Practice of Epidemiology

Assessment of Differential Item Functioning in the Experiences of Discrimination Index

The Coronary Artery Risk Development in Young Adults (CARDIA) Study

Timothy J. Cunningham*, Lisa F. Berkman, Steven L. Gortmaker, Catarina I. Kiefe, David R. Jacobs, Jr., Teresa E. Seeman, and Ichiro Kawachi

* Correspondence to Dr. Timothy J. Cunningham, Centers for Disease Control and Prevention, 4770 Buford Highway NE, Mailstop K-22, Atlanta, GA 30341 (e-mail: tjcunningham@cdc.gov).

The psychometric properties of instruments used to measure self-reported experiences of discrimination in epidemiologic studies are rarely assessed, especially regarding construct validity. The authors used 2000–2001 data from the Coronary Artery Risk Development in Young Adults (CARDIA) Study to examine differential item functioning (DIF) in 2 versions of the Experiences of Discrimination (EOD) Index, an index measuring self-reported experiences of racial/ethnic and gender discrimination. DIF may confound interpretation of subgroup differences. Large DIF was observed for 2 of 7 racial/ethnic discrimination items: White participants reported more racial/ethnic discrimination for the "at school" item, and black participants reported more racial/ethnic discrimination for the "getting housing" item. The large DIF by race/ethnicity in the index for racial/ethnic discrimination probably reflects item impact and is the result of valid group differences between blacks and whites regarding their respective experiences of discrimination. The authors also observed large DIF by race/ethnicity for 3 of 7 gender discrimination items. This is more likely to have been due to item bias. Users of the EOD Index must consider the advantages and disadvantages of DIF adjustment (omitting items, constructing separate measures, and retaining items). The EOD Index has substantial usefulness as an instrument that can assess self-reported experiences of discrimination.

African Americans; bias (epidemiology); observer variation; prejudice; psychometrics; questionnaires; reproducibility of results

Discrimination is a complex system of social relationships that unfairly limit the opportunities and agency of specific groups, such as racial/ethnic minorities and women. It manifests on at least 3 levels: institutional, personally mediated, and internalized (1, 2). Increasingly, psychometric instruments are being used in epidemiologic studies to measure self-reported experiences of discrimination and to explain disparities in population health in the United States and elsewhere (3–5). One commonly used psychometric instrument that quantifies self-reported experiences of both racial/ethnic and gender discrimination is the Experiences of Discrimination (EOD) Index (6–14). To date, the studies that have employed versions of the EOD Index have had mixed results, with investigators differentially reporting linear associations, nonlinear associations, and no associations between self-reported experiences of discrimination and health. Despite these mixed results, other versions of the EOD Index have been shown to demonstrate good construct validity, high internal consistency reliability, and test-retest reliability (6). However, the presence of several items displaying differential item functioning (DIF) in the EOD Index could still threaten its construct validity and confound the interpretation of subgroup differences (15, 16).

A systematic review showed that the psychometric properties of instruments used to measure self-reported experiences

of discrimination in existing epidemiologic studies are rarely assessed thoroughly (3). Only 1 previous study has examined DIF in any version of the EOD Index to date. Krieger et al. (6) found DIF by race/ethnicity in the ''getting service in a store or restaurant'' item of a 9-item version of the EOD Index for racial/ethnic discrimination. DIF is found when an item within an instrument behaves differently for different groups (17). The 2 types of DIF are item impact and item bias (18). A DIF item is considered to display item impact when the exogenous variable that differentiates the groups is relevant to the latent construct measured by the psychometric instrument. In contrast, a DIF item is considered to demonstrate item bias when the differences are illegitimate and unrelated to the latent construct. For example, Cole et al. (19) found that 2 items within the Center for Epidemiologic Studies Depression (CES-D) Scale displayed DIF by the exogenous variable race/ethnicity. The ''people unfriendly'' and ''people disliked me'' items had higher endorsement among blacks compared with whites. The authors concluded that these 2 items displayed item bias, were confounded with experiences of racial/ethnic discrimination, and were a threat to the construct validity of the CES-D Scale. If the latent construct measured by the CES-D Scale were experiences of racial/ethnic discrimination and we observed the same differences for the 2 items described above, then that DIF would have been considered item impact.

