural	-0.1296	0.4729	.78
Multicultural	0.1733	0.4369	.69
Cultural Mistrust	0.8228	0.4497	.07
Hypothesis 2			
Assimilated	-0.2001	0.7032	.78
Afrocentric	0.0291	0.3758	.94
Black American	-0.4361	0.4049	.28
Bicultural	0.4106	0.4075	.31
Multicultural	0.0254	0.3764	.95
Cultural Mistrust	-0.5042	0.3877	.19
Hypothesis 3			
Racial Salience	0.1899	0.0581	.001*
Hypothesis 4			
Racial Salience	0.0058	0.0483	0.90

[1] All models controlled for respondent gender, age, education, and income.
* Statistically significant at $p \leq .05$.

<u>Table 3.3</u>. Results of Linear Mixed Models Predicting Preferences for an Interviewer of One's Same Race and Ethnicity (Hypothesis 5) and Comfort with a White Interviewer (Hypothesis 6) by Type of Survey [1]

	Estimate	Standard Error	Pr > | t |
Hypothesis 5			
Change from Ethnic Identity to Motivation Survey	-1.0038	0.3069	.001*
Hypothesis 6			
Change from Ethnic Identity to Motivation Survey	0.5037	0.2491	.04*

[1] All models controlled for respondent gender, age, education, and income.
* Statistically significant at $p \leq .05$.

Table 3.4. Mean Importance of Having an Interviewer of One's Same Race and Ethnicity and Comfort with a White Interviewer by Ethnic Identity Component (Ethnic Identity Respondents Only)

Identity Component	Mean Score on Importance of Having an Interviewer of One's Same Race and Ethnicity	Mean Score on Predicted Comfort Level if the Interviewer Had Been White
Assimilated	4.4	7.0
Bicultural	5.1	7.5
Multicultural	5.2	7.2
Black American	6.5	6.8
Afrocentric	6.6	6.9
Cultural Mistrust	7.2	6.3

CHAPTER 4

LINGUISTIC EXPRESSIONS OF IDENTITY IN THE SURVEY INTERACTION: THE USE OF AFRICAN AMERICAN ENGLISH IN TELEPHONE SURVEYS

INTRODUCTION

Race is a profound force in American society. Thus, it may be unsurprising that the race of an interviewer can impact survey data (see Chapter 2). However, the impact of interviewer race is not fully understood and requires further investigation. One under-explored influence may be the cultural variability that exists within racial groups. Prior research has focused on comparing data obtained by African American interviewers to data obtained by White interviewers, which treats African American interviewers as a separate homogeneous group. Yet, a growing body of research indicates that African Americans vary in their feelings about race and ethnicity (Cross, 1991; Sellers et al., 1998). For some African Americans, being African American forms the structure upon which one's personal identity is constructed. For other African Americans, race plays only a tangential role in shaping personal identity in relation to other aspects of the self. These orientations lead to very different formulations of ethnic identity. To date, this diversity in African American ethnic identity has not been accounted for in studies of interviewer effects. But, just as interviewer race may influence survey data, within-group variability in African American interviewers' ethnic identity may also impact the responses that African American survey participants provide.

No research has been published on interviewer ethnicity effects among African Americans, but the possibility of interviewer ethnicity effects yields several interesting questions. For one, how is ethnic identity communicated, if at all, during a telephone survey interaction between African American interviewers and respondents? In a face-to-face survey, ethnic attitudes may be communicated visually through clothing, jewelry,

hairstyle, body language, or other style choices, although interviewers are generally trained to avoid such cues. In both face-to-face and telephone surveys, non-visual cues may be available that suggest an interviewer's ethnicity: (a) the text of the survey script itself, such as in the rare case where the script identifies the ethnic orientation of the interviewer; (b) an interviewer's name, which may or may not evoke ethnic connotations; and (c) aspects of the interviewer's speech that shape the way in which he or she delivers the script. Thinking about how ethnicity may be conveyed between African American interviewers and survey respondents prompts a second question: If present, do African American interviewers' expressions of ethnic identity influence the data that African American respondents provide? As discussed in Chapter 2, prior research suggests that survey items that overtly query racial or ethnic attitudes may be the most susceptible to interviewer race effects. By extension, is it reasonable to speculate that interviewer ethnicity similarly impacts responses to overtly racial or ethnic survey items? And, if ethnicity effects occur, are there other predictable patterns to suggest what contexts motivate African American interviewers to convey their ethnic identification to African American telephone survey respondents?

This paper describes a small study designed to examine a subset of the questions outlined above. Specifically, this exploratory research investigates: (a) whether and how ethnic identity is conveyed through the use of culturally affiliated linguistic features during a scripted telephone interaction between African American interviewers and African American respondents, (b) whether these expressions of identity impact the answers to survey questions that respondents provide, and (c) if certain types of survey items are more or less likely to evoke culturally associated speech patterns. The results of this work are limited by modest sample size. However, as the first research of this kind, it is hoped that this study will inform and generate further consideration of within-group cultural variability on the survey response.

BACKGROUND

How an individual uses language is one of the most powerful tools for expressing personal identity. This study evaluates the use of language by African American

interviewers and respondents living in a social environment in which two language systems are likely to prevail: American Standard English and African American English. American Standard English (ASE) is the predominant dialect of English spoken in the United States. ASE is the dialect taught in schools and is typically employed in business settings and in more formal interactions among strangers. Other English dialects are also in use in the United States, and most connote geographic origins or social class. However, unlike these other American English dialects, African American English (AAE)[2] is primarily identified with speakers from a single racial group – African Americans. Not all African Americans speak AAE, and not all speakers of AAE are African American. But, since most speakers of AAE are African American, AAE is primarily identified with African American people and culture.

Any discussion of AAE is inherently political, because discussing the language of a people is intertwined in debates over who is and is not a true member of the group (Morgan, 1994). As a consequence, both speakers and scholars alike disagree over many aspects of AAE, including what to call it. Many names exist for language associated with African Americans. Of these, Ebonics is perhaps the most popularly known by non-linguists. In response to Eurocentric claims that Africans "lost" their languages and culture when brought to the U.S. during slavery, Ebonics was coined by African American scholars in 1973 to mark the influence of African languages and communication styles on the distinctive speech associated with African slaves and their descendants (Williams, 1975). However, Ebonics is not well defined from a sociolinguistic perspective (Baugh, 2000), and, since Ebonics became the term of choice during the media's reporting of the 1996 Oakland school board controversy, it has acquired a misleading and often negative reputation in many Americans' minds (Baugh, 2000; Green, 2002; Rickford, 1999). For these reasons, the term Ebonics is not used in this paper. Another term applied to the speech of African Americans is Black English Vernacular (BEV). Many of the specifics of BEV, which was subsequently renamed African American Vernacular English (AAVE), were defined through the work of William Labov in the late 1960s and early 1970s (Labov, 1972). In an effort to dispel

[2] There is some debate in linguistics as to whether AAE and its variants are dialects of English or independent languages. See Smitherman (2000) and Bailey (2001) for further discussion of this issue.

racist myths about African American children's cognitive abilities, Labov's research focused on legitimizing AAVE by providing evidence for the systematic ways in which AAVE diverged from ASE. At that time, he determined that the divergences were most extreme in the vernacular speech of those speakers most associated with "street culture:" male African American adolescents residing in the inner cities of the Northern U.S. (Nguyen, 2006). Thus, AAVE primarily describes the informal, street culture speech of a sub-set of African American youth (Nguyen, 2006). Labov argued that the speech of other African Americans, including women and those from the middle class, was closer to that of Whites and, therefore, further from the authentic vernacular of African American people (Labov, 1972; Morgan, 1994; Nguyen, 2006). Labov termed these latter speakers "lames" (1972). Thus, while Labov made great strides in establishing AAVE as a legitimate dialect, he also divested the majority of African American speakers from African American culture. The term Black English was introduced in 1969 to describe the speech of African Americans of all ages and social classes and to dispel the negative connotations that some associated with the word "vernacular" (Wolfram & Fasold, 1969). In the 1990s, Black English was mostly retired as the term African American English became more politically correct. However, there is no single, agreed-upon terminology for the speech styles that have been associated with African Americans. Nor is there agreement upon the definitions of the various terms or who uses each speech style. Some people argue that no dialect should be identified, as this act legitimizes beliefs in racial differences. However, most linguists agree that there are distinguishable speech characteristics that can be primarily associated with African American people and culture. Since it seems to be the least misunderstood and most encompassing term for speech associated with African Americans of various ages and social classes, this paper uses the term African American English (AAE) from this point forward.

Morgan wrote that "if the history of a language speaks volumes, the history of African American English is deafening" (Morgan, 2002, p. 13). As with most aspects of AAE, the roots of AAE are contentious. Many hypotheses are debated in contemporary linguistic circles. The two primary hypotheses, the divergence hypothesis and the creolist hypothesis, are briefly described here because they demonstrate that languages are alive

and ever-changing and that their histories, however much debated, bear heavily upon people's attitudes about what the use of a language represents. The divergence hypothesis contends that the speech of African Americans and Whites was once more similar than it is today. This premise rests upon the belief that AAE developed relatively recently as a consequence of increasing segregation in the 20th century (Labov, 1972; Rickford & Rickford, 2000). The divergence perspective purports that African American slaves lost their African languages as they were forced to adopt the English of their captors. However, much later, industrialization prompted growing numbers of African Americans to move to urban, segregated neighborhoods in the Northern cities, and African Americans and Whites became increasingly linguistically isolated from one another. As a result, a distinctive dialect of African American speech emerged on the streets of the inner Northern cities (Labov, 1972). Insofar as residential and educational segregation persist, the divergence hypothesis predicts that AAE and ASE will continue to differentiate from one another over time (Rickford & Rickford, 2000). In contrast, the creolist hypothesis argues that the current dialect of AAE developed from a creole language based on English and African languages that has converged to become more similar to ASE over time. These convergences are believed to be the result of the historical influences of slavery and segregation, as well as times of hopefulness toward gaining new access to social and economic opportunities which encouraged African Americans to periodically decreolize their language and adopt White standards (Smitherman, 2000). Creolists hold that the original development of AAE resulted from slaveholders' purposeful mixing of African slaves who spoke different languages, which was done in order to prevent the organization of rebellions (Baugh, 1999). Since slaves were not allowed to learn how to read or write, many African Americans were socioeconomically and linguistically isolated (Baugh, 1999; Smitherman, 2000). As a consequence, a simple shared language, or pidgin English, developed as these diverse speakers co-mingled (Dillard, 1972). The pidgin language became a creole English as the next generation of African Americans were reared with the new language as their native tongue (Dillard, 1972). The creole language was comprised of predominantly English vocabulary with grammar and semantics based on African languages (Smitherman, 2000). A decreolization of AAE occurred around the time of

Emancipation, as many African Americans believed that use of AAE would limit their economic prospects (Smitherman, 2000). But, the linguistic isolation of many African Americans remained after Emancipation, due to the institutionalization of segregation and unequal opportunities for African Americans through the Jim Crow laws (Baugh, 1999; Mufwene, 2001; Smitherman, 2000). The mass migration of African Americans to segregated neighborhoods in the Northern cities during 1900-1960 contributed to increased linguistic isolation, as well as increased linguistic freedom (Morgan, 2002). However, World Wars I and II also created new economic opportunities for African Americans, resulting in the creation of an African American middle class. The pursuit of these opportunities was accompanied by a further decreolization of AAE (Smitherman, 2000). Proponents of both hypotheses would endorse the position that in the 1960s, use of AAE began to be a symbol of Black pride and solidarity (Dillard, 1972; Smitherman, 2000). Since that time, AAE has been embraced by many African American writers, as well as the Hip-Hop Nation, in an effort to emphasize in-group identity and the uniqueness of African American speech and culture (Smitherman, 2000). However, there may be deepening divisions between working-class African Americans, who are more likely to be monodialectical speakers of AAE, and middle-class African Americans, who are more likely to be fully bidialectical (Smitherman, 2000). This bidialectalism gives middle-class African Americans social advantages, and in U.S. society today, "(c)ommunicative resources … form an integral part of an individual's symbolic and social capital" (Gumperz & Cook-Gumperz, 1982, p. 5).

AAE shares many traits in common with ASE, but the two dialects differ in many systematic aspects that can lead to miscommunication (Hansell & Ajirotutu, 1982; Rickford, 1999; Smitherman, 2000; Speicher & McMahon, 1992). For one, many linguists have observed that AAE has some unique syntactic constructions (Dillard, 1972; Green, 2002; Koch, Gross, & Kolts, 2001; Martin & Wolfram, 1998; Rickford & Rickford, 2000). For instance, the ASE sentence "Charles is working" can be phrased in AAE as either "Charles be working" or "Charles working". In AAE, "Charles be working" indicates that Charles is generally working most of the time, whereas "Charles working" means that Charles is working at the moment, thereby indicating a different temporal reference (Baugh, 1999; Rickford & Rickford, 2000; Smitherman, 2000). AAE

may be distinguished from ASE by different vocabulary (Green, 2002; Ogbu, 1999), as well as different meanings for the same words (Green, 2002; Smitherman, 2000). In addition, AAE speakers characteristically imbue words and phrases with multiple meanings (Spears, 2001). This style may have originated with African communication forms, but it certainly continued out of a need to communicate in ASE-sounding speech without being understood by speakers of ASE (Morgan, 2002; Smitherman, 2000). Until the 1960s, the oppressive environment for African Americans living in the South required the use of deferential speech styles that implied an acceptance of White superiority as a survival tactic (Morgan, 2002). Through the use of multiple meanings, or indirectness, African American speakers were able to maintain a social face in front of other African Americans within earshot, thereby indicating a resistance to oppression that was only detectable to in-group members (Morgan, 2002). According to Smitherman (2000), "This Africanized form of speaking became a code for Africans in America to talk about Black business, publicly or privately, and in the enslavement period, even to talk about 'ole Massa' himself right in front of his face" (p. 19). Other communication norms that have been associated with AAE include the use of directness, signifying or playing the dozens (ritualized games of oral insults designed to demonstrate and test the speakers' verbal skills), the valuing of rhythm, call and response patterns, low voice pitch to indicate sincerity, and specific speaker-audience interactions that determine the meaning of words (Green, 2002; Morgan, 2002; Mufwene, 2001; Smitherman, 2000; Spears, 2001). AAE is also distinguished by the systematic use of certain phonological features, which concern how sounds are pronounced (Koch et al., 2001; Rickford, 1999; Rickford & Rickford, 2000). Due to the constraints of using a scripted survey, phonological features may be the most likely linguistics features to be used in a standardized survey interaction. As with speakers of all languages, AAE speakers vary in which AAE features they use and in how often and in what types of situations they use them (Rickford, 1999). Thus, there is natural variability both among and within AAE speakers.

