A Mixed-Methods Approach to Developing a Self-Reported Racial/Ethnic Discrimination Measure for Use in Multiethnic Health Surveys

Objective: The development of measures of self-reported racial/ethnic discrimination is an active area of research, but few measures have been validated across multiple racial/ethnic and language groups. Our goal is to develop and evaluate a discrimination measure that is appropriate for use in surveys of racially and ethnically diverse populations.

Methods: To develop our measure, we employ a mixed-methods approach for survey research, drawing from both qualitative and quantitative traditions, including literature review, cognitive testing, psychometric analyses, behavior coding as well as two rounds of field testing using a split-sample design. We tested our new measure using two different approaches to elicit self-reported experiences of racial/ethnic discrimination.

Results: Our new measure captures four dimensions of racial/ethnic discrimination: 1) frequency of encounters with discrimination across several domains (e.g., medical care, school, work, street and other public places); 2) timing of exposure (e.g., recent, lifetime); 3) appraisal of discrimination as stressful; and 4) responses to discrimination.

Conclusions: Because of the growing interest in measurement of racial/ethnic discrimination in health surveys, we think this report on the methods informing the development and testing of the discrimination module that will be used on the California Health Interview Survey would be useful to other researchers. The application of mixed methods to rigorously test the validity and reliability of our instrument proves to be a good roadmap for measuring racial/ethnic discrimination in multicultural and multilingual populations. (Ethn Dis. 2009;19:447–453)

Key Words: Cross-Cultural Comparison, Discrimination, Survey Methods

INTRODUCTION

Racial/ethnic discrimination is associated with negative health outcomes, ranging from depression to coronary calcification to mortality.1-4 Few measures of self-reported racial/ethnic discrimination have been validated across multiple racial/ethnic and language groups.5 Moreover, it is not clear that the experience of discrimination is qualitatively equivalent across different racial/ethnic and language groups. The growing multicultural demographic landscape of the United States and the documented relationship between discrimination and health outcomes highlight the need for new, brief measures that are valid across multiple populations and that can be used in health surveys. The National Cancer Institute (NCI) is collaborating with the UCLA Center for Health Policy Research to improve discrimination measures in the California Health Interview Survey (CHIS). In this article, we describe our mixed-methods design for evaluating two approaches to elicit self-reported experiences of racial/ethnic discrimination as part of telephone health surveys.

Mixed-Methods Approach

In 2006, an NCI-led workgroup, called the CHIS Discrimination Module Workgroup, began developing an instrument for measuring self-reported racial/ethnic discrimination using the CHIS. Our goal was to develop and evaluate a discrimination measure that is appropriate for use in surveys of racially and ethnically diverse populations. The CHIS is a statewide, random digit-dial telephone health survey administered in the nation’s most ethnically-diverse state. The CHIS question-
naire was developed in collaboration between the CHIS staff and academic researchers, policymakers, funders and several standing committees, including the CHIS Multicultural Technical Advisory Committee (TAC). These features, and the fact that the CHIS is fielded in the nation’s most populous and linguistically and ethnically diverse state, make the CHIS an ideal vehicle for testing measurement issues of racial/ethnic discrimination. (www.chis.ucla.edu)

We drew from both qualitative and quantitative traditions for our mixed-methods approach. The CHIS DM Workgroup received input from the TAC as well as from an external group of advisors at multiple steps in the process. Each methodological step in our process of evaluation was led by an internal expert, reviewed by the entire workgroup, and then reviewed again by the panel of external advisors. The multiple methods used are shown in Figure 1 and summarized below:

1) Structured review of existing racial/ethnic discrimination scales and studies from the literature;
2) Secondary analyses of existing racial/ethnic discrimination items fielded from the National Institute of Mental Health-funded National Latino and Asian American Study and the National Survey of American Life; these analyses focused on the psychometric properties of the Everyday Discrimination Scale including the validity and reliability of the individual items and scale in multiracial/ethnic samples;
3) Cognitive testing (with primary data) to identify cross-cultural equivalence of item wording;
4) Field-testing two alternative ways of asking about racial/ethnic discrimination in English using the CHIS 2007, and analyses comparing the two approaches;
5) Behavior coding of a subset of field-test interviews to identify strategies to improve survey administration of the measures;
6) Cultural review of the items by cultural/linguistic experts to further improve cross-cultural item and translation equivalence;
7) Multi-step, team-based translation (ie, team of translators working
Together instead of a single translator;  
8) Field-testing two alternative ways of asking about racial/ethnic discrimination in English, Spanish, Vietnamese, Korean, Cantonese and Mandarin using the CHIS 2009, and comparing the two approaches; and  
9) Feedback from members of the scientific community (eg, external advisors, and CHIS Multicultural TAC).