An obstacle to examining DIF has been implementation using standard statistical packages. Recent advances in user-friendly software permit more comprehensive analysis of psychometric instruments used to measure latent constructs in epidemiologic studies. In this study, we examined DIF related to the exogenous variables race/ethnicity, gender, age, and educational attainment among participants in the Coronary Artery Risk Development in Young Adults (CARDIA) Study, using 2 frequency versions of a 7-item EOD Index for racial/ethnic and gender discrimination. We utilized JMetrik (created by J. Patrick Meyer, University of Virginia), which includes statistical software for DIF based on the Cochran-Mantel-Haenszel (CMH) procedure. Consistent with the previous study (6), we hypothesized that DIF by race/ethnicity would be observed in the EOD Index for racial/ethnic discrimination.

MATERIALS AND METHODS

Study participants

CARDIA is a prospective, multicenter investigation of the development of cardiovascular disease risk factors in young adulthood. In 1985–1986, 5,115 persons were recruited from 4 study centers: Birmingham, Alabama; Chicago, Illinois; Minneapolis, Minnesota; and Oakland, California. A stratified random sampling procedure was used to achieve balance at each center by race/ethnicity (black, white), gender (women, men), age (18–24 years and 25–30 years), and educational attainment (high school graduate or less, some college or more) at baseline. Of eligible participants, 50% enrolled in the study. Additional details about study participants are available elsewhere (20).

This analysis used data collected in year 15 (2000–2001) from participants with complete data for the two 7-item frequency versions of the EOD Index for racial/ethnic and gender

discrimination. At year 15, 74% of the surviving participants were examined. The institutional review board at each center approved the CARDIA study protocol, and informed consent was obtained from each participant.

The EOD Index

Seven-item frequency versions of the EOD Index were used to measure self-reported experiences of discrimination. The EOD Index is a self-report measure of whether an individual's group membership increases his or her propensity to experience discrimination attributed to membership in that group. Figure 1 illustrates an effect indicator model, whereby items are seen as effects of the latent construct. Since institutional discrimination and personally mediated discrimination act together and may harm health through multiple pathways across the life course, double-headed arrows are used to signify this complex interaction. The single-headed arrows signify their direct effect on the effect indicators, which correspond with EOD Index items.

Participants were asked about their experiences of discrimination due to race/ethnicity and gender using the appropriate version of the EOD Index (shown in Figure 2). Participants were asked whether they had ''ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior'' because of their race/ethnicity or gender in any of 7 domains. If participants answered ''yes'' to any of these questions, they were asked how often this had occurred (rarely, sometimes, or often). For each item, 0 points were assigned for the answer ''no,'' 1 point for ''rarely,'' 2 points for ''sometimes,'' and 3 points for ''often.'' We were primarily interested in the items measuring experiences of discrimination and did not assess DIF in the responses to items measuring unfair treatment.