In most societies, dialects spoken by a minority population tend to disappear over time (Filmer, 2003). However, use of AAE remains strong. Although debatable (Morgan, 2002), it has been roughly estimated that 80% of African Americans use AAE

as their primary dialect (Baugh, 1999; Dillard, 1972). Many African Americans learn AAE as a first language (Baugh, 1999), and it remains the language that they use within the home and community (Ogbu, 1999). For these speakers, ASE is acquired at school and is used in educational and occupational settings (Ogbu, 1999). ASE is associated with White people and power systems and is the dialect of choice for use in playing the "game for survival" in interactions with outsiders to the community such as police officers, social workers, and people passing out flyers (Ogbu, 1999). To many AAE speakers, use of ASE is believed to be necessary to open doors to educational and economic opportunities (Green, 2002; Ogbu, 1999; Speicher & McMahon, 1992). As such, the ability to use both ASE and AAE, or code-switch, is seen by many African American as a valuable social skill. However, many other African Americans feel that acquisition of ASE is not sufficient to overcome the barriers imposed by racism and discrimination and, therefore, have little motivation to use ASE (Ogbu, 1999). For many African Americans, use of AAE signifies support and appreciation of the fortitude and resilience of African Americans in their survival of slavery, segregation, and discrimination. Use of AAE signifies in-group solidarity and racial pride and can be used to connect across class boundaries. For instance, in a study of second-generation middle-class African American youth living in neighborhoods with a minority of African American residents, Linnes (1998) observed that native ASE speakers used AAE features to gain access to African American peer groups. As Linnes writes, "For these younger speakers, acquisition of AAE represents an ethnic membership card allowing them fuller access to a culture they have only peripherally experienced" (p. 349). In a qualitative study of African American attitudes toward AAE, Speicher and McMahon (1992) similarly present evidence of use of AAE by African American professionals as a means of making a cultural connection with one another. In contrast, use of ASE in AAE-dominated speech communities represents "talking White" and a rejection of African American culture (Baugh, 1999; Linnes, 1998; Ogbu, 1999; Rickford & Rickford, 2000; Smitherman, 2000). In these speech communities, use of ASE is not viewed as the acquisition of an additional language; it is viewed as a displacement of AAE and the assimilation of the speaker into a society whose values are being imposed onto African Americans (Ogbu, 1999). As a result, some children who use ASE at home or in the

community may be criticized by their parents and peers for "acting White" (Speicher & McMahon, 1992), whereas other children who resist learning ASE may be praised for their fealty to AAE (Ogbu, 1999). Even highly educated adult members of these speech communities who use ASE in their professional lives may eschew its use at home or in the community (Ogbu, 1999).

But, as with other aspects of AAE, opinions are both varied and strong. Since AAE is associated with a population targeted by discrimination, it has become a stigmatized speech form to many ASE speakers, both African American and White (Green, 2002; Morgan, 2002; Mufwene, 2001). Use of AAE has been associated with laziness, low social class, low education, lack of intelligence, and less likeability (Baugh, 1999; Doss & Gross, 1994; Filmer, 2003; Green, 2002; Koch et al., 2001; Ogbu, 1999; Speicher & McMahon, 1992). Some of these perceptions result from the fact that, although AAE speakers derive from all social classes, most monodialectical speakers are lower and working class African Americans (Rickford, 1999; Smitherman, 2000). Negative stereotypes may also derive from the initial descriptions of AAVE by Labov and others, which, as a result of focusing on adolescent male speakers from the street culture, were characterized by a high degree of profanity, sexual references, and ritualized linguistic insult games that are highly offensive to many AAE and ASE speakers alike (Speicher & McMahon, 1992). Many non-linguists, including bidialectical African American speakers (Brown, 2006; Ogbu, 1999; Speicher & McMahon, 1992), might not realize that AAE is a systematic dialect; therefore, they believe that AAE speakers are merely speaking "slang" or poor English (Green, 2002). As a result of these negative connotations, many African Americans are critical of AAE speakers and embrace ASE as an expression of educational and economic attainment. In contrast, other African Americans, such as those affiliated with the Nation of Islam, may avoid AAE because it implies the continued subservience of African Americans to Whites since the times of slavery (Morgan, 1994). Other African American monodialectical ASE speakers may feel that their use of ASE signifies a victory over racism and oppression (Morgan, 2002).

Because there are so many value-laden perspectives on language use among African Americans, those African Americans with the opportunity to choose their

primary language face complex choices. As a participant in Speicher and McMahon's (1992) study noted, the varying attitudes about AAE make language a source of division instead of a source of unity for African Americans. The African American linguist John Baugh describes some of these tensions in writing about his early linguistic experiences as a youth in Philadelphia: "At a young age … I received mixed messages about language; some were overt, advocating that I 'speak properly' and avoid 'bad language,' whereas others were more subtle, reflected by the hippest Sisters and Brothers who emphatically rejected '(W)hite speech' " (Baugh, 2000, p. 5). What is clear is that many African Americans feel strongly about language use (Baugh, 1999; Morgan, 1994, 2002; Ogbu, 1999; Speicher & McMahon, 1992), and, as a consequence, the language used by African American interviewers and respondents may have implications for the survey interaction.

The traditional use of ASE in survey scripts may impact the survey interaction in several ways. For instance, adult AAE speakers in Ogbu's (1999) Oakland study reported that ASE speakers would assume that they were "ignorant" and would not believe what they said. These beliefs discouraged them from interacting with native speakers of ASE, including their children's teachers (Ogbu, 1999). Do monodialectical AAE speakers react in a similar way to ASE-speaking survey interviewers, thereby leading to their lower representation in survey data? A person's native dialect may also impact their comfort levels and cognition during the survey interaction. In at least one study, bidialectical African American speakers noted that they felt more comfortable using AAE and referred to ASE as an "alien dialect" that was used less frequently in their daily activities (Ogbu, 1999). How is a respondent impacted by having to use an "alien dialect" when participating in a survey? And, do differences in dialect lead to misunderstandings during interviews, data analyses, and the interpretation of survey findings? For example, Rickford (1999, p. 322) reports that many Whites incorrectly interpret the sentence "She BIN married" as meaning that the woman in question was married at one time but is no longer married. How many non-AAE speaking survey researchers would recognize that this sentence means that the woman is still married? Variations in communication norms may also impact the survey interaction. Morgan (2002) suggests that discussions of trustworthiness may be affected by culturally bound

interpretations. She contends that to many African Americans, talking about trust in an attempt to establish one's trustworthiness has the opposite effect – it makes the listener question the honesty of the speaker (Morgan, 2002). If this interpretation is prevalent, what, then, do the assurances of confidentiality that are routinely provided at the beginning of a survey convey to African American respondents?

There is no extant literature on language use during surveys with African American respondents. Thus, the purpose of this study is to address this gap by exploring one of many aspects of the role of language in the survey process: the use of AAE in survey interactions between African American interviewers and respondents. Since the use of AAE is so value-laden among African Americans, this study pursues a deeper understanding of whether African American interviewers use AAE during surveys and, if so, how such use impacts the survey data that respondents provide. African American interviewers who routinely use or do not use AAE features may be associated with role-independent interviewer effects, which are defined as measurement errors that arise from personal characteristics associated with a particular interviewer or set of interviewers that may systematically impact the answers to survey questions that respondents provide. As discussed in Chapter 2, interviewer effects can have a significant impact on survey findings. Thus, the exploration of systematic linguistic differences leading to interviewer effects is an important pursuit.

This study has three main hypotheses. First, since language is an expression of personal identity, Hypothesis 1 predicts that African American interviewers with more pro-Black ethnic identity types will demonstrate higher use of AAE-associated features. Hypothesis 2 postulates that African American interviewers will use AAE features more when administering certain types of survey content. A first sub-hypothesis is based on previous findings that interviewer race effects tend to be associated with racially topical survey items (see Chapter 2). Specifically, it predicts that interviewers will exhibit more AAE features during racially topical sections than during other parts of the survey script. A second sub-hypothesis presumes that because interviewers are charged with building rapport, they will use AAE features more during recruitment sections than in other parts of the survey script. The survey literature indicates that respondents often search for cues from interviewers on how to put their best selves forward. In a telephone survey,

respondents may search for linguistic cues from their interviewers to determine the most socially desirable response for a given context. Because sentiment about use of AAE may be high among African American respondents, interviewers' use AAE features may impact the answers that respondents provide. With these dynamics in mind, Hypothesis 3 predicts that when interviewers use AAE features during racially topical parts of the interview script, respondents will provide more pro-Black answers to racially topical survey questions. Since this is an exploratory study, other patterns of use of AAE features will also be investigated. If other patterns exist and impact survey data, these patterns may provide important information for survey research and practice.

METHODS

Study Sample

The participants for this study consisted of African American telephone interviewers and African American telephone survey respondents.

African American interviewers were recruited from two sources. The first group of interviewers was recruited from the 13 African American interviewers who conducted baseline telephone surveys for a randomized controlled trial called *Eat for Life*. *Eat for Life* was a five-year trial designed to test the effectiveness of personalizing health program materials on ethnic identity to increase fruit and vegetable intake among African American adults. The *Eat for Life* trial is described elsewhere (Resnicow et al., under review). The *Eat for Life* interviewers worked for a survey call center in Washington State. The interviewers were sent a recruitment packet containing an introductory letter, consent form, and self-administered survey. Enrollees were encouraged to mail back their completed consent forms and surveys to the study coordinator using an enclosed pre-stamped envelope. Interviewers were originally offered $75 as a post-incentive for participating in this study. However, due to very low initial enrollment rates, the incentive was increased to $255 for completion of all study procedures (described below). Eight of 13 Washington (WA) interviewers enrolled in the study, all of whom were female.

In order to increase the sample of interviewers, a second group of interviewers was recruited from a potential pool of eight female African American interviewers working at a survey call center associated with a public university in Michigan. The Michigan (MI) interviewers were recruited via the posting of flyers in their workplace. Potential participants called or emailed the study coordinator, and, if the caller was female and self-identified as Black or African American, she was mailed a recruitment packet consisting of a similar recruitment letter and consent form and an identical self-administered survey as those sent to the WA interviewers. Six potential participants contacted the study coordinator, and all six were mailed recruitment packets. As with the WA interviewers, the MI interviewers were encouraged to mail back their completed consent forms and surveys to the study coordinator using an enclosed pre-stamped envelope. The MI interviewers received a post-incentive of $255 for completion of all study procedures. Six of the eight MI interviewers enrolled in the study.

The African American respondents consisted of those persons who had completed baseline telephone surveys as part of their participation in the *Eat for Life* trial. *Eat for Life* participants were recruited from two integrated health care delivery systems, one in Atlanta and one in Detroit. Participants were randomly assigned to an experimental group or a control group for the purposes of the *Eat for Life* trial. However, all participants completed the same baseline survey, and the random assignments were not made until after completion of the baseline survey. Thus, for the purposes of this study, the experimental and control group participants are considered as a single study group.

The Detroit participants consisted of randomly selected healthcare system members whose medical records identified them as African American. The Atlanta healthcare system did not record racial or ethnic information on its members. Therefore, the Atlanta respondents were randomly selected from healthcare system members with home addresses in Census blocks in which 80% or more of the residents were African American at the time of the 2000 Census. Recruitment letters containing pre-incentives in the form of $2 bills were mailed to potential participants in one-week waves until enrollment targets were achieved. These letters were followed by recruitment telephone calls conducted between September 2006 and July 2007. During the eligibility portion of these calls, interviewers ensured that participants were between the ages of 21 and 70,

self-identified as Black or African American and were not Hispanic or multiracial, ate fewer than ten servings of fruit and vegetables per day, were not currently hospitalized or living in skilled care facilities, had lived in the U.S. for more than half of their lives, and had no mental or physical conditions that would inhibit or be endangered by their participation in *Eat for Life*. A total of 2,018 recruitment letters were mailed to potentially eligible respondents, yielding 625 completed baseline surveys by eligible participants (AAPOR Response Rate 1 = 34.5%) (American Association for Public Opinion Research, 2003). Of these 625 respondents, only the 492 participants who were considered to have complete survey data and who were interviewed by one of the eight WA interviewers who participated in this study were included in the present analyses. All *Eat for Life* participants who completed a baseline survey received a thank-you letter containing a $5 bill within a couple of weeks of completing the survey and tailored health education materials over the next three months. Participants were subsequently contacted to complete a follow-up survey and receive an additional incentive of a $15 gift card to a retail store.

Eat for Life participants were randomly assigned to the WA interviewers. The *Eat for Life* participants did not interact with the MI interviewers. All of the ethnic identity study interviewers were African American, and all respondents were cued to the fact that their interviewer was African American through the following language in the recruitment script: "I am calling as part of a team of African American interviewers …" All respondents also heard one reference to "our community" near the end of the recruitment script and two more references to "our community" at about two-thirds through the interview during the introduction to the survey section where participant ethnic identity was assessed.

This study was approved by human subjects review committees at the University of Michigan, the two participating integrated healthcare delivery systems in Detroit and Atlanta, and the institutions where the interviewers were employed.

Study Procedures

Each participating WA and MI interviewer completed a self-administered survey measuring sociodemographics and ethnic identity. These measures are described below. *Eat for Life* participants completed a telephone survey assessing sociodemographics, ethnic identity, and beliefs and behaviors related to fruit and vegetable consumption.

Early in the study design phase, it was hoped that the actual *Eat for Life* baseline telephone interviews could be digitally recorded. Those recordings would have been used in the present analyses. However, concerns about protection of personal information made the health care delivery systems wary about allowing the collection of these types of data. Recording only the interviewers' side of the survey interactions was also considered. But, there was no available means of recording only those interviews conducted by interviewers participating in the study while assuring the confidentiality of their study participation from supervisors and co-workers. As a result, it was not possible to record the actual *Eat for Life* telephone interviews.

In an attempt to obtain information about how the *Eat for Life* interviewers may have used language during the actual *Eat for Life* interviews, simulated survey interactions were conducted. In these simulated interviews, each of the 8 participating WA interviewers and the 6 participating MI interviewers administered three digitally recorded telephone interviews using the *Eat for Life* baseline survey script. All interviewers surveyed the same three standardized respondents. The standardized respondents were not *Eat for Life* program participants or affiliated with any healthcare system. The three standardized respondents were African American females of a similar age and education level. One of the respondents exhibited low use of AAE features in a sample recording of her natural speech, while the other two respondents demonstrated high use of AAE features. The respondents were asked to use their natural speech during the simulated interviews; thus, no acting was involved. Forty-two interviews comprised the corpus of recordings. Each interview lasted approximately 40 minutes. The interviews were recorded using either an Olympus DM-20 or DS-40 Digital Voice Recorder. Olympus TP7 Telephone Recording Devices were used to connect the recorders to the telephone lines used by the standardized respondents. Thus, the

recordings captured both the interviewers' and the standardized respondents' speech exactly as it would have been heard by either party on the line. However, no survey data from the standardized respondents were used in any analyses.

Since a full disclosure of the study objectives and design might have influenced the interviewers' recorded speech, the WA and MI interviewers received only a vague description of the study objectives prior to participating in the study. However, once data collection was complete, all participating interviewers were sent a letter describing the goals of the study in more detail. The WA and MI interviewers' participation in the study ranged between July and September 2007.

Measures

Linguistic Features

This research focused on telephone-administered surveys in order to isolate linguistic variables from visual cues. This research further explored the use of AAE phonological features in standardized telephone interviews, since standardized interviewing likely comprises the majority of telephone-administered surveys. When using an ASE standardized survey script, interviewers are limited in their opportunities to incorporate most AAE features and communication styles into their vocal delivery. However, in this context, an interviewer may easily utilize AAE-affiliated phonological features. Unlike some AAE features, AAE phonological features have been associated with African American speakers from varying social classes (Nguyen, 2006). So, they have the potential to be present in surveys conducted by African American interviewers from a range of backgrounds. Phonological features are also interesting to include in an exploratory analysis because individual phonological features, while all associated with AAE, may have additional associations with class and other sociodemographic characteristics (Nguyen, 2006). As a consequence, an examination of interviewers' use of AAE-affiliated phonological features has the potential to yield rich findings about the exploration of language use in the survey interaction.