**LITERATURE REVIEW**

Three well-known, validated instruments that measure a variety of dimensions and domains of discrimination were identified in our literature review as particularly relevant to our project. These were the Everyday Discrimination Scale (EDS), Experiences of Discrimination Scale, and General Ethnic Discrimination Scale. The first two lend themselves to telephone administration in omnibus public health surveys such as the CHIS because they are short and have been administered in various health surveys in multicultural and multilingual settings. The third instrument was not designed for brief telephone administration. However, this instrument was important in informing the types of dimensions that could be measured, including the appraisal of discrimination as being stressful. Other instruments were reviewed but not used due to their length or limited validation across multiple racial/ethnic groups.

Based on the review, the workgroup agreed that the new measure should capture four dimensions: 1) frequency of encounters with discrimination across a number of domains (eg, medical care, school, work, street, other public places, police, and courts); 2) timing of exposure (eg, recent, lifetime); 3) appraisal of discrimination as stressful; and 4) responses to discrimination. These decisions were supported by our panel of external advisors and the CHIS Multicultural TAC.

**ANALYSIS OF EXISTING MEASURES**

Subsequent to our literature review we conducted secondary and primary data analyses to evaluate and refine discrimination items from the EDS. The EDS had been administered in many studies, including the National Latino and Asian American Study (NLAAS) and the National Survey of American Life (NSAL), to three racial/ethnic groups (Blacks, Latinos and Asian Americans). We conducted factor analyses on the EDS, which had been conceptualized as a unidimensional scale in previous research. Our factor analysis confirmed that the EDS can be conceptualized as a unidimensional scale. We also found that two items (“You are treated with less courtesy than other people” and “You are treated with less respect than other people”) were highly correlated. Qualitative results from cognitive testing supported our hypothesis that the term respect is more meaningful to respondents than courtesy. Accordingly, the courtesy item was dropped. Descriptive analyses showed that few people reported experiencing frequent discrimination (at least once a week) and that Black participants reported experiencing the most unfair acts of any racial/ethnic group we studied. We found that the item, “You are called names or insulted,” functioned differentially for Asian Americans and Blacks. Differential item functioning was also found for the item, “People act as if they are afraid of you,” between men and women.

The cognitive testing revealed that the majority of items performed as intended. Respondents did not raise any major or consistent concerns with the phrasing of these items. However, a few items were not interpreted as intended, or produced cognitive difficulties. “You received poorer service than other people,” was perceived as being too vague because respondents felt that everyone has received poor service at some point in time. Confirming other cross-cultural research, our cognitive testing showed that respondents preferred vague quantifiers (eg, once, often) to quantitative categories (eg, once, twice, three or more times, or once a year, once a month, once a week). We used behavioral coding to identify any unanticipated reactions during the course of interviews. Details are described in the next section.

**FIELD TESTS**

Quantitative and qualitative results informed instrument development for our field test. Our external advisors confirmed that the main objective of the field test should be to assess the two most common ways to ask about self-reported experiences of racial/ethnic discrimination. The first approach, early attribution, is to ask participants specifically about discrimination based on race/ethnicity (eg, Have you been unfairly treated because of your race/ethnicity?). The second, late attribution, is to first ask participants about unfair treatment, and then ask participants about the reasons for this unfair treatment (eg, Have you been unfairly treated? What was the reason?).