DIF analyses

We examined DIF according to 4 exogenous variables at year 15 (reference group, focal group): 1) race/ethnicity (white, black); 2) gender (women, men); 3) age (33–39 years, 40–50 years); and 4) educational attainment (high school graduate or less, some college or more). We scored both versions of the EOD Index as a summated rating scale, with scores ranging from 0 to 21. DIF analyses were conducted with JMetrik, which is a free and open-source statistical software package (available for download at www.ItemAnalysis.com). JMetrik includes a data management system, point-and-click operation, and a user-friendly interface. This analysis was based on the CMH procedure, which examines the strength of associations by comparing the observed and expected totals in 3-way contingency tables (21). When the CMH procedure is applied to DIF, reference and focal groups are matched on the latent construct of interest (22). For polytomous items, DIF is assessed by examining the standardized mean difference, which is the difference between the unweighted item mean of the focal group and the weighted item mean of the reference group. The weights applied to the reference group are applied so that the weighted number of reference group participants is the same as that in the focal group with the same total score. The effect size for the standardized mean
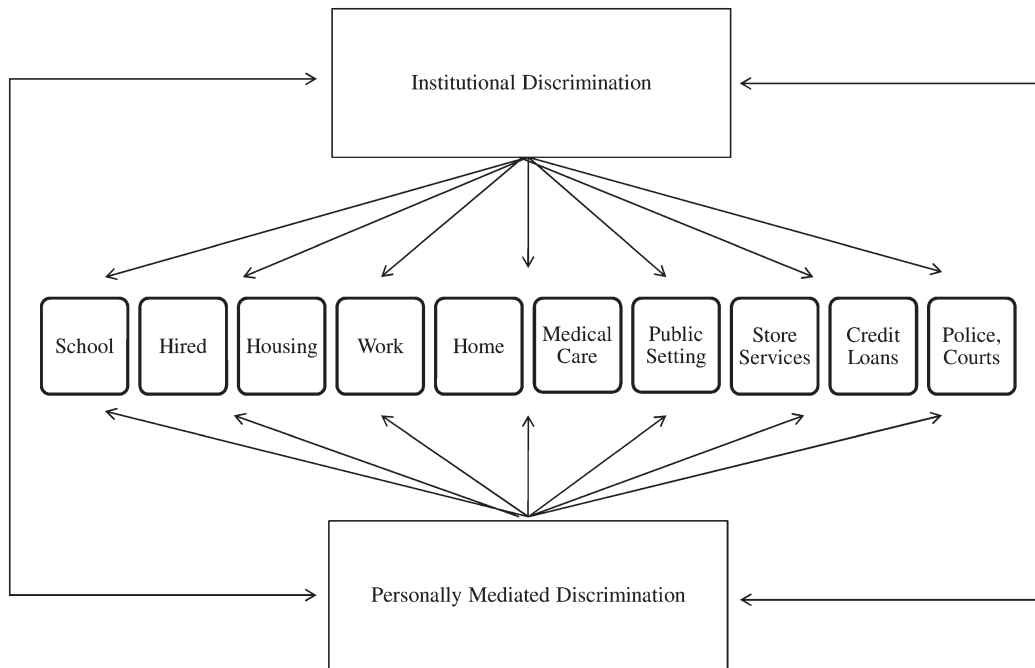
Figure 1. Model representing experiences of discrimination. Response effect indictors are represented by the row of 10 boxes in the center and correspond with items from different versions of the Experiences of Discrimination Index. The double-headed arrows signify the complex interaction between institutional and personally mediated discrimination, and the single-headed arrows signify their direct effects.

difference is computed by dividing the standardized mean difference by the total group-item standard deviation. Missing item responses are scored as 0 points.

According to the degree of DIF present for each item, JMetrik uses the Educational Testing Service (Princeton, New Jersey) classification system for dichotomous items and the National Assessment of Educational Progress (National Center for Education Statistics, US Department of Education) classification system for polytomous items. A polytomous item is classified into one of 3 categories—AA, BB, or CC (23)—according to whether the observed DIF is negligible, intermediate, or large:

1) category AA if either Mantel's chi-square is not significantly different from zero ($P \geq 0.05$) or the absolute value of the effect size is less than or equal to 0.17;
2) category BB if Mantel's chi-square is significant ($P < 0.05$) and the absolute value of the effect size is over 0.17 and less than or equal to 0.25; and
3) category CC if Mantel's chi-square is significant and the absolute value of the effect size is over 0.25.

Internal consistency analyses

Internal consistency reliability analyses were examined to estimate how well the items that reflect the same construct yield similar results with 1 measurement occasion. For both versions of the EOD Index, we examined the Cronbach's alpha (a measure of inter-item consistency, which ranges from 0.0 to 1.0), item-total correlations (Pearson correlation

of an item with the remainder of the index with that item omitted), and inter-item correlations (Pearson correlation between each pair of items) by race/ethnicity in SAS 9.2 (SAS Institute Inc., Cary, North Carolina). According to classical psychometric theory, one can select empirical indicators from various possible items to measure latent constructs (17). Cronbach's alpha reflects the degree to which a given participant provides correlated responses to the various items of the index and should only be used for effect indicator models measuring unidimensional latent constructs.