Seven phonological features were selected for inclusion in this research. These features were selected because each variable: (a) had a contemporary association with AAE in the linguistics literature (Bailey & Thomas, 1998; Green, 2002; Labov, 1972; Rickford, 1999), (b) had multiple opportunities for occurring in the *Eat for Life* survey script, and (c) was likely to be audible in a telephone interview (for example, consonant sounds are easier to hear than vowel sounds in a telephone environment). Examples of phonological features that were excluded from coding for these reasons are the pronunciation of "brother" (ASE) as "brotha" (AAE) and the pronunciation of "business" (ASE) as "bidness" (AAE).

Although phonological features were the focus of this research, two other types of linguistic features were also coded. AAE syntactic features concern how words are put together and were coded throughout the interviews. An example of an AAE syntactic feature that was coded is the use of the invariant "be", which is invoked when a speaker uses the AAE construction "you be tired" instead of "you are tired". Discourse features, or off-script responses to respondents, were also coded if they consisted of more than an utterance such as "uh-huh" or "hmm" and had potential cultural associations. The occurrence of culturally associated syntactic or discourse features was expected to be rare, given the use of a standardized interviewing script and the professional nature of the interaction being recorded. But, if present in these circumstances, culturally associated syntactic and discourse features were expected to be meaningful. However, only one interviewer used an AAE-associated syntactic feature, and she used it only two times. No significant discourse features were observed in the recorded interviews. Thus, syntactic and discourse variables were dropped from the analyses.

As a consequence of the many practical constraints associated with this study, some linguistic features that people may associate with AAE or Ebonics were not present in the *Eat for Life* script or, as a result, the current analyses. Examples of features that were not included in these analyses are the use of distinctively AAE syntactic features (e.g., the use of double negatives, as in "don't nobody like Brussels sprouts") and AAE-associated vocabulary (e.g., "ashy" for dry skin). However, since ASE syntax and vocabulary are typically used in survey scripts, it is unlikely that interviewers employ AAE syntax and vocabulary with high frequency in real-world survey interactions. Thus,

the current dataset may likely reflect typical language use by professional African American interviewers.

Each opportunity of occurrence, or token, of a phonological feature was clipped into an individual auditory token file using Audacity 1.2.6 (Audacity, 2006). All tokens were double-coded. The coding was conducted using Praat 4.1.11 for Windows (Boersma & Weenink, 1992-2003). The primary coder had a doctoral degree in Linguistics from the University of Michigan, and the secondary coder was a doctoral student in Linguistics at the University of Michigan. Each token was coded as a "1" if it was articulated as the AAE variant of the sound being coded. Tokens that were not articulated as AAE variants were coded as "0". The coding protocol was designed to be conservative; thus, if a token was difficult to hear or the articulation was difficult to determine, a code of "0" was assigned. The non-AAE code was, therefore, the default code. Each coder entered the codes into an Excel spreadsheet. The primary and secondary coders were blinded to the codes that the other had assigned. In cases where the primary and secondary coder disagreed, a tertiary coder with a doctoral degree in Linguistics from the University of Michigan provided the final code. Only the tertiary coder had access to the codes provided by the other coders. The Excel spreadsheets were then imported into SAS 9.1.3 for Windows for analysis (SAS Institute, 2002-2003).

The seven phonological features that were selected for coding include the "ask" metathesis, syllable-final [d][3], "-ing" endings, consonant cluster reduction, the pin/pen merger, initial "th", and [ai] monophthongization. These features will be described in turn.

A metathesis is the substitution of one sound for another. To non-linguists, the metathesis of "ask" to "aks" may be one of the most readily identifiable phonological characteristics of AAE (Rickford, 1999). The "ask" metathesis is also found in traditional Southern White speech; however, it is disappearing among younger, Southern

[3] Letters and symbols in brackets connote sounds in the International Phonetic Alphabet (IPA). The letters in the brackets are not necessarily found in words containing the indicated sound. For example, the phonetic symbol [ai] may refer to the pronunciation of the vowel "i" in the word "design", even though the word "design" does not contain the letter "a". Thus, it is important to remember that IPA symbols represent sounds that may not have an obvious correlation to English spelling.

White speakers (Bailey & Thomas, 1998). In this study, all audible occurrences of the word "ask" in the survey script were clipped. However, other conjugations of "ask" such as "asking" and "asks" were not included as tokens. The metathesized variant of "ask" was coded as "1".

Syllable-final [d] occurs when a [d] at the end of a word is not articulated. For example, "food" would be pronounced as "foo". This feature is found in both AAE and Southern White vernacular speech; however, the glottalized context in which it occurs in AAE appears to be unique to AAE (Bailey & Thomas, 1998). A recent study indicates that use of the AAE variants of syllable-final [d], [ʔ] and [Ø], may be increasing over time among African American speakers in Detroit (Nguyen, 2006). For this project, both the [ʔ] and [Ø] variants of syllable-final [d] were coded as "1". Only tokens consisting of post-vocalic occurrences were clipped. All tokens also preceded a pause or consonant, but tokens preceding [t], [d], [dʒ], and [tʃ], and [j] were excluded. For example, the phrase "good fruit" would have been clipped as a token, but "good apple" would not have been clipped. The AAE variants were instances in which syllable-final [d] was pronounced as a glottal stop, [ʔ], or when a creaky voice or glottalization was not followed by a glottal or alveolar stop, [Ø].

The "-ing" variable is characterized by the reduction of words ending in "-ing", [ŋ], to "in", [n]. For example, the verb "eating" might be articulated as "eatin", and "sleeping" might sound like "sleepin". This speech variant is found in AAE as well as other English vernaculars and has a strong affiliation with lower social class; thus, it is considered to be a generally stigmatized variable across racial and ethnic groups (Green, 2002; Trudgill, 1974). For the present analysis, the "in" form was coded as "1". Tokens of words ending in "-ing" were clipped if the "-ing" functioned as a separate morpheme.

Consonant cluster reduction is realized when two or three consonants are grouped together in a cluster but only one consonant is articulated. In AAE, it is usually the last consonant that is dropped, which makes it distinctive from other dialects (Bailey & Thomas, 1998). For example, the word "send" might be pronounced as "sen", while "fast" may be articulated as "fas". Consonant cluster reduction is found in the speech of

African Americans and Whites, but both the frequency and settings of its use are more numerous among African Americans (Bailey & Thomas, 1998; Rickford, 1999). Among African American speakers in Detroit, consonant cluster reduction has been observed to have an inverse relationship to social class (W. Wolfram, 1969). In the present data, a token was coded as "1" if the first consonant in a two-consonant cluster was the only consonant that was articulated. Both consonant and stop cluster tokens were clipped, but only those tokens consisting of monomorphemic clusters that were followed by consonant sounds (as opposed to vowel sounds or pauses) were clipped. Tokens were not clipped if they were succeeded by homorganic stops. Since the occurrence of consonant clusters was very frequent in the *Eat for Life* survey script, consonant cluster tokens were only clipped for a subsample of survey sections at the beginning and end of the survey script.

The pin/pen merger describes the use of the [ɪ] variant, as in "tin", in instances where the [ɛ] variant, as in "ten", might be used before nasal consonants in other dialects. When the pin/pen merger is invoked, words such as "pin" and "pen" or "tin" and "ten" are pronounced exactly the same. The pin/pen merger is found in both AAE and Southern White vernacular speech (Bailey & Thomas, 1998; Labov, 1972), but often in greater frequency in AAE (Rickford, 1999). Tokens of [ɛ] were clipped if they were stressed and preceded a nasal consonant. Tokens of "any" and "many" were not clipped, as they were predicted to be too difficult to code accurately. Tokens representing [ɪ] were coded as "1".

The initial "th" variable refers to the use of [d] in the place of [ð]. For example, the word "the" would be pronounced as "de", while "that" would be articulated as "dat". Articulations that are stops will be coded as "1", while fricatives will be coded as "0". Only words that begin in "th" and follow pauses, vowel sounds, or liquids were clipped as tokens.

The monophthongization of [ai] refers to the articulation of the [ai] dipthong to a reduced pronunciation of the two vowel sounds. This occurs when the word "five", for example, is pronounced like "fahv". Another example is the articulation of "time" as

"tahm". This feature is found in both AAE and Southern White vernacular speech (Bailey & Thomas, 1998), but it typically occurs to greater extent among AAE speakers (Rickford, 1999). The monophthongized variant of [ai] has been associated with AAE and, when present in the current data, was coded as a "1". Tokens of [ai] were excluded from this project if they preceded voiceless construents or occurred at the end of words, due to the tendencies of most speakers to monophthongize [ai] in these settings. In several recent studies of African American comedians' portrayals of African Americans and natural African American speakers, the monophthongized variant of [ai] has been associated with lower class speakers, while the [ai] dipthong has found to be more common among middle class speakers (Nguyen, 2006; Rahman, 2007).

Kappa statistics measuring inter-rater agreement between the primary and secondary coders and the percentages of tokens of each phonological variable that required a third coding are as follows: (1) "ask" metathesis – .69 Kappa / 7.4% third coding; (2) syllable-final [d] – .36 Kappa / 7.8% third coding; (3) "-ing" endings – .77 Kappa / 6.8% third coding; (4) consonant cluster reduction – .72 Kappa / 14.1% third coding; (5) pin/pen merger – .35 Kappa / 22.2% third coding; (6) initial "th" – .12 Kappa / 26.9% third coding; and (7) [ai] monophthongization – .46 Kappa / 8.9% third coding. These Kappa statistics may seem somewhat low, particularly for the syllable-final [d], pin/pen merger, and initial "th" variables. However, from a linguistics perspective, the overall levels of inter-rater agreement are considered to be relatively high for most of the variables and still in the acceptable range for the three variables with the lowest Kappa statistics. Further, unlike the primary and secondary coders, the tertiary coder was specifically an expert in AAE. Thus, more confidence was attributed to the validity of her coding of the data, and her codes were accepted as the final codes.

Section Content

For classification purposes, the survey script was divided into 40 sections, each of which included approximately 1-8 survey questions. Each section contained items of a similar nature. For example, questions about daily fruit and vegetable intake comprised one section, questions about weekly fruit and vegetable intake comprised another section,

etc. The sections were classified as to whether or not they contained each of three types of survey content: Recruitment Sections, Racially Topical Sections, and Sensitive Sections. These sections were not exclusive; a section could be classified as having all three types of survey content.

Those sections labeled as Recruitment Sections contained scripting designed to persuade the respondent to participate in an imminent or future activity. Six Recruitment Sections were located at the beginning of the survey, during which the interviewer attempted to persuade the respondent to participate in the baseline survey. The remaining two Recruitment Sections were located at the end of the survey, when the interviewer attempted to persuade the respondent to verify his or her mailing address to receive his or her health program materials, consent to a short survey to be conducted in the succeeding few days, and provide contact information for a follow-up survey to be conducted in three months. As a percentage of total words used in the *Eat for Life* survey script, the Recruitment Sections comprised approximately 30% of words in the coded survey script.

Ten sections were coded as Racially Topical Sections. These sections contained questions that specifically queried the respondent's racial and ethnic self-identification, racial attitudes, preferred racial and ethnic self-labels, and preferences regarding interviewer race and ethnicity. A two-paragraph introduction to the ethnic identity measure was also coded as a Racially Topical Section. The Racially Topical Sections were located at the beginning of the survey, in a large section in the latter third of the survey, and at the very end. The Racially Topical Sections constituted about 26% of the words spoken in the recorded interview script.

The initial study hypotheses focused only on responding for racial survey items. However, during the coding process, the coders observed that the interviewers appeared to use AAE features more when asking respondents about their body weight and household income. These types of items are often affected by socially desirable responding and item nonresponse. Since this was an exploratory study designed primarily to examine if and when interviewers use AAE features, an evaluation of their use of AAE features when asking nonracial sensitive questions was added to the analysis plan. These observations seemed to have some support in the interviewer effects literature, which suggests that interviewer race and ethnicity effects may be more

prominent in both racial and nonracial but sensitive survey questions (see Chapter 2). Thus, a code was created for Sensitive Sections. There were few nonracial sensitive items in the *Eat for Life* survey. Those items that were coded as sensitive included items querying whether a respondent had a compromised immune system, was undergoing cancer treatment, had or had previously had various health conditions such as cancer or diabetes, current body weight, and household income. Four sections were coded as Sensitive Sections. These sections were located at the beginning, middle, and end of the survey script. Approximately 7% of the words in the recorded *Eat for Life* survey script were coded as Sensitive Sections.

Ethnic Identity

The WA interviewers, MI interviewers, and *Eat for Life* participants were each classified into one of 16 ethnic identity types using the Black Identity Classification Scale (BICS) (Davis et al., under review). The BICS contains 32, 7-point "Strongly Disagree"/"Strongly Agree" items, a single-item racial salience question, and two health material preference questions. A listing of BICS items is included as Appendix A. The 32, 7-point items are divided into six core identity component subscales: Afrocentric (7 items), Black American (8 items), Bicultural (4 items), Multicultural (4 items), Racial Salience (6 items), and Cultural Mistrust (3 items). The single racial salience question, which is separate from the Racial Salience subscale, asks "How important is being Black to your overall identity?" Responses range from zero ("Not at All Important") to ten ("Very Important"). This item was used to classify interviewers and participants as Assimilated or non-Assimilated at an early stage in the algorithm. The 16 identity types represented by the BICS include:

- Assimilated
- Black American
- Black American with Cultural Mistrust
- Bicultural
- Multicultural
- Bicultural/Multicultural
- Black American/Bicultural

- Black American/Bicultural with Cultural Mistrust
- Black American/Multicultural
- Black American/Multicultural with Cultural Mistrust
- Afrocentric/Black American
- Afrocentric/Black American with Cultural Mistrust
- Afrocentric/Bicultural
- Afrocentric/Bicultural with Cultural Mistrust
- Afrocentric/Multicultural
- Afrocentric/Multicultural with Cultural Mistrust

As indicated above, an individual may have one, two, or three identity components. A person with an Assimilated component is defined as having low racial salience and placing little importance on being a member of a racial or ethnic group. In contrast, being Black is a defining aspect of personal identity for persons with Black American, Afrocentric, Bicultural, and Multicultural identity components. Of these, the Black American and Afrocentric types are considered the most strongly affiliated with African American people and culture, or "pro-Black". A person with a Black American component is connected to Black American people and culture while an Afrocentric person feels a strong connection to Africa. In contrast, a Bicultural person tends to perceive the world as a Black/White dichotomy whereas a Multicultural person views the world as a collection of many valued cultures. Persons with Black American and Afrocentric identity components can have an additional Cultural Mistrust identity component, which is defined as a generalized mistrust of Whites and White society (Terrell & Terrell, 1981).

Other Measures

Several health and demographic characteristics were measured and used as either dependent or control variables in the various analyses. Where queried of both interviewers and *Eat for Life* participants, the question wording was identical. These additional variables included gender ("Are you male or female?"), age ("How old are you?"), income ("Approximately what was the total income of your household last year before taxes?"), and education ("What is the highest grade or degree you have completed?").

Analysis Plan

All analyses were conducted using SAS 9.1.3 for Windows (SAS Institute, 2002-2003). Descriptive statistics were first calculated to characterize the *Eat for Life* participants (n=492), WA interviewers (n=8), and MI interviewers (n=6) who participated in the study. Statistics were subsequently computed to describe the use of AAE phonological features by the WA and MI interviewer groups. Next, analyses were performed to test the three hypotheses outlined below. To clarify which participants were included in each analysis, Table 4.1 lists the participants evaluated in testing each of the three hypotheses.