Two studies suggest that results may vary with the way questions on racial/ethnic discrimination are asked. The first study compared explicit versus generic measures of discrimination (parallel to the two approaches in our field test, early and late attribution, respectively) and found that prevalence, correlates, and associations with mental health vary depending on question framing. The second study, using data from the NLAAS, found that these two different approaches yielded discrimination scores that were only modestly correlated ($r = .43$). Among the participants who reported that they had experienced racial/ethnic discrimination...
<table>
<thead>
<tr>
<th>Table 1. Self-reported racial/ethnic discrimination field test instrument sections and sample items by early and late attribution approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Early Attribution—ask specifically about discrimination based on race/ethnicity</strong></td>
</tr>
<tr>
<td><strong>Section A. Racial/Ethnic Background</strong></td>
</tr>
<tr>
<td>A1. Do you think of yourself as [FILL from previous race/ethnicity items], or is there some other term that you think better describes you?</td>
</tr>
<tr>
<td>Never, Rarely, Sometimes, OR Often</td>
</tr>
<tr>
<td>B2. In the past 12 months, how often have you been treated unfairly or been discriminated against at restaurants or stores because you are [FILL]? Would you say…</td>
</tr>
<tr>
<td>Never, Rarely, Sometimes, OR Often</td>
</tr>
<tr>
<td>B5. In the past 12 months, how often have people acted as if they are afraid of you because you are [FILL]? Would you say…</td>
</tr>
<tr>
<td>Never, Rarely, Sometimes, OR Often</td>
</tr>
<tr>
<td>B9: Now I’m going to ask you why you may have been treated unfairly. Over the past five years, were you treated unfairly because of your…</td>
</tr>
<tr>
<td>Ancestry or national origin: [Yes/No]</td>
</tr>
<tr>
<td>Gender or sex: [Yes/No]</td>
</tr>
<tr>
<td>Race or skin color: [Yes/No]</td>
</tr>
<tr>
<td>Age: [Yes/No]</td>
</tr>
<tr>
<td>The way you speak English (language/accent): [Yes/No]</td>
</tr>
<tr>
<td>Some other reason: [Yes/No]; If Yes, specify_____________</td>
</tr>
<tr>
<td>IF yes to more than one: Which of these do you think is the main reason why you were treated unfairly…</td>
</tr>
<tr>
<td>IF yes to one or more items, B1–B8: How stressful have these experiences of unfair treatment usually been for you? Would you say…</td>
</tr>
<tr>
<td>Not at all stressful, A little stressful, Somewhat stressful, OR Extremely stressful</td>
</tr>
<tr>
<td><strong>Section C. Lifetime Experiences of Discrimination &amp; Appraisal of Discrimination as Stressful</strong></td>
</tr>
<tr>
<td>C1. Over your entire lifetime, how often have you been treated unfairly or been discriminated against at school because you are [FILL]? Would you say…</td>
</tr>
<tr>
<td>Never, Rarely, Sometimes, OR Often</td>
</tr>
<tr>
<td>C3. Over your entire lifetime, how often have you been treated unfairly or been discriminated against when getting medical care because you are [FILL]? Would you say…</td>
</tr>
<tr>
<td>Never, Rarely, Sometimes, OR Often</td>
</tr>
<tr>
<td>C6: Now I’m going to ask you why you may have been treated unfairly. Over your entire lifetime, were you treated unfairly because of your…</td>
</tr>
<tr>
<td>Ancestry or national origin: [Yes/No]</td>
</tr>
<tr>
<td>Gender or sex: [Yes/No]</td>
</tr>
<tr>
<td>Race or skin color: [Yes/No]</td>
</tr>
<tr>
<td>Age: [Yes/No]</td>
</tr>
<tr>
<td>The way you speak English (language/accent): [Yes/No]</td>
</tr>
<tr>
<td>Some other reason: [Yes/No]; If Yes, specify_____________</td>
</tr>
<tr>
<td>IF yes to more than one: Which of these do you think is the main reason why you were treated unfairly…</td>
</tr>
<tr>
<td>IF yes to one or more items, C1–C5: How stressful have these experiences of unfair treatment usually been for you? Would you say…</td>
</tr>
<tr>
<td>Not at all stressful, A little stressful, Somewhat stressful, OR extremely stressful</td>
</tr>
</tbody>
</table>
in response to the direct question (early attribution), nearly one-third stated that they had not experienced unfair racial/ethnic treatment in response to the two-stage question (late attribution). Conversely, among participants who, in response to the late-attribute question, stated that they had experienced unfair treatment, fully half stated, in response to the early-attribute question, that they had not experienced racial/ethnic discrimination. In the NLAAS example, the early attribution approach had 3 items, whereas the late attribution approach had 9 items and there was incomplete overlap in the content between the two approaches. For instance, early attribution included a question on unfair treatment of friends that was not asked in late attribution. Because the questions directly assessing racial/ethnic discrimination and those assessing unfair treatment were not identical, we do not know the contribution of variation in question wording to the observed differences. This is an important limitation of these studies.

Our study will directly address this limitation. To do this, we have developed one instrument with two versions (each with 8 items on recent experiences and 5 items on lifetime experiences). The two versions differ only in their approach for eliciting self-reported experiences of racial/ethnic discrimination. Respondents will be randomized into the early and late attribution versions of the module. Responses will be compared across racial/ethnic groups. Thus, findings from our field test will enable researchers to directly address this critical methodological question.