RESULTS

Responses were skewed towards the "no" category for all items in the EOD Index for racial/ethnic discrimination for both blacks and whites (Table 1) and the EOD Index for gender discrimination for women and men (Table 2). Scores for both versions could range from 0 to 21. The median total EOD Index score for racial/ethnic discrimination was 4 (interquartile range (IQR), 1–7) among blacks and 0 (IQR, 0–0) among whites. The median total EOD Index score for gender discrimination was 2 (IQR, 0–5) among women and 0 (IQR, 0–2) among men.

DIF analyses

Items functioned differently by race/ethnicity for both EOD indices. For the EOD Index for racial/ethnic discrimination, items 1 and 3 functioned differently by race/ethnicity, with

We are going to ask you a number of questions related to discrimination. Please select one response on questions 1 and 2.

1. If you feel you have been treated unfairly, do you usually:
__ Accept it as a fact of life?
__ Try to do something about it?

2. And if you have been treated unfairly, do you usually:
__ Talk to other people about it?
__ Keep it to yourself?

Please read each statement and check either the "No" or "Yes" box. If you check "Yes", please indicate how often you have experienced this (rarely, sometimes, or often).

3. Have you ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior in any of the following seven situations because of your **race or color?**

3a. At school	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3b. Getting a job	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3c. Getting housing	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3d. At work	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3e. At home	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3f. Getting medical care	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
3g. On the street or in a public setting	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often

4. Have you ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior in any of the following seven situations because of your **gender?**

4a. At school	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4b. Getting a job	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4c. Getting housing	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4d. At work	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4e. At home	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4f. Getting medical care	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often
4g. On the street or in a public setting	____ No ____ Yes	If Yes, how often? ____Rarely ____Sometimes ____Often

Figure 2. Racial/ethnic discrimination and gender discrimination versions of the Experiences of Discrimination Index.

white participants reporting more racial/ethnic discrimination for the "at school" item and black participants reporting more racial/ethnic discrimination for the "getting housing" item (Table 3). Items 1 and 3 were flagged as CC by the CMH procedure, because the absolute value of the effect size exceeded the 0.25 cutpoint. Items did not function differently by gender, age, or educational attainment in this index. For the EOD Index for gender discrimination, items functioned differently by race/ethnicity (Table 3) and educational attainment (not shown). Items 1, 2, 3, and 5 functioned differently by race/ethnicity, with white participants reporting more gender discrimination for the "at school" and "at home" items and black participants reporting more gender discrimination for "getting a job" and "getting housing." Items 1, 3, and 5 were flagged as CC by the CMH procedure, because the absolute value of the effect size exceeded the 0.25 cutpoint. Item 2 was

Table 1. Item Responses (%) for the Racial/Ethnic Discrimination Version of the Experiences of Discrimination Index, by Race/Ethnicity, CARDIA Study, 2000–2001

Item	Blacks				Whites			
	No	Rarely	Sometimes	Often	No	Rarely	Sometimes	Often
At school	69.20	8.37	18.97	3.45	92.86	3.62	2.84	0.67
Getting a job	53.86	10.48	28.34	7.32	94.67	3.36	1.76	0.21
Getting housing	72.25	8.08	15.05	4.63	99.38	0.20	0.40	0.00
At work	50.64	11.77	28.86	8.72	93.38	3.41	2.59	0.60
At home	96.00	1.76	1.76	0.47	99.59	0.16	0.21	0.05
Getting medical care	85.48	4.10	8.96	1.46	99.43	0.16	0.36	0.05
On the street or in a public setting	40.11	16.39	34.89	8.61	82.89	11.17	5.53	0.41

Abbreviation: CARDIA, Coronary Artery Risk Development in Young Adults.