Hypothesis 1

Hypothesis 1 explores the premise that African American interviewers with more pro-Black ethnic identity types will demonstrate higher use of AAE variants of phonological features. Specifically, Hypothesis 1 states that WA and MI interviewers will exhibit more frequent use of AAE phonological features if they have Afrocentric or Black American ethnic identity components than if they do not and that no effects will be found for interviewers with or without Bicultural or Multicultural ethnic identity components. Since this hypothesis is tested on a very small sample of interviewers (n=14), 32 separate logistic regression models were used. The dependent variables in these models represented the odds of the AAE variants of each of the seven phonological features occurring. For example, the dependent variable for the "ask" model was the odds that the "ask" metathesis would occur. Twenty-four models tested similarly constructed dependent variables for syllable-final [d], "-ing" endings, consonant cluster reduction, the pin/pen merger, initial "th", and [ai] monophthongization. The final four models tested the odds that any AAE feature would occur.

The main independent variables used to test Hypothesis 1 consist of binary variables representing the presence or absence of the following core BICS ethnic identity components: Afrocentric, Black American, Bicultural, and Multicultural. Each of these variables was coded as "1" if an interviewer had than identity component and "0" if she

did not. These categories were somewhat overlapping, since an individual interviewer could have up to two of these core identity types. All models controlled for interviewer education level (scored 1-8) and age (continuous). The Hypothesis 1 models were tested using the SAS proc logistic procedure. Since the sample size was low, p-values of $p \leq .05$ and $p \leq .15$ are both reported.

Hypothesis 2

Hypothesis 2 examines the use of AAE during different portions of an interview script. This hypothesis consists of three sub-hypotheses testing the WA and MI interviewers' (n=14) combined use of AAE phonological features while administering the *Eat for Life* survey script to the three standardized respondents. Hypothesis 2a predicts that the interviewers expressed more AAE phonological features during the recruitment sections of the script than during other parts of the survey script. Hypothesis 2b tests whether the interviewers exhibited higher use of these AAE features during racially topical sections than during sections that did not query racial topics. Hypothesis 2c predicts that the interviewers used more AAE phonological features during the sensitive item portions of the survey script than during sections covering other types of content.

To test these hypotheses, the SAS proc genmod procedure was used to examine the proportion of use of AAE phonological features while controlling for the clustering of data by interviewers and the three standardized respondents. As for Hypothesis 1, separate models were used to explore the odds that an AAE variant of the feature was expressed. Seven models were tested. Six of these models tested the odds of expression for the "ask" metathesis, the pin/pen merger, [ai] monophthongization, and the AAE variants of syllable-final [d] sounds, "-ing" endings, and words with an initial "th". The seventh model tested the odds that any AAE variable would occur. Since consonant cluster reduction was only coded for a subset of sections located at the beginning and end of the survey script, no model was tested for consonant cluster reduction. The occurrence of consonant clusters is exceedingly numerous in English, as well as in the *Eat for Life* survey script. Thus, tokens of consonant cluster reduction were only clipped for a few representative sections located at the beginning and end of the survey script. As a result,

the consonant cluster tokens were not sufficiently spaced throughout the survey sections to permit an analysis of their use by section type.

The independent variables in these models consisted of three binary variables representing the occurrence of AAE features during the Recruitment, Racially Topical, and Sensitive Sections of the survey script (each coded "1" if yes and "0" if no). Age and education have been associated with the use of AAE. Thus, all seven models controlled for the standardized respondent being surveyed (coded 1-3), interviewer education status (coded 1-8), and interviewer age (continuous). But, since interviewer age and income were highly correlated (0.83) and age has been suggested as a strong correlate of language use among African Americans, interviewer income level was not included in the models.

Hypothesis 3

Hypothesis 3 presumes that interviewers' use of AAE variants of phonological features will be positively associated with more pro-Black reporting on racially topical survey items by *Eat for Life* program participants. This hypothesis was tested by modeling the impact of use of each of the seven AAE phonological features and the AAE Index variable on *Eat for Life* participants' means of the following five BICS subscales: Afrocentric, Black American, Bicultural, Multicultural, and Racial Salience. Each subscale mean was scored from lowest to highest (1-7). The main independent variable in each model was the percent of use of a specific AAE phonological feature by the WA interviewers (n=8) or an overall AAE Index score. It was previously intended that a summary variable would be constructed to indicate use of the seven AAE features combined, thereby representing overall use of AAE as a single underlying construct. However, preliminary analyses revealed that the bivariate correlations among the seven features were generally very low. The correlation of the "ask" metathesis and the AAE variant of "-ing" endings was 0.63. But, no other correlations were higher than 0.35, and several correlations were negative. Thus, the data did not support the premise that the seven individual phonological features represent a single construct. However, a combined index variable was still believed to be useful as an indication of overall use of

the seven independent phonological features. An AAE index variable was therefore constructed to represent the proportion of overall use of the seven phonological variables across all coded tokens. This index variable was created by summing all of the "1" codes for all seven phonological features and dividing this sum by the total number of tokens for all features. The resulting variable is called AAE Index.

It should be noted that the linguistic variables were creating using the interviewers' expression of AAE during the standardized recordings as proxies for their expression of AAE during the live *Eat for Life* interviews. Each proxy variable represented an interviewers' use of the AAE feature in question as an average across all of the coded tokens in the three standardized interviews. Thus, these variables do not indicate use of AAE features during specific BICS subscale administrations. This latter approach was impossible, because each feature would have had to occur as code-able tokens in the wording for items from each BICS subscale. Since tokens did not occur in every BICS subscale wording, the AAE variables therefore represent the overall language context in which the live *Eat for Life* interviews likely occurred. For example, the first model tested explored the influence of a proxy variable representing the percent of use of interviewers' overall use of the "ask" metathesis during the three standardized interviews on *Eat for Life* participants' mean scores on the BICS Afrocentric subscale.

A final model was run for each of the eight dependent phonological variables to explore whether an interviewer having a Black American core ethnic identity component was positively related to *Eat for Life* participants' reporting on the BICS subscale means. The Black American ethnic identity type was selected because it was the most prevalent core ethnic identity component among the WA interviewers. Exactly half of the WA interviewers had the Black American identity component, and exactly half did not. The Black American ethnic identity variable was coded as a binary variable with "1" representing the presence of the Black American component and "0" representing its absence.

Forty-five separate models were tested. Due to the large numbers of models under scrutiny, a more stringent significance level of $p = .01$ was applied. All models controlled for *Eat for Life* participants' gender (coded 0/1), age (continuous), education level (coded 1-8), and income (coded 1-7). Unlike the pattern found in the interviewer

data, *Eat for Life* participant education and income were not highly correlated, which allowed both variables to be in the models.

In order to control for the clustering of *Eat for Life* participant data by interviewers, 29 of the models were tested as linear mixed models using the SAS proc mixed procedure. However, the remaining models were associated with almost no variability across the eight interviewers, which both prohibited and obviated the use of a clustered model. These 16 models were tested using ordinary least squares regression via the SAS proc reg procedure.

RESULTS

Description of the Study Sample

Selected sociodemographic characteristics for the 492 *Eat for Life* participants who were included in the present analyses are described in Table 4.2. The mean age of the participants was 48.8 years. The participants were approximately evenly divided between the Detroit and Atlanta health care plans. The majority of participants were female (70.9%), had a high school-level education or higher (97.3%), and lived in a household earning $20,001-$60,000 per year (56.3%). The percent of *Eat for Life* participants with each of the six core ethnic identity components was as follows: 13.0% Assimilated, 31.5% Afrocentric, 54.7% Black American, 37.6% Bicultural, 30.0% Multicultural, and 12.0% Cultural Mistrust (participants could be in more than one category).

Eight WA interviewers and six MI interviewers participated in the current study. All of the WA and MI interviewers were female and identified themselves as Black or African American and non-Hispanic. However, one WA interviewer and one MI interviewer identified herself as being biracial, with American Indian listed as the second race. No other races or ethnicities were reported. On average, the interviewers were younger than the *Eat for Life* participants with a mean age of 34.6 years and 35.0 years for the WA and MI interviewers, respectively. All of the WA and MI interviewers had a high school-level education or higher. Only two of the WA and two of the MI interviewers lived in a household earning more than $60,000. Of the 16 BICS ethnic

identity types, the WA interviewers were distributed in the following manner: 1 Assimilated, 2 Black American/Bicultural, 1 Afrocentric/Black American, 1 Bicultural, 1 Black American/Bicultural, 1 Afrocentric/Multicultural, and 1 Multicultural. The MI interviewers were comprised of the following 16 BICS ethnic identity types: 1 Black American, 1 Afrocentric/Black American, 1 Bicultural, 1 Afrocentric/Multicultural with Cultural Mistrust, 1 Bicultural/Multicultural, and 1 Multicultural.

Use of AAE Phonology in the Study Sample

As shown in Table 4.3, use of AAE phonology varied considerably across the seven phonological features under scrutiny. The variable with the highest percentage of use per token for both the WA and MI interviewers was consonant cluster reduction. Among occurrences of consonant clusters, the mean occurrence of the AAE consonant cluster reduction feature per total tokens of this feature was 44.8% for the WA interviewers and 47.5% for the MI interviewers. Since consonant cluster reduction was the only variable that was not clipped throughout the entire interview script, consonant clusters comprised the lowest number of tokens in the dataset. After consonant cluster reduction, the percentages of use of the remaining six AAE phonological variants decreased considerably. The WA interviewers expressed the pin/pen merger in 27.1% of possible occurrences and the AAE variant of "-ing" endings in 16.9% of tokens. Comparable percentages for these variables for the MI interviewers were 17.7% for the pin/pen merger and 20.6% for the AAE variant of "-ing" endings. The WA interviewers exhibited far less use of the "ask" metathesis than the MI interviewers at 8.3% versus 17.5%, respectively. The two groups also varied in their expression of [ai] monophthongization, which ranged from 12.6% for the WA interviewers to 4.1% for the MI interviewers. However, usage of the AAE variants for syllable-final [d] and initial "th" sounds was similar across the two groups. The AAE variant for syllable-final [d] occurred in 6.2% of tokens for the WA interviewers and in 8.5% of tokens for the MI interviewers, while the initial "th" AAE variant was used in 12.3% and 16.3% of tokens for the WA and MI interviewers, respectively.

The range of use of the AAE phonological features had wide variation among interviewers within the two groups, but overall usage was below 50% for all 14 interviewers for the seven phonological features. The distributions were most extreme for the "ask" metathesis, which was not used at all by six of the WA interviewers or two of the MI interviewers. All of the remaining features were used by all of the WA interviewers, but the percentages were generally low. None of the WA interviewers used the AAE variant of syllable-final [d] or initial "th" more than 20% of the time. The highest expression of [ai] monophthongization among the WA interviewers was 25.4%. Use of AAE "-ing" endings, consonant cluster reduction, and the pin/pen merger was below 50% for the majority of the WA interviewers. Usage of these variables was different but similarly low for most of the MI interviewers. Only two of the MI interviewers exhibited [ai] monophthongization with any frequency, at 10.5% and 12.1%, respectively. Usage of the pin/pen merger, the AAE variant of initial "th", and the AAE variant of syllable-final [d] was below 50% for all of the MI interviewers. One MI interviewer had high expression of the AAE variant for "-ing" endings (84.1%); however, expression of this variable among the remaining five MI interviewers ranged from 0-21.1%. Use of consonant cluster reduction among the MI interviewers had a tighter distribution ranging from 35.7-66.7%.

Although not shown, the data also indicate inconsistent patterns in use of the seven variables within each individual speaker. All interviewers use some features more than other features, and no one interviewer dominates the highest usage for more than three or four variables. Overall, these data suggest a complex pattern of phonological expression.

The percentage of use of the seven phonological features represented by the AAE Index variable is also listed in Table 4.3. The AAE Index score for the WA interviewers was 15.5%. This score was slightly higher than the score for the MI interviewers, which was 14.5%. However, the MI interviewers had a much wider range of index scores.

Interviewer Ethnic Identity Type and Use of AAE Phonology

Hypothesis 1 presumed that the WA and MI African American interviewers' ethnic identity types would be related to their use of AAE variants of phonological features. It was anticipated that interviewers with Afrocentric or Black American ethnic identity components would exhibit higher expression of AAE features than interviewers without these ethnic identity components. Conversely, no differences in AAE use were expected between interviewers with and without Bicultural or Multicultural ethnic identity components. The results of the 32 models testing Hypothesis 1 are listed in Table 4.4 and suggest a complicated picture that varies across the seven phonological features, as well as the odds of using any of the seven features.

Interviewers with an Afrocentric ethnic identity component are significantly more likely to express the pin/pen merger ($p < .0001$), significantly more likely to express the initial "th" feature ($p = .02$), and borderline more likely to use [ai] monophthongization ($p = .13$) than interviewers without an Afrocentric identity component. However, interviewers with an Afrocentric identity component are significantly less likely to use the AAE variant of "-ing" endings ($p < .0001$) than interviewers without this component, and no difference is observed for use of the "ask" metathesis, syllable-final [d], or consonant cluster reduction between interviewers with and without the Afrocentric identity component. The odds of using any AAE phonological feature were significantly lower for interviewers with an Afrocentric identity component than interviewers without an Afrocentric component ($p = .003$), providing overall evidence against Hypothesis 1.

Interviewers with a Black American ethnic identity component are significantly more likely to exhibit the initial "th" AAE variant ($p = .03$) and borderline more likely to use [ai] monophthongization ($p = .15$) than interviewers without a Black American identity component. In contrast, interviewers with a Black American identity component are significantly less likely than interviewers without a Black American identity component to express the AAE variants of "-ing" endings ($p < .0001$) and the pin/pen merger ($p = .03$). No other significant patterns were found for the Black American identity type for individual features. However, interviewers with a Black American identity component were overall less likely to use one of the phonological features

studied than interviewers without this component ($p < .0001$), thereby contradicting Hypothesis 1.

Interviewers with the Bicultural ethnic identity component exhibited significantly higher use of the AAE variants for "-ing" endings ($p < .0001$), initial "th" ($p = .002$), and [ai] monophthongization ($p = .04$) than interviewers without this component. Bicultural interviewers were also borderline more likely to express consonant cluster reduction ($p = .12$). No other significant contrasts were evident. Overall, these findings indicate higher use of AAE features among Bicultural interviewers, which is opposite from the hypothesized pattern. The findings for overall use of any AAE feature also contradict Hypothesis 1. Interviewers with a Bicultural identity component were significantly more likely to use any of the seven AAE features than interviewers without this component ($p < .0001$).

A complex picture also emerged for the Multicultural interviewers. Interviewers with a Multicultural ethnic identity component were significantly more likely to use the "ask" metathesis ($p < .0001$) and the AAE variant for "-ing" endings ($p < .0001$) but significantly less likely to express the AAE variant of initial "th" ($p = .007$) or [ai] monophthongization ($p < .0001$) than interviewers without a Multicultural identity component. No pattern emerged for use of syllable-final [d], consonant cluster reduction, or the pin/pen merger. However, the overall model of language use indicates that interviewers with a Multicultural identity component were significantly more likely to use AAE features than interviewers without this identity component ($p < .0001$). This finding is the opposite of that which was expected.