Our first field test, on CHIS 2007, was in English and yielded a multi-racial/ethnic sample who identified as: Latino (n=2056), Asian American and Pacific Islander (n=1357), African-American (n=1069), Native American (n=763), Non-Hispanic White (n=2037), or multiracial (n=223). Data was collected from October 2007 through February 2008. In addition to conducting psychometric analysis on the split-sample described above, the two instruments will be evaluated for validity, cultural comparability and administration properties using qualitative and quantitative techniques. One example is behavior coding. We used this technique to conduct a qualitative study designed to identify patterns related to question performance along lines of racial/ethnic and sex groups with a subsample of approximately 500 respondents (100 for each of the 5 racial/ethnic groups). Findings from this analysis will complement psychometric and other findings.

The two approaches will be tested in multiple languages on the CHIS 2009 and compared. For this second round of field testing, both versions of the instrument are undergoing English simplification, cultural review, and translation in order to field the two approaches in Spanish and some Asian languages. We will use these data to refine the two approaches for asking about racial/ethnic discrimination. We anticipate that our final module will be ready for administration in CHIS 2011 and will be released for public use.

**Next Steps**

We seek a concise, reliable and valid way to elicit reports of racial/ethnic discrimination experiences that will work equally well in all racial/ethnic populations. Although this module is intended for health-related population-based surveys, it may also be used to measure racial/ethnic discrimination in other types of studies. Our instrument is undergoing English simplification, cultural review, and translation in order to field the two approaches in Spanish and some Asian languages. We will use these data to refine the two approaches for asking about racial/ethnic discrimination. We anticipate that our final module will be ready for administration in CHIS 2011 and will be released for public use.

**Although this module is intended for health-related population-based surveys, it may also be used to measure racial/ethnic discrimination in other types of studies.**
includes dimensions that are theoretically important (eg, appraisal of discrimination as stressful) but are not routinely measured in instruments administered by telephone. A goal of our evaluation is to further balance content validity (eg, lifetime vs recent experiences, occurrences in various settings) with practical considerations related to parsimony, respondent burden, and costs of administration. Although the discrimination module is still under development, we feel that the growing interest in the measurement of racial/ethnic discrimination in health surveys warranted this early report. We hope that our preliminary results will contribute to development and refinement of measuring racial/ethnic discrimination in multicultural and multilingual populations. Our intention is to provide periodic updates to the scientific community as the data are analyzed and showcase a roadmap on best practice methods for the assessment or racial/ethnic discrimination in health surveys.

ACKNOWLEDGMENTS
We would like to thank the NIH Office of Behavioral Social Sciences Research for funding the field test and NCI Division of Cancer Control and Population Sciences for providing staff for this project. We would also like to acknowledge the Cancer Prevention Fellowship Program, Office of Preventive Oncology, National Cancer Institute for providing financial support for S. Shariff-Marco in her post-doctoral fellowship; Westat for their work on the cognitive interviews and behavior coding; and the Collaborative Psychiatric Epidemiology Surveys project for providing access to data from the National Survey of American Life, and the National Latino and Asian American Survey. Support for this work was also provided by the National Center for Minority Health and Health Disparities (MD 000508).

We would also like to acknowledge Dr. David Takeuchi for his participation as a member of the panel of external advisors for this research project and Drs. Martin Brown and Rachel Ballard-Barbash for their thoughtful reviews of our manuscript.

REFERENCES
27. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Qual Life Res. 2007;16(suppl 1):5–18.
28. Teresi JA. Statistical methods of examination of differential item functioning with applica-


**AUTHOR CONTRIBUTIONS**

**Design concept of study:** Shariff-Marco, Gee, Breen, Willis, Reeve, Ponce, Krieger, Alegria, Johnson, Brown

**Acquisition of data:** Shariff-Marco, Gee, Breen, Brown

**Data analysis and interpretation:** Shariff-Marco, Breen, Willis, Reeve, Ponce, Krieger,Williams, Mays

**Manuscript draft:** Shariff-Marco, Gee, Breen, Willis, Reeve, Ponce, Krieger, Landrine, Alegria, Mays, Johnson

**Statistical expertise:** Ponce, Krieger

**Acquisition of funding:** Shariff-Marco, Breen, Willis, Grant, Williams, Mays, Johnson

**Administrative, technical, or material assistance:** Shariff-Marco, Breen, Willis, Grant, Williams, Mays, Johnson

**Supervision:** Shariff-Marco, Gee, Breen, Ponce, Brown