Table 2. Item Responses (%) for the Gender Discrimination Version of the Experiences of Discrimination Index, by Gender, CARDIA Study, 2000–2001

Item	Women				Men			
	No	Rarely	Sometimes	Often	No	Rarely	Sometimes	Often
At school	77.65	8.01	12.33	2.01	87.45	5.03	6.40	1.12
Getting a job	71.22	9.58	16.50	2.70	79.12	7.02	11.37	2.49
Getting housing	88.46	3.49	6.48	1.57	90.12	3.23	4.41	2.24
At work	59.97	12.48	22.79	4.76	75.64	8.83	12.80	2.73
At home	87.77	4.37	6.78	1.08	95.03	2.61	1.86	0.50
Getting medical care	89.54	3.63	5.75	1.08	94.78	1.80	2.36	1.06
On the street or in a public setting	56.04	17.19	23.04	3.73	89.54	3.63	5.75	1.08

Abbreviation: CARDIA, Coronary Artery Risk Development in Young Adults.

flagged as BB because the absolute value of the effect size was over 0.17 but less than 0.25. Item 3 functioned differently by educational attainment, with participants with high school graduation or less education reporting more gender discrimination for the "getting housing" item. Items did not function differently by gender or age in this index.

Internal consistency analyses

As measured by Cronbach's alpha, the internal consistency reliability of the EOD Index for racial/ethnic discrimination was 0.82 for all participants, 0.79 for blacks, and 0.66 for whites (Table 4). The Cronbach's alpha values that would result if a given item were deleted and the item-total correlation for each item within the EOD Index for racial/ethnic discrimination for blacks and whites are also presented. As measured by Cronbach's alpha, the internal consistency reliability of the EOD Index for gender discrimination was 0.78 for all participants, 0.78 for blacks, and 0.74 for whites (Table 4). The Cronbach's alpha values that would result if a given item were deleted and the item-total correlation for each item for the EOD Index for gender discrimination for blacks and whites are also presented (Table 4). The item-total correlation ranged from 0.23 to 0.68 for the EOD Index for racial/ethnic discrimination and from 0.29 to 0.66 for the EOD Index for gender discrimination (Table 4). In addition,

Table 3. Differential Item Functioning by Race/Ethnicity (White vs. Black) for the Racial/Ethnic Discrimination and Gender Discrimination Versions of the Experiences of Discrimination Index, CARDIA Study, 2000–2001

Item	Mantel χ^2	P Value	Effect Size[a]	NAEP Category[b]
Racial/ethnic discrimination				
At school	75.61	<0.01	−0.49	CC
Getting a job	8.78	<0.01	0.16	AA
Getting housing	40.72	<0.01	0.42	CC
At work	4.94	0.03	−0.03	AA
At home	1.23	0.27	−0.49	AA
Getting medical care	2.20	0.14	0.03	AA
On the street or in a public setting	4.16	0.04	0.08	AA
Gender discrimination				
At school	91.66	<0.01	−0.35	CC
Getting a job	27.02	<0.01	0.18	BB
Getting housing	88.49	<0.01	0.40	CC
At work	5.76	0.02	−0.08	AA
At home	28.49	<0.01	−0.28	CC
Getting medical care	0.78	0.38	−0.05	AA
On the street or in a public setting	12.38	<0.01	0.11	AA

Abbreviations: CARDIA, Coronary Artery Risk Development in Young Adults; NAEP, National Assessment of Educational Progress.

[a] The effect size for the standardized mean difference is computed by dividing the standardized mean difference by the total group-item standard deviation.

[b] NAEP categories indicate whether the differential item functioning is negligible (AA), intermediate (BB), or large (CC).