Although not shown in Table 4.4, several statistically significantly effects were found for the control variables. In the Multicultural interviewer model, interviewers with higher education were significantly less likely to exhibit the "ask" metathesis. Use of the "ask" metathesis was also inversely related to age for the Afrocentric, Black American, and Bicultural interviewer models. Use of the AAE variant for "-ing" endings increased with interviewer education for the Afrocentric, Black American, and Bicultural models and with interviewer age for all "-ing" ending models. Use of the AAE variant for initial "th" sounds also increased with interviewer education and age for all models. However, use of the pin/pen merger decreased with increasing interviewer education and age across

all models. Expression of [ai] monophthongization also decreased with interviewer age, but only for the Multicultural interviewer model. No significant control variable effects were found for models testing use of the AAE variants for syllable-final [d] or consonant cluster reduction.

Interviewers' Use of AAE During Recruitment, Racially Topical, and Sensitive Survey Sections

Hypotheses 2a, 2b, and 2c predicted that the 14 WA and MI African American interviewers participating in this study would express higher use of AAE phonological features during parts of the survey script containing recruitment, racially topical, and nonracial but sensitive survey content, respectively, than during sections of the script not querying each of these types of content.

The total numbers of tokens and the numbers of tokens located in the Recruitment Sections, Racial Sections, and Sensitive Sections of the *Eat for Life* survey script are listed in Table 4.5. Since consonant cluster reduction was not modeled for Hypothesis 2, those numbers are not included in Table 4.5.

The results of testing Hypothesis 2a are detailed in Table 4.6. The odds of use of syllable-final [d] ($p = .003$), AAE "-ing" endings ($p=.02$), and [ai] monophthongization ($p = .003$) were significantly lower during the Recruitment Sections than in non-recruitment parts of the survey script. Decreased expression of the "ask" metathesis ($p = .10$) and the pin/pen merger ($p=.07$) was borderline significant. No change in the use of initial "th" was indicated for the Recruitment Sections versus non-recruitment sections of the script. The model testing overall use of any AAE feature during Recruitment Sections was also nonsignificant.

Findings for Hypothesis 2b were mixed (Table 4.6). Only one feature, initial "th", was expressed significantly more during Racially Topical Sections than in non-racial survey sections ($p = .02$). Syllable-final [d] ($p = .001$) and [ai] monophthongization ($p = .0001$) were used significantly less during Racially Topical Sections than during non-racial sections. No significant patterns of use were evident for the "ask" metathesis, "-ing" endings, or the pin/pen merger. No significant effects were

found when comparing overall use of any AAE feature between Racially Topical Sections and non-racial sections.

In line with the coders' observations and Hypothesis 2c, the interviewers in this study did exhibit a higher use of some AAE features during Sensitive Sections than during non-sensitive parts of the survey script (Table 4.6). Expressions of the "ask" metathesis ($p = .03$), syllable-final [d] ($p = .0006$), and the pin/pen merger ($p < .0001$) were all significantly more frequent during the Sensitive Sections of the script than during non-sensitive parts of the script. No significant effects were found for "-ing" endings, initial "th", or [ai] monophthongization. In addition, no significant effects were found for the model testing overall use of any AAE feature.

These analyses also explored whether the interviewers' use of AAE phonological features varied according to which of the three standardized respondents they were interviewing. Only one significant difference in expression was found across the three standardized survey respondents: the interviewers used initial "th" significantly less when surveying the respondent with low use of AAE than when interviewing the two respondents with high use of AAE (results not shown). Only two AAE features were impacted by interviewer education level. The pin/pen merger was used less as interviewer education increased, and initial "th" was used more as interviewer education increased. Older interviewers were significantly less likely to express the "ask" metathesis and borderline significantly more likely to use the AAE variant of initial "th".

Overall, these numbers provide no support for Hypotheses 2a or 2b but limited support for Hypothesis 2c. The findings suggest a pattern of lower use of AAE phonological features during Recruitment Sections, which is opposite from that which was predicted by Hypothesis 2a. However, this pattern was not supported when assessing overall use of any AAE feature. No consistent pattern of use of AAE phonological features is observable during Racially Topical Sections (Hypothesis 2b). However, a pattern of higher use of AAE variants emerged for Sensitive Sections for half of the phonological features assessed, which provides partial support for Hypothesis 2c. These mixed findings indicate substantial variance within AAE phonological features, which is most evident in the varied pattern observed for the racially topical survey segments.

Association Between Interviewers' Use of AAE and *Eat for Life* Participants' Reporting on Racially Topical Survey Items

Results of the 45 models testing associations between WA interviewers' use of AAE phonological features and *Eat for Life* participants' mean scores for five of the BICS subscales are detailed in Table 4.7. Table 4.7 also documents findings for the eight models examining the relationship between whether interviewers had a Black American ethnic identity component and *Eat for Life* participants' mean scores on the same five BICS subscales. None of the 45 models tested was statistically significant at the $p \leq .01$ level, indicating that neither interviewers' use of the seven AAE phonological variables examined in this study nor interviewers' Black American ethnic identity status had an impact on *Eat for Life* participants' answers to racially topical survey items.

DISCUSSION

The relationship between African American telephone interviewers' ethnic identity types and use of AAE-associated phonological features suggests a complex picture of language use that varies by phonological feature (Table 4.4). After controlling for interviewer education and age, the findings suggest that interviewers with Bicultural or Multicultural identity components may exhibit higher overall expression of AAE phonological features. Multicultural interviewers were also the only group to exhibit significantly higher use of the "ask" variable, which is the most stigmatized and, to non-linguists, the most recognizable AAE feature that was included in the study. Use of this feature alone may have been sufficient to convey African American in-group social identity. The AAE variant of "-ing" endings, which was used more by both the Bicultural and Multicultural identity types, is a generally stigmatized phonological variable that is strongly associated with lower social class across racial and ethnic groups. Thus, it is interesting that, even after controlling for interviewer education and age, the Bicultural and Multicultural interviewers used this feature more. Interviewers with an Afrocentric identity component were more likely to use the pin/pen merger, AAE variant of initial "th", and [ai]monophthongization. But, the overall evidence contradicts the hypothesis that Afrocentric interviewers would use AAE features more than interviewers

without this identity type. The available evidence also contradicts the hypothesis that interviewers with a Black American identity component would exhibit higher use of AAE features than those without the Black American component.

Williams (1968) noted that a skilled interviewer is able to simultaneously enact two separate and often contradictory roles: (a) creating and maintaining rapport with a respondent, and (b) conveying sufficient objectivity so as not to sway a respondent's answers. The interviewers in this study were all highly trained, experienced professional telephone interviewers at respected survey call centers. As such, it is likely that this sample of interviewers was skilled in balancing the competing tasks of conveying objectivity and rapport-building with survey respondents. The findings from this study indicate that these interviewers may have used AAE to fulfill their dual survey roles. Bicultural and Multicultural interviewers may have used AAE phonology to convey their African American in-group membership in order to build rapport and put African American respondents at ease. In contrast, interviewers with Afrocentric and Black American identity components may not have felt a need to "prove" themselves as in-group members. In fact, interviewers with these identity components may have purposefully avoided the most stigmatized AAE phonological features – the "ask" metathesis and "-ing" endings – in order to convey their objectivity to African American survey respondents. However, it must be cautioned that all interpretations of these data are tentative, due to the small number of interviewers included in these analyses. Further research is warranted to better understand whether and how African American telephone interviewers use language cues to fulfill competing role objectives. It would also be interesting to explore whether African American interviewers intentionally use linguistic features to accomplish specific interviewing tasks.

The examination of when African American interviewers use AAE phonological features during a survey interaction also yielded a complicated pattern that varied by linguistic feature (Table 4.6). However, even with a limited sample size, some broad trends were evident. The strongest trend was that after controlling for interviewer effects, respondent effects, interviewer education level, and interviewer age, the interviewers used certain AAE phonological features significantly less during the recruitment sections of the survey script. This pattern is the opposite from that which was predicted. The

recruitment sections were located at both the very beginning and very end of the survey script. So, this pattern of lower use is not merely the result of language use as it unfolds in the first minutes of a conversation with a stranger. These findings suggest that interviewers use specific linguistic features to cue more formal, less stigmatized, more educated, and more middle to upper class ASE speech when trying to persuade respondents to participate in research activities. Future research might consider investigating whether interviewers who use these speech patterns yield higher survey response rates.

African American interviewers' use of AAE was hypothesized to increase during racially topical sections of the survey script. However, the findings do not support this hypothesis. In fact, only one AAE phonological feature – initial "th" – was exhibited more during racially topical portions of the interviews than during other parts of the survey script. Use of two AAE features significantly decreased, and no differences were found for the three remaining features, including the highly stigmatized "ask" metathesis. No significant effects were found for the model testing overall use of any AAE feature. This lack of a consistent overall pattern may actually be meaningful, as it suggests that the interviewers neither strongly gravitated to use of AAE or ASE during these types of survey sections. Since it is probably well understood among many African Americans that dialect choice is a laden topic, these results may imply that African American interviewers use language features to particularly convey objectivity when in the act of administering racial questions to African American respondents. Thus, the interviewers in this study may be exhibiting the deft use of language as a skill in fulfilling their interview roles.

It was further hypothesized that African American interviewers would use AAE phonological features more often during sections of the survey script containing traditionally sensitive items such as weight and income. This hypothesis was partially borne out. Half of the AAE features assessed were used significantly more when the interviewers administered nonracial sensitive survey sections than during other sections of the script. These three features include the stigmatized "ask" metathesis, the pin/pen merger, and the AAE variants of syllable-final [d], which is the only variable in this study that is solely affiliated with AAE. No effects were found for the remaining three

linguistic features or the model testing overall use of any AAE feature. The interviewers' increased use of some of the more prominent phonological features of AAE may suggest that they were using these features to build rapport and establish trust when querying respondents about sensitive topics. It is likely that the interviewers expected these items to be associated with more socially desirable responding and higher item nonresponse. However, additional research is warranted to determine whether use of in-group linguistic cues helps build rapport and put African American respondents at ease.

This study found no evidence that African American telephone interviewers' use of seven AAE-associated phonological variables impacts African American respondents' reporting on racially topical survey items (Table 4.7). There are several possible explanations for these findings, which were based solely on the language use displayed by the eight WA interviewers. The first explanation is that the WA interviewers' use of the AAE variants of the phonological features under study (Table 4.3) was too low to have an impact on respondents' answers. The mean percentage of use of each of the seven phonological features measured was below 45%. And, in most individual interviews, use of the features studied was far below 50%. The AAE Index variable also reflected a low percentage of overall use at 15.5% for the WA interviewers. Although cues were provided that the WA interviewers were African American in the introductory survey script, the WA interviewers' use of AAE phonological characteristics may have been too low to convey strong within-group cultural identity. The linguistics literature provides few statistics-based estimates of use of individual phonological features associated with AAE, but, what estimates exist suggest that the WA interviewers' use of AAE phonological features may have been below average. In his research with African American speakers in Detroit in the 1960s, Wolfram (1969) found a 51-84% occurrence for consonant cluster reduction, with higher use inversely correlating with social class. Thirty years later, Catherine Chappell's 1999 replication of the Wolfram study found that lower- and middle-class African American speakers in Oakland exhibited consonant cluster reduction in 72-84% of occurrences (as cited in Rickford & Rickford, 2000). These estimates are substantially higher than the percentage of use of consonant cluster reduction for the WA interviewers, which ranged from 38.5-66.7%. A 1992 study reported a 60% use of [ai] monophthongization among working class African American

speakers in Detroit (Edwards, 1992). This estimate of [ai] monophthongization is also lower than the usage found in the current study, which ranged from 4.6-25.4% for the WA interviewers. A case study comparing language use between two African American brothers reported percentages of use of the AAE variant of "-ing" endings of 10% and 86% (Valentin-Marquez & Nguyen, 2002, October, 2004, April), which is also higher than the range of 0.3-66.4% found for the WA interviewers in this study. The low use of AAE by the WA interviewers may be attributable to the fact that they are all Washington State residents. Use of AAE was at least slightly higher among the MI interviewers for five of the seven phonological variables. Data on the interviewers' geographical origins were not collected, and no linguistic analyses have been published on African Americans living in Washington State. In fact, little data is available at all on regional variation in AAE use. Additional research on how AAE varies by geography would help parse out regional influences from other sociodemographic characteristics. The current study found great variability of use of AAE variants of phonological features not only across, but also within the WA interviewers. There was also no consistent pattern of increased or decreased AAE use during the racial survey sections. As a result of this overall varied and low usage of AAE, in combination with an ASE-worded script and the loaded implications of AAE use, the respondents in this study may not have heard strong or consistent enough phonological cues to suggest a clear direction for socially desirable reporting. A more robust and uniform pattern of use of AAE phonology may have been required to convey a strongly pro-Black African American ethnic identity.

A second explanation for the nonsignificant findings for the impact of AAE use on survey data is that AAE phonology alone, in the absence of other AAE features and communication norms, was insufficient to convey in-group cultural identity. Phonology is an important component of language. But, a language is also defined by distinctive syntax, vocabulary, meaning, and communication norms, and some authors have suggested that AAE is most distinguishable from other English dialects in these latter aspects (Bailey & Thomas, 1998). The interviewers who participated in this study were well-trained professionals. They did not interject much of their own wording into the interview script, and, as a consequence, a different population of interviewers may have been more likely to use nonstandard syntactic and discourse features.

It is also possible that the WA interviewers' low use of AAE phonology was sufficient to convey in-group membership and put the *Eat for Life* participants at ease. All participants were cued that their interviewer was African American. And, as seen in Chapter 3, having an African American interviewer was important to many *Eat for Life* participants. Given this context, it may be that a low threshold of AAE use was required for the WA interviewers to demonstrate their in-group membership. As a result, participants may have felt sufficiently comfortable to answer racially topical survey questions without socially desirable editing.

Conversely, the WA interviewers' low and selective use of AAE phonological features may have conveyed two messages that worked in unison to establish a single tone: (a) that the interviewers were members of the in-group but (b) that the interview interaction was not an appropriate setting for in-group expression. Since they were the first speakers in the interactions and utilized a standardized survey script, it is likely that the interviewers established the linguistic norms for the survey interactions. And, since the *Eat for Life* survey was designed from an ASE frame of reference, it is likely that ASE was conveyed as the appropriate dialect for the interactions. Many *Eat for Life* participants reported strongly racially salient answers to racial survey items. However, it is possible that some participants were affected by the interviewers' language use and would have reported more pro-Black answers if the survey script had been designed using AAE. This question cannot be addressed with the present data but could be explored in a survey using varied dialect expression. Whether or not survey designers wish to establish a particular dialect environment for surveys with African Americans, however, is discussed below.

Lastly, it is possible that African American interviewers' language choices have no impact on reporting by African American survey respondents. It is reasonable to assume that many African Americans are accustomed to interacting in ASE language situations. This may be particularly true for this study population, which was recruited from the memberships of two integrated health care delivery systems and was primarily middle aged with some college education. This population, as well as many other populations of African Americans in the U.S., may therefore be less influenced by survey language context than other populations who might be less accustomed to interacting in

multiple dialects. If dialect choice has no impact on African American respondents' reporting of survey data, then survey designers have one less source of measurement error to consider.

This study also found no evidence that an African American interviewer's Black American ethnic identity status impacts racially topical or sensitive survey data from African American respondents. However, this study was a limited test of this hypothesis, as it only examined a single ethnic identity type and used a small sample of interviewers. Additional research with a larger sample of interviewers is needed to determine whether other, within-group ethnic identity types influence a wider range of types of survey data. The influence of matching interviewers and respondents by ethnic identity type should also be explored. The outcomes of the *Eat for Life* trial indicate that African American respondents who receive ethnically matched dietary health intervention materials may be more likely to increase their fruit and vegetable consumption than group-targeted intervention materials (Resnicow et al., 2008). Ethnically matched interviewers may have a similar impact on survey data.

This research has a number of limitations. Foremost among these limitations is the fact that this study was conducted within the constraints of a larger study, the *Eat for Life* trial. The research objectives of the two studies were not always in concert, and the design of the current study had to adapt to the *Eat for Life* study design. For example, both the number of survey respondents and the sample of interviewers who were available to be recruited for this study were limited by the parameters established by the *Eat for Life* trial. Where possible, all statistical models in the present analyses controlled for individual interviewer effects. However, only tentative conclusions may be drawn from such a small sample, which likely represented a limited range of speakers. It is also possible that the inclusion of male interviewers may have altered the study findings. Male speakers tend to have different speech characteristics, and the involvement of a male interviewer may have different implications for the survey interaction.

The survey topics studied may also have influenced study findings. Prior research indicates that interviewer race and ethnicity effects may be most prominent in mixed-race interviewer-respondent pairings and for racially topical survey items in which the respondent is queried about his attitudes toward members of another race (see Chapter 2).

The African American respondents in this study may have been less sensitive to socially desirable responding for questions about African American culture and people, such as those contained in the present dataset, than they would have been to racial attitude items querying their attitudes toward Whites. The BICS measure itself may not have been as vulnerable to socially desirable response editing as other, pre-existing measures of racial and ethnic identity. As described in more detail elsewhere (Davis et al., under review), the BICS was developed for administration by healthcare delivery systems to their customers. Thus, any survey items or language that was considered as potentially inflammatory was vetted and removed from the BICS scale. Stronger differences in interviewers' dialect use and respondents' reporting to racial and ethnic survey items may emerge for a survey containing more sensitive racial content. Future research should also consider exploring the influence of language use and ethnic identity on more sensitive nonracial data, as recent research indicates that interviewer race and ethnicity effects may be particularly operant in surveys of sensitive topics such as substance use and physical abuse (see Chapter 2).

This study used proxy variables for language use that were constructed using data from standardized interviews. It is possible that interviewers' language use varied between the standardized and live *Eat for Life* interviews. Additional research should also examine whether language use has a more immediate and temporary effect on survey data. In this study, language use was operationalized as use of AAE phonological features in survey sections and the interviews as a whole. It may be, however, that how an interviewer delivers a specific question is more influential than the overall language context of the interview. Studying language use and effects at this micro level was not possible in this study, as this could only be achieved though the use of a script where appropriate phonological variables were embedded in the wording of every survey item of interest. These findings are also tempered by the limited range of language features studied. The use of phonological features in isolation of other language components such as syntax and vocabulary may have a negligent impact on survey data. It may be, too, that other phonological variables such as vowel sounds may be more important in conveying AAE language use. As previously mentioned, this study may have been adversely impacted by the low use of AAE features among the WA interviewers.

However, this study did test the likely use of AAE phonology in real-world interview contexts. And, as such, the lack of significant findings on survey data may present glad tidings to survey designers concerned about measurement error.

Although the WA interviewers' use of AAE phonological features in this study was relatively low, it should be noted that this low usage is likely a realistic depiction of the typical use of AAE in standardized, scripted telephone surveys. As previously discussed, many African Americans view ASE as the language to be used in conversations with strangers, formal interactions, and linguistic interchanges with outsiders to one's community and/or cultural group (Ogbu, 1999). Further, it is likely that most, if not all, surveys are written in ASE and assume ASE communication norms. But, should survey designers consider the use of other English dialects when communicating with African American populations? Numerous linguists have noted the potential for miscommunication between ASE and AAE speakers (Hansell & Ajirotutu, 1982; Rickford, 1999; Smitherman, 2000; Speicher & McMahon, 1992). Differences in language and communication norms may impact respondents' interpretation of survey questions, translation of the intensity of response scale labeling, estimation of socially desirable responses, and communication of answers. Dialect differences may also color interviewers' interpretation and recording of respondents' answers. For example, if speaker-audience interactions determine the meaning of words in AAE, an AAE-speaking respondent or interviewer in a mixed-dialect dyad may misinterpret something voiced by the other. Dialect miscommunications may further influence researchers' interpretation and analyses of survey data, particularly if the researchers are monodialectical. Another consideration in dialect choice is the range of expression desired. Several researchers have noted that AAE has distinctive vocabulary and syntax that may provide communication options that have no meaningful equivalents of expression in ASE. Thus, it is possible that the use of ASE as the language norm constrains respondents' expression of the meaning or intensity of their answers to survey questions. A further consideration is whether survey designers wish to establish an ASE or other communication environment for a particular survey. For instance, would a survey yield substantially different findings if conducted in AAE verses ASE?

All of these language factors are considered – and, often, assumed to matter – when making similar language choices in survey designs for other populations. Spanish and English versions of surveys are routinely offered for Mexican American populations. But, what is the real impact of conducting surveys with African Americans in ASE? If American ASE speakers were surveyed in British English, chances are that differences in vocabulary and communication norms would complicate the communication process. Further, the experience of answering a survey administered in British English might put an American respondent in a different frame of mind as he constructs his answers and self-edits them for the interviewer. All of these factors might affect the survey data obtained. The appropriateness of the sole use of ASE in surveys is likely related to the dialectical characteristics of the respondent population. For bidialectical African Americans, language choice may not be critical. However, linguists have estimated that there is a large population of monodialectical AAE speakers in the U.S. And, for these respondents, use of ASE versus AAE may have a significant impact on the survey interaction.

No answers to questions about dialect appropriateness are available in the extant scientific literature. What is clear is that many African Americans have strong feelings about language choices. Exclusive use of AAE or ASE may convey strong messages to African American respondents that may have an impact on the survey process, and this possibility warrants further research. However, any research involving AAE is bound to be highly political (Morgan, 1994). Thus, research on language use in surveys with African Americans should proceed with great sensitivity and caution. A first essential step may be to explore how African American respondents themselves feel about dialect choices in the survey interaction. Use of AAE can be highly controversial among AAE and ASE speakers alike, as evidenced by the 1996 Oakland school district Ebonics incident. But, in contrast with school-based education, the goal of surveys is not to change participants, but to understand how they think, feel, and live. To this end, the onus is on us as survey designers to create appropriate environments where respondents of all cultural backgrounds will feel safe in communicating their knowledge and trust that they will be heard.

Table 4.1. Study Participants Included in Hypothesis Testing[1]

	Washington Interviewers (n=8)	**Michigan Interviewers (n=6)**	***Eat for Life* Participants (n=492)**
Hypothesis 1	X	X	
Hypothesis 2	X	X	
Hypothesis 3	X		X

[1] All hypotheses used variables representing the use of African American English phonology during the recorded interviews that each interviewer completed with the three standardized respondents. However, no survey data from the standardized respondents were used in any analyses.

Table 4.2. Descriptive Statistics for Study Participants

	***Eat for Life* Participants (n=492)**	**Washington Interviewers (n=8)**	**Michigan Interviewers (n=6)**
Race and Ethnicity			
Black or African American (%)	100	100	100
American Indian (%)	0	12.5	16.7
Other Race or Ethnicity (%)	0	0	0
Female (%)	70.9	100	100
Mean Age in Years (SD)	48.8 (10.9)	34.6 (13.9)	35.0 (14.1)
Health Plan Affiliation (%):			
Detroit	46.7	NA	NA
Atlanta	53.3	NA	NA
Educational Status (%):			
Less Than High School	2.7	0	0
High School or GED	23.8	25.0	25.0
Training Other Than College	6.3	0	0
Some College/2-Year College Graduate	31.7	50.0	50.0
College Graduate	19.8	25.0	25.0
Graduate School	15.8	0	0
Income (%):			
$20,000 or Less	8.0	25.0	25.0
$20,001 to $40,000	28.3	25.0	25.0
$40,001 to $60,000	28.0	25.0	25.0
$60,001 to $80,000	17.8	12.5	12.5
More Than $80,000	17.8	12.5	12.5
Ethnic Identity (% with Component): [1]			
Assimilated	13.0	12.5	0
Afrocentric	31.5	25.0	33.3
Black American	54.7	50.0	33.3
Bicultural	37.6	37.5	33.3
Multicultural	30.3	37.5	50.0
Cultural Mistrust	12.0	0	16.7

[1] Participants may be classified as having more than one ethnic identity component.

Table 4.3. Percent of Use of African American English Variants of Seven Phonological Features by African American Interviewers from Washington and Michigan (n=14) (Means and Ranges)

	Washington Interviewers (n=8)	Michigan Interviewers (n=6)
"Ask" Metathesis (n=646)	8.3% (0-63.6%)	17.5% (0-64.6%)
Syllable-Final [d] (n=2246)	6.2% (1.2-11.5%)	8.5% (0-18.0%)
"-ing" Endings (n=4073)	16.9% (0.3-66.4%)	20.6% (0-84.1%)
Consonant Cluster Reduction (n=198)	44.8% (38.5-66.7%)	47.5% (35.7-66.7%)
Pin/Pen Merger (n=2452)	27.1% (2.3-50.9%)	17.7% (8.8-25.7%)
Initial "th" (n=2304)	12.3% (5.1-18.6%)	16.3% (2.5-37.3%)
[ai] Monophthongization (n=2924)	12.6% (4.6-25.4%)	4.1% (0.4-12.1%)
Use of Any AAE Feature	15.5% (9.1-24.0%)	14.5% (2.6-34.3%)

<u>Table 4.4</u>. Tests of Association Between Interviewers' (n=14) Ethnic Identity Type Components and Their Use of African American English (AAE) Variants of Seven Phonological Features Using Logistic Regression (Odds Ratio Estimates and 95% Confidence Limits) [1]

AAE Feature (n=Total Number of Tokens)	Afrocentric Ethnic Identity Component	Black American Ethnic Identity Component	Bicultural Ethnic Identity Component	Multicultural Ethnic Identity Component
"Ask" Metathesis (n=646)	0.790 (0.438-1.423)	0.918 (0.518-1.628)	0.785 (0.470-1.312)	79.901** (18.494-345.205)
Syllable-Final [d] (n=2246)	0.928 (0.642-1.343)	1.013 (0.715-1.434)	1.235 (0.863-1.767)	1.000 (0.709-1.410)
"-ing" Endings (n=4073)	0.266** (0.212-0.335)	0.305** (0.252-368)	2.800** (2.303-3.404)	11.736** (9.412-14.634)
Consonant Cluster Reduction (n=198)	0.795 (0.411-1.537)	0.858 (0.478-1.540)	1.678 * (0.869-3.241)	0.810 (0.440-1.490)
Pin/Pen Merger (n=2452)	1.962** (1.581-2.434)	0.801** (0.656-0.977)	0.941 (0.755-1.173)	1.008 (0.817-1.244)
Initial "th" (n=2304)	1.376** (1.063-1.780)	1.322** (1.036-1.687)	1.581** (1.188-2.104)	0.696** (0.536-0.903)
[ai] Monophthongization (n=2924)	1.250* (0.935-1.671)	1.214* (0.930-1.585)	1.369** (1.020-1.837)	0.546** (0.407-0.732)
Use of Any AAE Feature	0.851** (0.766-0.946)	0.723** (0.657-0.795)	1.535** (1.382-1.705)	1.910** (1.731-2.108)

[1] All models controlled for interviewer education and age.
** = Statistically significant at $p \leq .05$
* = Borderline significant at $p \leq .15$

<u>Table 4.5</u>. Number of Tokens of African American English (AAE) Variants of Seven Phonological Features Occurring in Recruitment, Racially Topical, and Sensitive Survey Sections (Percent of Total Tokens)

AAE Feature (n=Total Number of Tokens)	Recruitment Sections	Racially Topical Sections	Sensitive Sections
"Ask" Metathesis (n=646)	202 (32.2%)	165 (25.5%)	86 (13.3%)
Syllable-Final [d] (n=2246)	548 (24.4%)	510 (22.7%)	212 (9.4%)
"-ing" Endings (n=4073)	687 (16.9%)	896 (22.0%)	238 (5.8%)
Pin/Pen Merger (n=2452)	411 (16.8%)	1135 (46.3%)	65 (2.7%)
Initial "th" (n=2304)	408 (17.7%)	1184 (51.4%)	81 (3.5%)
[ai] Monophthongization (n=2924)	457 (15.6%)	721 (24.7%)	397 (13.6%)

Table 4.6. Changes in the Proportion of African American Interviewers' (n=14) Use of African American English (AAE) Phonological Features During Recruitment, Racially Topical, and Sensitive Sections of a Survey Script Using Generalized Estimating Equations (Parameter Estimates and Standard Errors) [1]

AAE Feature (n=Total Number of Tokens)	Recruitment Sections	Racially Topical Sections	Sensitive Sections
"Ask" Metathesis (n=646)	-0.3635* (0.2201)	0.0387 (0.2149)	0.4395** (0.1990)
Syllable-Final [d] (n=2246)	-0.4171** (0.1379)	-0.6847** (0.2086)	0.6028** (0.1762)
"-ing" Endings (n=4073)	-0.2862** (0.1216)	-0.0625 (0.0969)	-0.0117 (0.1745)
Pin/Pen Merger (n=2452)	-0.2895* (0.1605)	0.0791 (0.1125)	1.1010** (0.2411)
Initial "th" (n=2304)	0.1864 (0.1365)	0.3844** (0.1573)	-0.5413 (0.3608)
[ai] Monophthongization (n=2924)	-0.3747** (0.1258)	-0.8640** (0.2277)	-0.2216 (0.1562)
Use of Any AAE Feature	-0.0800 (0.0766)	0.0024 (0.0756)	-0.1057 (0.0863)

[1] All models controlled for the standardized respondent involved in each interview, interviewer education, interviewer age, and the clustering of linguistic data by interviewers.
** Statistically significant at $p \leq .05$.
* Borderline significant at $p \leq .10$

<u>Table 4.7</u>. Results of Linear Mixed Models Testing Associations Between Interviewers' (n=8) Black American Ethnic Identity Status and Use of African American English Phonological Features and Ethnic Identity Subscale Means for *Eat for Life* Participants (n=492) (Estimates and Standard Errors) [1] [2]

	Afrocentric Subscale Mean	Black American Subscale Mean	Bicultural Subscale Mean	Multicultural Subscale Mean	Racial Salience Subscale Mean
"Ask" Metathesis	0.0014 (0.0044)	0.0001 (0.0022)	-0.0012 (0.0025)	**-0.0007 (0.0021)**	-0.0030 (0.0030)
Syllable-Final [d]	-0.0064 (0.0283)	0.0119 (0.0156)	-0.0059 (0.0167)	**-0.0232 (0.0147)**	**-0.0263 (0.0192)**
"-ing" Endings	-0.0061 (0.0053)	0.0003 (0.0031)	0.0007 (0.0033)	**-0.0041 (0.0031)**	-0.0024 (0.0045)
Consonant Cluster Reduction	0.0145 (0.0059)	**0.0049 (0.0041)**	-0.0065 (0.0047)	**0.0014 (0.0044)**	**0.0127 (0.0057)**
Pin/Pen Merger	0.0060 (0.0077)	0.0009 (0.0041)	0.0002 (0.0045)	**0.0040 (0.0039)**	0.0003 (0.0061)
Initial "th"	-0.0235 (0.0219)	-0.0142 (0.0114)	0.0028 (0.0143)	**-0.0037 (0.0117)**	**-0.0282 (0.0153)**
[ai] Monophthongization	0.0189 (0.0098)	**0.00694 (0.0058)**	-0.0042 (0.0076)	**0.0049 (0.0063)**	**0.0173 (0.0081)**
AAE Index	-0.0090 (0.0205)	0.0052 (0.0105)	-0.0018 (0.0119)	**-0.0087 (0.0103)**	-0.0061 (0.0157)
Black American Ethnic Identity Component	0.2822 (0.1627)	0.1080 (0.0942)	**-0.1910 (0.0814)**	**-0.1058 (0.0964)**	0.0609 (0.1491)

[1] Numbers in bold type indicate models tested using ordinary least squares (OLS) regression. All OLS regression models controlled for *Eat for Life* participant gender, age, education, and income. The unbolded numbers indicate linear mixed models. All linear mixed models controlled for *Eat for Life* participant gender, age, education, and income, as well as the clustering of participant data by interviewer.
[2] No models were statistically significant at $p \leq .01$.