Table 4. Cronbach's Alpha After Omission of Each Item and Item-Total Correlation for the Racial/Ethnic Discrimination and Gender Discrimination Versions of the Experiences of Discrimination Index, by Race/Ethnicity, CARDIA Study, 2000–2001

Item	Blacks		Whites	
	Cronbach's α if Item Deleted	Item-Total Correlation[a]	Cronbach's α if Item Deleted	Item-Total Correlation[a]
Racial/ethnic discrimination				
At school	0.75	0.56	0.59	0.47
Getting a job	0.72	0.68	0.59	0.47
Getting housing	0.76	0.52	0.65	0.32
At work	0.75	0.56	0.57	0.52
At home	0.80	0.23	0.66	0.27
Getting medical care	0.77	0.47	0.65	0.27
On the street or in a public setting	0.74	0.58	0.61	0.46
Total[b]	0.79		0.66	
Gender discrimination				
At school	0.75	0.53	0.68	0.56
Getting a job	0.72	0.66	0.69	0.52
Getting housing	0.75	0.50	0.74	0.26
At work	0.75	0.55	0.67	0.58
At home	0.79	0.29	0.73	0.33
Getting medical care	0.76	0.45	0.72	0.38
On the street or in a public setting	0.74	0.57	0.68	0.53
Total[b]	0.78		0.74	

Abbreviations: CARDIA, Coronary Artery Risk Development in Young Adults; EOD, Experiences of Discrimination.
[a] Cronbach's α for the 7-item EOD Index.
[b] The item total correlation is the Pearson correlation between the specific item and the remainder of the EOD Index.

the inter-item correlations ranged from 0.07 to 0.47 for the EOD Index for racial/ethnic discrimination and from 0.13 to 0.54 for the EOD Index for gender discrimination (Table 5).

DISCUSSION

This study demonstrates the importance of assessing DIF and internal consistency with 2 versions of the EOD Index. We observed large DIF by race/ethnicity for some items in both indices. In the EOD Index for racial/ethnic discrimination, 2 items showed large DIF after matching on overall self-reported experiences of racial/ethnic discrimination: Whites reported more racial/ethnic discrimination for the "at school" item, and blacks reported more racial/ethnic discrimination for the "getting housing" item. In other words, whites and blacks with the same total score for discrimination, on average, endorsed more discrimination at school and more discrimination in getting housing, respectively. In the EOD Index for gender discrimination, 3 items showed large DIF after matching on overall self-reported experiences of gender discrimination: Whites reported more gender discrimination for the "at school" and "at home" items, and blacks reported more gender discrimination for the "getting housing" item. We did not observe any evidence of DIF by gender or age for either version of the EOD Index. Additionally, we observed reasonable results among blacks and whites from the analyses

of internal consistency reliability, item-total correlations, and inter-item correlations.

Similar to the previous DIF study by Krieger et al. (6) using a 9-item version of the EOD Index for racial/ethnic discrimination and the multiple-indicator, multiple-cause approach, we found evidence of DIF by race/ethnicity and no evidence of DIF by gender and age. In that version of the EOD Index, Krieger et al. only observed DIF by race/ethnicity in the "getting service in a store or restaurant" item for blacks and not other items. This item was not included in the 7-item version of the EOD Index for racial/ethnic discrimination assessed in our analysis. Although these DIF observations do not remain consistent over variations in participants and settings, it appears that blacks experience more racial/ethnic discrimination in service-oriented contexts than whites. While we did not find any evidence of DIF by educational attainment for the racial/ethnic discrimination index, we did observe intermediate DIF by educational attainment for the "getting housing" item within the gender discrimination index, which was endorsed more by participants whose educational attainment was high school graduation or less. The absence of large DIF by gender, age, or educational attainment reflects a lack of confounding.

Though our observations indicate that large DIF by race/ethnicity is present in both the EOD Index for racial/ethnic discrimination and the EOD Index for gender discrimination, we suspect that the nature of the DIF varies. The large DIF by

Table 5. Inter-Item Correlation for the Racial/Ethnic Discrimination and Gender Discrimination Versions of the Experiences of Discrimination Index, by Race/Ethnicity, CARDIA Study, 2000–2001