CHAPTER 5

CONCLUSION

The purpose of this dissertation was to explore interviewer-associated measurement error in surveys between African American interviewers and African American respondents. This dissertation had four primary objectives. The first objective was to review the literature on interviewer effects (Chapter 2). The second objective was to investigate whether African American respondents prefer to be interviewed by African American interviewers and whether this preference is related to respondents' ethnic identity orientations (Chapter 3). The third objective was to explore how racial and ethnic identity are conveyed by African American interviewers during telephone survey administrations by using linguistic analyses of interviewers' use of African American English (AAE) (Chapter 4). The fourth objective was to inform and encourage future research. Few, if any, empirical investigations have been published to date on interviewer effects associated with within-group cultural variability, and more research is needed on the presence of interviewer effects in public health survey data.

The Chapter 2 review of interviewer effects research indicates that much remains to be learned about the impact of interviewers' sociodemographic characteristics on health-related survey data. Interviewer effects appear to be most likely to occur when survey items query attitudes about sociodemographic characteristics or respondents' engagement in sensitive behaviors such as substance use. Interviewer effects can have a substantial impact on survey data and may be particularly operant in public health surveys, where respondents are likely to be queried about racial attitudes, sensitive behaviors, and other topics prone to socially desirable responding. Additional research is needed to elucidate many issues, including the influence of interviewers' sociodemographic characteristics on health-related topics, the role of within-group interviewer variability on survey data, and the simultaneous impact of multiple interviewer characteristics. The findings of such research would provide much-needed guidance to public health professionals on whether or not to match interviewers and respondents on key sociodemographic characteristics.

As a first step in initiating the research suggested above, Chapter 3 presents data from a study on African American telephone survey respondents' preferences for interviewer race. The findings from this study imply that African American respondents vary in their preferences for having an African American interviewer. Respondents with Afrocentric, Black American, or Cultural Mistrust ethnic identity components appear to prefer an African American interviewer for a survey with racial content. In contrast, interviewer race may be unimportant to respondents with Assimilated, Bicultural, or Multicultural ethnic identity components. A respondent's preference for an African American interviewer may be positively correlated with how important being Black is to his or her personal identity. However, no association is apparent between respondent ethnic identity type and reporting of hypothesized comfort with a White interviewer. Respondents with higher educational attainment appear to be more likely to report being less comfortable with a White interviewer. These findings may indicate a measurement effect, with respondents of higher educational status feeling less social desirability to express putative multicultural attitudes. The type of survey content appears to impact African American respondents' feelings about interviewer race. Respondents may prefer to be surveyed by African American interviewers and be less comfortable with White interviewers for surveys containing racial attitude questions. However, further research is required to assess whether or not the fulfillment of respondents' preferences has an impact on measurement error.

Chapter 4 explored African American interviewers' use of AAE phonological features as one means of conveying in-group social identity with African American telephone survey respondents. Study findings indicate that AAE use in telephone surveys varies by language feature, as well as across and within speakers. African American interviewers with a Bicultural or Multicultural ethnic identity component may be more likely to use select AAE phonological features as cues to convey in-group membership to telephone respondents. In contrast, interviewers with a Black American identity component may feel less compelled to convey their in-group membership through language use. Interviewers appeared to use AAE phonological features less when engaged in recruitment tasks and more when administering potentially sensitive survey items. In contrast, no clear role is evident for their use of AAE phonology during racially

topical survey sections. This lack of pattern for racial items may result from the highly politicized nature of language use among African Americans. These varied patterns of expression of AAE during survey interviews may reflect interviewers' attempts to fulfill their sometimes contradictory tasks of building rapport with and conveying objectivity to survey respondents. Interviewers' use of AAE phonological features, however, had no impact on telephone survey respondents' answers to racially topical or nonracial but sensitive survey items. This lack of effects may stem from the overall low use of AAE phonology by the interviewers in the study, who were all from Washington State. The low usage of AAE is reflective of the preferential use of American Standard English in standardized survey script. Whether or not surveys should be administered in AAE or other dialects may have consequences for the survey interaction and remains an important issue to be explored.

This dissertation was a first attempt to investigate several previously unexplored issues. Findings from this and future research will inform the conduct of surveys of African Americans, as well as public health programming with other racially or culturally defined groups. The elimination of race-based health disparities is a daunting task, and it may well require that public health professionals reconfigure data collection paradigms in manners beyond our current imagination.

APPENDIX A:

THE BLACK IDENTITY CLASSIFICATION SCALE (BICS)

Many of the items in the BICS have been adapted from other instruments. For a full description of the development of the BICS and the origins of the BICS items, please see (R. E. Davis et al., under review). All items are scored from one to seven, with one representing "Strongly Disagree" and seven representing "Strongly Agree".

Afrocentric Subscale

1. It is important to me to celebrate Kwanzaa.
2. I feel a strong emotional connection to Africa.
3. Black people should give their children African names.
4. I am involved in Black political activities.
5. It is important to learn about African culture.
6. I believe that it is important for African Americans to learn about spiritual beliefs in Africa.
7. It is important for African Americans to get back to their African roots.

Black American Subscale

8. When I watch television, I usually watch Black television shows.
9. When I read magazines, I read mostly Black magazines such as Jet and Ebony.
10. It is important for Black people to educate their children about Black art, music, and literature.
11. A thorough knowledge of Black history is very important for Blacks today.
12. I have a strong sense of belonging to the Black community.
13. It is important to be involved in the Black community.
14. Most of my friends are Black.
15. When I listen to the radio, I usually listen to Black radio shows.

Bicultural Subscale

16. I am proud of my ability to succeed in both the Black and White worlds.
17. I feel at ease with Whites and Blacks.
18. I feel comfortable in both worlds.
19. I feel comfortable interacting with both Blacks and Whites.

Multicultural Subscale

20. I feel strongly about American social issues such as women's rights, the environment, and animal rights.
21. I feel strongly about international human rights issues in places such as the Middle East and Tibet.
22. I care deeply about the needs of other groups such as Native Americans, Whites, Latinos, and Asian Americans.
23. I respect the cultural traditions of many groups – for example, Native Americans, Whites, Latinos, and Asian Americans.

Racial Salience Subscale

24. Being Black has a lot to do with how I feel about myself.
25. I think of myself as African American more than American.
26. Being Black is an important part of my self-image.
27. Many things that are important to me are connected to my Black identity.
28. Both in my public and private thoughts, race is an important part of who I am.
29. Many things that make me happy are connected to the fact that I am Black.

Cultural Mistrust Subscale

30. When I think about race relations in America, I get angry.
31. Many White politicians deliberately pass laws designed to block the progress of Blacks.
32. The United States government is trying to make things better for Blacks.

Racial Salience Single Item

33. For this next question, I'd like you to think of a number between zero and ten with zero meaning "Not at All Important" and ten meaning "Very Important". By choosing any number between zero and ten, how important is being Black to your overall identity?

Short-Cut Ethnic Identity Items (Used as Tie-Breakers)

34. As you know, our program involves creating personalized newsletters about health and diet. For your newsletter, which types of newsletters would you like to receive? Please answer "yes" or "no" to each of the types I describe. Would you like to receive a newsletter that focuses on the food and culture of: (all response options are yes/no)

    34a. Black people in America

    34b. African Americans and their connections to Africa

    34c. Both Black and White Americans

    34d. People of many racial and cultural backgrounds, including those such as Latinos and Asian Americans

    34e. Americans, without any references to Black culture

35. This question was only asked if a respondent answered "yes" to more than one of Questions 34a-34e: You've indicated that you would like to receive a newsletter that focuses upon the food and culture of (Black people in America / African Americans and their connections to Africa / Both Black and White Americans / people of many racial and cultural backgrounds, including those such as Latinos and Asian Americans / Americans, without any references to Black culture). Of these categories, which type of newsletter would you <u>most</u> prefer to receive?

BIBLIOGRAPHY

Achter, P., Parrott, R., & Silk, K. (2005). African Americans' opinions about human-genetics research. *Politics and the Life Sciences, 23*(1), 60-66.

American Association for Public Opinion Research, A. (2003, May). AAPOR Outcome Rate Calculator. Retrieved June 18, 2008, from http://www.aapor.org/uploads/Response_Rate_Calculator.xls.

Anderson, B. A., Silver, B. D., & Abramson, P. R. (1988a). The effects of race of the interviewer on measures of electoral participation by Blacks in SRC National Election Studies. *Public Opinion Quarterly, 52*(1), 53-83.

Anderson, B. A., Silver, B. D., & Abramson, P. R. (1988b). The effects of the race of the interviewer on race-related attitudes of Black respondents in SRC/CPS National Election Studies. *Public Opinion Quarterly, 52*(3), 289-324.

Aspinall, P. J. (2001). Operationalising the collection of ethnicity data in studies of the sociology of health and illness. *Sociology of Health & Illness, 23*(6), 829-862.

Atkinson, D. R., Furlong, M. J., & Poston, W. C. (1986). Afro-American preferences for counselor characteristics. *Journal of Counseling Psychology, 33*(3), 326-330.

Audacity. (2006). Audacity (Version 1.2.6) [Computer software]. Retrieved from http://audacity.sourceforge.net/.

Axinn, W. G. (1991). The influence of interviewer sex on responses to sensitive questions in Nepal. *Social Science Research, 20*(3), 303-318.

Bailey, G., & Thomas, E. (1998). Some aspects of African-American vernacular English phonology. In S. S. Mufwene, J. R. Rickford, G. Bailey & J. Baugh (Eds.), *African-American English: Structure, history, and use* (pp. 85-109). New York: Routledge.

Baugh, J. (1999). *Out of the mouths of slaves: African American language and educational malpractice*. Austin, TX: University of Texas Press.

Baugh, J. (2000). *Beyond Ebonics: Linguistic pride and racial prejudice*. New York: Oxford University Press.

Becker, S., Feyisetan, K., & Makinwa-Adebusoye, P. (1995). The effect of the sex of interviewers on the quality of data in a Nigerian family planning questionnaire. *Studies in Family Planning, 26*(4), 233-240.

Benney, M., & Geiss, P. (1950). Social class and politics in Greenwich. *The British Journal of Sociology, 1*(4), 310-327.

Benney, M., Riesman, D., & Star, S. A. (1956). Age and sex in the interview. *The American Journal of Sociology, 62*(2), 143-152.

Boersma, P., & Weenink, D. (1992-2003). Praat: doing phonetics by computer (Version 4.1.11) [Computer software]. Retrieved in 2003 from http://www.praat.org/.

Brenner, M. (1982). Response-effects of 'role restricted' characteristics of the interviewer. In W. Dijkstra & J. van der Zouwen (Eds.), *Response behaviour in the survey-interview* (pp. 91-130). New York: Academic Press.

Brown, D. W. (2006). Girls and guys, ghetto and bougie: Metapragmatics, ideology and the management of social identities. *Journal of Sociolinguistics, 10*(5), 596-610.

Campbell, B. A. (1981). Race-of-interviewer effects among Southern adolescents. *Public Opinion Quarterly, 45*(2), 231-244.

Carpenter, J. S., Andrykowski, M. A., Cordova, M. J., Cunningham, L. L. C., & Studts, J. L. (1999). Do participants' reports of symptom prevalence or severity vary by interviewer gender? *Nursing Research, 48*(5), 276-279.

Carr, L. G. (1971). The Srole items and acquiescence. *American Journal of Sociology, 36*(2), 287-293.

Catania, J. A., Binson, D., Canchola, J., Pollack, L. M., Hauck, W., & Coates, T. J. (1996). Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behavior. *Public Opinion Quarterly, 60*(3), 345-375.

Cleary, P. D., Mechanic, D., & Weiss, N. (1981). The effect of interviewer characteristics on responses to a mental health interview. *Journal of Health and Social Behavior, 22*(2), 183-193.

Cokley, K. (2007). Critical issues in the measurement of ethnic and racial identity: A referendum on the state of the field *Journal of Counseling Psychology, 54*(3), 224-234.

Colombotos, J., Elinson, J., & Loewenstein, R. (1969). Effect of interviewers' sex on interview responses. *The Milbank Memorial Fund Quarterly, 47*(1, Part 2), 227-231.

Corbie-Smith, G., Thomas, S. B., Williams, M. V., & Moody-Ayers, S. (1999). Attitudes and beliefs of African Americans toward participation in medical research. *Journal of General Internal Medicine, 14*(9), 537-546.

Cotter, P. R., Cohen, J., & Coulter, P. B. (1982). Race-of-interviewer effects in telephone interviews. *Public Opinion Quarterly, 46*(2), 278-284.

Cross, W. E., Jr. (1991). *Shades of Black: Diversity in African-American identity*. Philadelphia: Temple University Press.

Dailey, R. M., & Claus, R. E. (2001). The relationship between interviewer characteristics and physical and sexual abuse disclosures among substance users: A multilevel analysis. *Journal of Drug Issues, 31*(4), 867-888.

Davis, D. W. (1997a). The direction of race of interviewer effects among African-Americans: Donning the black mask. *American Journal of Political Science, 41*(1), 309-322.

Davis, D. W. (1997b). Nonrandom measurement error and race of interviewer effects among African Americans. *Public Opinion Quarterly, 61*(1), 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*(1), 33-45.

Davis, R. E., Alexander, G., Calvi, J., Wiese, C., Greene, S., Nowak, M., et al. (under review). A new audience segmentation tool for African Americans: The Black Identity Classification Scale.

DeLamater, J. D. (1974). Methodological issues in the study of premarital sexuality. *Sociological Methods & Research, 3*(1), 30-61.

Dijkstra, W. (1983). How interviewer variance can bias the results of research on interviewer effects. *Quality & Quantity, 17*, 179-187.

Dillard, J. L. (1972). *Black English: Its history and usage in the United States*. New York: Random House, Inc.

Dohrenwend, B. S., Colombotos, J., & Dohrenwend, B. P. (1968). Social distance and interviewer effects. *Public Opinion Quarterly, 32*(3), 410-422.

Dohrenwend, B. S., Williams, J. A., Jr., & Weiss, C. H. (1969). Interviewer biasing effects: Toward a reconciliation of findings. *Public Opinion Quarterly, 33*(1), 121-129.

Doss, R. C., & Gross, A. M. (1994). The effects of Black English and code-switching on intraracial perceptions. *Journal of Black Psychology, 20*(3), 282-293.

Edwards, W. F. (1992). Sociolinguistic behavior in a Detroit inner-city Black neighborhood. *Language in Society, 21*(1), 93-115.

Ehrlich, J. S., & Riesman, D. (1961). Age and authority in the interview. *Public Opinion Quarterly, 25*(1), 39-56.

Feldman, J. J., Hyman, H., & Hart, C. W. (1951-1952). A field study of interviewer effects on the quality of survey data. *Public Opinion Quarterly, 15*(4), 734-761.

Fendrich, M., Johnson, T., Shaligram, C., & Wislar, J. S. (1999). The impact of interviewer characteristics on drug use reporting by male juvenile arrestees. *Journal of Drug Issues, 29*(1), 37-58.

Filmer, A. A. (2003). African-American Vernacular English: Ethics, ideology, and pedagogy in the conflict between identity and power. *World Englishes, 22*(3), 253-270.

Finkel, S. E., Guterbock, T. M., & Borg, M. J. (1991). Race-of-interviewer effects in a preelection poll: Virginia 1989. *Public Opinion Quarterly, 55*(3), 313-330.

Ford, K., & Norris, A. E. (1997). Effects of interviewer age on reporting of sexual and reproductive behavior of Hispanic and African American youth. *Hispanic Journal of Behavioral Sciences, 19*(3), 369-376.

Freeman, J., & Butler, E. W. (1976). Some sources of interviewer variance in surveys. *Public Opinion Quarterly, 40*(1), 79-91.

Freitag, C. B., & Barry, J. R. (1974). Interaction and interviewer bias in a survey of the aged. *Psychological Reports, 34*(3), 771-774.

Fry, R. P. W., Rozewicz, L. M., & Crisp, A. H. (1996). Interviewing for sexual abuse: Reliability and effect of interviewer gender. *Child Abuse & Neglect, 20*(8), 725-729.

Gamble, V. N. (1997). Under the shadow of Tuskegee: African Americans and health care. *American Journal of Public Health, 87*(11), 1773-1778.

Green, L. J. (2002). *African American English: A linguistic introduction*. New York: Cambridge University Press.

Grimes, M. D., & Hansen, G. L. (1984). Response bias in sex-role attitude measurement. *Sex Roles, 10*(1/2), 67-72.

Groves, R. M. (2004). *Survey errors and survey costs*. Hoboken, NJ: John Wiley & Sons, Inc.

Groves, R. M., & Fultz, N. H. (1985). Gender effects among telephone interviewers in a survey of economic attitudes. *Sociological Methods & Research, 14*(1), 31-52.

Groves, R. M., & Magilavy, L. J. (1980). *Estimates of interviewer variance in telephone surveys.* Paper presented at the The Section on Survey Research Methods of the American Statistical Association.

Groves, R. M., & Magilavy, L. J. (1986). Measuring and explaining interviewer effects in centralized telephone surveys. *Public Opinion Quarterly, 50*(2), 251-266.

Gumperz, J. J., & Cook-Gumperz, J. (1982). Introduction: Language and the communication of social identity. In J. J. Gumperz (Ed.), *Language and social identity* (pp. 1-21). New York: Cambridge University Press.

Hagenaars, J. A., & Heinen, T. G. (1982). Effects of role-independent interviewer characteristics on responses. In W. Dijkstra & J. van der Zouwen (Eds.), *Response behaviour in the survey-interview* (pp. 91-130). New York: Academic Press.

Hansell, M., & Ajirotutu, C. S. (1982). Negotiating interpretations in interethnic settings. In J. J. Gumperz (Ed.), *Language and social identity* (pp. 85-94). New York: Cambridge University Press.

Hanson, R. H., & Marks, E. S. (1958). Influence of the interviewer on the accuracy of survey results. *Journal of the American Statistical Association, 53*(283), 635-655.

Hatchett, S., & Schuman, H. (1975-1976). White respondents and race-of-interviewer effects. *Public Opinion Quarterly, 39*(4), 523-528.

Heeb, J. L., & Gmel, G. (2001). Interviewers' and respondents' effects on self-reported alcohol consumption in a Swiss health survey. *Journal of Studies on Alcohol, 62*(4), 434-442.

Hox, J. J., de Leeuw, E. D., & Kreft, I. G. G. (1991). The effect of interviewer and respondent characteristics on the quality of survey data: A multilevel model. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 439-461). New York: John Wiley & Sons, Inc.

Huddy, L., Billig, J., Bracciodieta, J., Hoeffler, L., Moynihan, P. J., & Pugliani, P. (1997). The effect of interviewer gender on the survey response. *Political Behavior, 19*(3), 197-220.

Hutchinson, K. L., & Wegge, D. G. (1991). The effects of interviewer gender upon response in telephone survey-research. *Journal of Social Behavior and Personality, 6*(3), 573-584.

Hyman, H. H., Cobb, W. J., Feldman, J. J., Hart, C. W., & Stember, C. H. (1954/1975). *Interviewing in social research*. Chicago: The University of Chicago Press.

Johnson, T. P., & Parsons, J. A. (1994). Interviewer effects on self-reported substance use among homeless persons. *Addictive Behaviors, 19*(1), 83-93.

Johnson, W. T., & Delamater, J. D. (1976). Response effects in sex surveys. *Public Opinion Quarterly, 40*(2), 165-181.

Kane, E. W., & Macaulay, L. J. (1993). Interviewer gender and gender attitudes. *Public Opinion Quarterly, 57*(1), 1-28.

Katz, D. (1942). Do interviewers bias poll results? *Public Opinion Quarterly, 6*(2), 248-268.

Kish, L. (1962). Studies of interviewer variance for attitudinal variables. *Journal of the American Statistical Association, 57*(297), 92-115.

Koch, L. M., Gross, A. M., & Kolts, R. (2001). Attitudes toward Black English and code switching. *Journal of Black Psychology, 27*(1), 29-42.

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*(4), 364-383.

Labov, W. (1972). *Language in the inner city: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Landis, J. R., Sullivan, D., & Sheley, J. (1973). Feminist attitudes as related to sex of the interviewer. *Pacific Sociological Review, 16*(3), 305-314.

Lenski, G. E., & Leggett, J. C. (1960). Caste, class, and deference in the research interview. *The American Journal of Sociology, 65*(5), 463-467.

Linnes, K. (1998). Middle-class AAVE versus middle-class bilingualism: Contrasting speech communities. *American Speech, 73*(4), 339-367.

Livert, D., Kadushin, C., Schulman, M., & Weiss, A. (1998, May 14-17). *Do interviewer-respondent race effects impact the measurement of illicit substance use and related attitudes?* Paper presented at the 53rd Annual Conference of the American Association for Public Opinion Research, St. Louis, MO.

Lueptow, L. B., Moser, S. L., & Pendleton, B. F. (1990). Gender and response effects in telephone interviews about gender characteristics. *Sex Roles, 22*(1/2), 29-42.

Malat, J. R., van Ryn, M., & Purcell, D. (2006). Race, socioeconomic status, and the perceived importance of positive self-presentation in health care. *Social Science & Medicine, 62*(10), 2479-2488.

Martin, S., & Wolfram, W. (1998). The sentence in African-American vernacular English. In S. S. Mufwene, J. R. Rickford, G. Bailey & J. Baugh (Eds.), *African-American English: Structure, history, and use* (pp. 11-36). New York: Routledge.

McGlone, M. S., Aronson, J., & Kobrynowicz, D. (2006). Stereotype threat and the gender gap in political knowledge. *Psychology of Women Quarterly, 30*(4), 392-398.

Morgan, M. (1994). Theories and politics in African American English. *Annual Review of Anthropology, 23*, 325-345.

Morgan, M. (2002). *Language, discourse, and power in African American culture*. New York: Cambridge University Press.

Morten, G., & Atkinson, D. R. (1983). Minority identity development and preference for counselor race. *Journal of Negro Education, 52*(2), 156-161.

Mufwene, S. S. (2001). What is African American English? In S. L. Lanehart (Ed.), *Sociocultural and historical contexts of African American English* (Vol. 27, pp. 21-51). Philadelphia: John Benjamins Publishing Company.

Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly, 60*(1), 58-88.

Nguyen, J. G. (2006). *The changing social and linguistic orientation of the African American middle class.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Ogbu, J. U. (1999). Beyond language: Ebonics, proper English, and identity in a Black-American speech community. *American Educational Research Journal, 36*(2), 147-184.

Olson, K., & Peytchev, A. (2007). Effect of interviewer experience on interview pace and interviewer attitudes. *Public Opinion Quarterly, 71*(2), 273-286.

Parham, T. A., & Helms, J. E. (1981). The influence of Black students' racial identity attitudes on preferences for counselor's race. *Journal of Counseling Psychology, 28*(3), 250-257.

Pollner, M. (1998). The effects of interviewer gender in mental health interviews. *The Journal of Nervous and Mental Disease, 186*(6), 369-373.

Rahman, J. (2007). An ay for an ah: Language of survival in African American narrative comedy. *American Speech, 82*(1), 65-96.

Reese, S. D., Danielson, W. A., Shoemaker, P. J., Chang, T. K., & Hsu, H. L. (1986). Ethnicity-of-interviewer effects among Mexican-Americans and Anglos. *Public Opinion Quarterly, 50*(4), 563-572.

Resnicow, K., Davis, R. E., Zhang, G., Konkel, J., Strecher, V. J., Shaikh, A. R., et al. (2008). Tailoring a fruit and vegetable intervention on novel motivational constructs: Results of a randomized study. *Annals of Behavioral Medicine, 35*(2), 159-169.

Resnicow, K., Davis, R. E., Zhang, G., Saunders, E., Strecher, V. J., Tolsma, D., et al. (under review). Tailoring a fruit and vegetable intervention on ethnic identity: Results of a randomized study.

Resnicow, K., Soler, R. E., Braithwaite, R. L., Selassie, M. B., & Smith, M. (1999). Development of a racial and ethnic identity scale for African American adolescents: The Survey of Black Life. *Journal of Black Psychology, 25*(2), 171-188.

Rhodes, P. J. (1994). Race-of-interviewer effects: A brief comment. *Sociology 28*(2), 547-558.

Rickford, J. R. (1999). *African American Vernacular English: Features, evolution, educational implications*. Malden, MA: Blackwell Publishers Inc.

Rickford, J. R., & Rickford, R. J. (2000). *Spoken Soul: The story of Black English*. New York: John Wiley & Sons, Inc.

SAS Institute, I. (2002-2003). (Version 9.1.3). Cary, NC: SAS Institute, Inc.

Schaeffer, N. C. (1980). Evaluating race-of-interviewer effects in a national survey. *Sociological Methods & Research, 8*(4), 400-419.

Schuman, H., & Converse, J. M. (1971). The effects of Black and White interviewers on Black responses in 1968. *Public Opinion Quarterly, 35*(1), 44-68.

Sellers, R. M., Rowley, S. A. J., Chavous, T. M., Shelton, J. N., & Smith, M. A. (1997). Multidimensional Inventory of Black Identity: A preliminary investigation of reliability and construct validity. *Journal of Personality & Social Psychology, 73*(4), 805-815.

Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A. J., & Chavous, T. M. (1998). Multidimensional Model of Racial Identity: A reconceptualization of African American racial identity. *Personality and Social Psychology Review 2*(1), 18-39.

Singer, E., Frankel, M. R., & Glassman, M. B. (1983). The effect of interviewer characteristics and expectations on response. *Public Opinion Quarterly, 47*(1), 68-83.

Smitherman, G. (2000). *Talkin that talk: Language, culture, and education in African America*. New York: Routledge.

Spears, A. K. (2001). Directness in the use of African American English. In S. L. Lanehart (Ed.), *Sociocultural and historical contexts of African American English* (Vol. 27, pp. 239-259). Philadelphia: John Benjamins Publishing Company.

Speicher, B. L., & McMahon, S. M. (1992). Some African-American perspectives on Black English Vernacular. *Language in Society, 21*(3), 383-407.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629.

Stember, H., & Hyman, H. (1949-1950). Interviewer effects in the classification of responses. *Public Opinion Quarterly, 13*(4), 669-682.

Stock, J. S., & Hochstim, J. R. (1951). A method of measuring interviewer variability. *Public Opinion Quarterly, 15*(2), 322-334.

Summers, G. F., & Hammonds, A. D. (1966). Effect of racial characteristics of investigator on self-enumerated responses to a Negro prejudice scale. *Social Forces, 44*(4), 515-518.

Terrell, F., & Terrell, S. (1981). An inventory to measure cultural mistrust among Blacks. *The Western Journal of Black Studies, 5*(3), 180-185.

Tourangeau, R., Couper, M. P., & Steiger, D. M. (2003). Humanizing self-administered surveys: Experiments on social presence in Web and IVR surveys. *Computers in Human Behavior, 19*(1), 1-24.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. London: Cambridge University Press.

Valentin-Marquez, W., & Nguyen, J. G. (2002, October). *Am I my brother's keeper?: The role of family ties and standardizing factors in the phonology of African American English*. Paper presented at the Paper presented at the New Ways of Analyzing Variation 31 Pre-Conference on African American Vernacular English, Palo Alto, CA.

Valentin-Marquez, W., & Nguyen, J. G. (2004, April). *Beyond demographics: A case study of intra-family variation in African American English*. Paper presented at the Paper presented at the Sociolinguistics Symposium 15, Newcastle, Australia.

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*(5), 334-342.

Webster, C. (1996). Hispanic and Anglo interviewer and respondent ethnicity and gender: The impact on survey response quality. *Journal of Marketing Research, 33*(1), 62-72.

Weeks, M. F., & Moore, R. P. (1981). Ethnicity-of-interviewer effects on ethnic respondents. *Public Opinion Quarterly, 45*(2), 245-249.

Weinreb, A. A. (2006). The limitations of stranger-interviewers in rural Kenya. *American Sociological Review, 71*(6), 1014-1039.

Welch, S., Comer, J., & Steinman, M. (1973). Interviewing in a Mexican-American community: An investigation of some potential sources of response bias. *Public Opinion Quarterly, 37*(1), 115-126.

West, B. T., Welch, K. B., & Gałecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. New York: Chapman & Hall.

Williams, J. A., Jr. (1964). Interviewer-respondent interaction: A study of bias in the information interview. *Sociometry, 27*(3), 338-352.

Williams, J. A., Jr. (1968). Interviewer role performance: A further note on bias in the information interview. *Public Opinion Quarterly, 32*(2), 287-294.

Williams, R. (1975). *Ebonics: The true language of Black folks*. St. Louis, MO: The Institute of Black Studies.

Wilson, D. C. (2007). Conversations with a stranger: Race of interviewer perceptibility. Unpublished Unpublished manuscript.

Wilson, S. R., Brown, N. L., Mejia, C., & Lavori, P. W. (2002). Effects of interviewer characteristics on reported sexual behavior of California Latino couples. *Hispanic Journal of Behavioral Sciences, 24*(1), 38-62.

Wolford, M. L., Brown, R. E., Marsden, A., Jackson, J. S., & Harrison, C. (1995, August 13-17). *Bias in telephone surveys of African Americans: The impact of perceived race of interviewer on responses.* Paper presented at the Annual Meeting of the American Statistical Association, Orlando, FL.

Wolfram, W. (1969). *A sociolinguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.

Wolfram, W. A., & Fasold, R. W. (1969). Towards reading materials for speakers of Black English: Three linguistically appropriate passages. In J. C. Baratz & R. Shuy (Eds.), *Teaching Black children to read* (pp. 138-155). Washington, DC: Center for Applied Linguistics.

Zimmerman, M. A., Caldwell, C. H., & Bernat, D. H. (2002). Discrepancy between self-report and school-record grade point average: Correlates with psychosocial outcomes among African American adolescents. *Journal of Applied Social Psychology, 32*(1), 86-109.