Item	Blacks							Whites						
	At School	Getting a Job	Getting Housing	At Work	At Home	Getting Medical Care	On the Street or in a Public Setting	At School	Getting a Job	Getting Housing	At Work	At Home	Getting Medical Care	On the Street or in a Public Setting
Racial/ethnic discrimination														
At school														
Getting a job	0.46							0.26						
Getting housing	0.35	0.49						0.26	0.15					
At work	0.41	0.55	0.30					0.33	0.47	0.21				
At home	0.19	0.16	0.17	0.14				0.25	0.20	0.25	0.22			
Getting medical care	0.35	0.38	0.38	0.29	0.29			0.18	0.26	0.21	0.15	0.24		
On the street or in a public setting	0.44	0.51	0.38	0.46	0.10	0.32		0.39	0.30	0.20	0.37	0.07	0.14	
Gender discrimination														
At school														
Getting a job	0.45							0.37						
Getting housing	0.35	0.45						0.21	0.24					
At work	0.35	0.54	0.31					0.42	0.54	0.18				
At home	0.26	0.20	0.20	0.15				0.27	0.17	0.13	0.27			
Getting medical care	0.30	0.35	0.35	0.26	0.26			0.28	0.25	0.20	0.25	0.22		
On the street or in a public setting	0.40	0.48	0.35	0.45	0.18	0.33		0.50	0.34	0.14	0.41	0.26	0.28	

Abbreviation: CARDIA, Coronary Artery Risk Development in Young Adults.

race/ethnicity observed in the EOD Index for racial/ethnic discrimination is probably due to item impact and reflects valid group differences between blacks and whites in the latent construct of self-reported experiences of racial/ethnic discrimination. We recognize the complexity of how people experience, perceive, cope with, and report racial/ethnic discrimination; however, it is apparent that racial/ethnic group membership influences those experiences (4).

Racial/ethnic discrimination experienced by racial/ethnic minority groups, especially blacks, in the national housing market has been well-documented through observational studies (24, 25) and quasi-experimental audit studies (26–28) and may account for the large DIF observed for the "getting housing" item. For example, Yinger (27) used a quasi-experimental design known as "fair housing audits" to study the incidence and intensity of housing discrimination experienced by racial/ethnic minorities in terms of differential treatment by realtors regarding the numbers, types, and locations of housing units shown. Furthermore, there is a long history of racial/ethnic discrimination in the US housing market. Housing discrimination was propagated through the National Housing Act of 1934 and redlining and was permissible until passage of the Fair Housing Act of 1968. An explanation for the large DIF observed for the "at school" item among whites is less clear. One potential explanation is the documented attitude and belief that affirmative action in higher education discriminates against whites. Past studies have shown that some whites think affirmative action is institutional discrimination against their racial/ethnic group that entails quota policies (29–33). These differences for blacks and whites appear to be legitimately relevant to the large DIF by

race/ethnicity observed in the EOD Index for racial/ethnic discrimination. Conversely, since race/ethnicity is not as relevant to the EOD Index for gender discrimination, the large DIF by race/ethnicity observed for this instrument is more likely due to item bias and reflects confounding with racial/ethnic discrimination.

Since we observed DIF, we performed additional item bias analyses by examining internal consistency reliability using 3 approaches (Cronbach's alpha, item-total correlations, and inter-item correlations) for blacks and whites separately (34). The presence of DIF did not appear to substantially affect the internal consistency reliability of either version of the EOD Index. Additionally, the differences observed for the Cronbach's alphas for the EOD Index for racial/ethnic discrimination appear to reflect racial/ethnic differences in the prevalence of the latent construct, although the coherence of that index may be the same for blacks and whites (17). Whereas omitting items with evidence of DIF resulted in lower Cronbach's alphas in some instances, the reduction was too small to be considered a result of more than a mathematical artifact. The average item-total correlations and average inter-item correlations for both versions of the EOD Index were reasonable as well. Therefore, the deletion or modification of the items exhibiting DIF may not be necessary, and DIF may not have a major impact on the psychometric properties of these versions of the EOD Index.

Some limitations warrant consideration in the interpretation our results. One limitation of our study is that while we were able to detect uniform DIF, the CMH procedure is not sensitive to nonuniform DIF (35). Uniform DIF is present if the differences in the probability of answering an item equivalently

are constant across different levels of the latent construct. Nonuniform DIF is present if the differences in the probability of answering an item equivalently are not constant across different levels of the latent construct. Another limitation is our use of a marginal DIF approach that assessed DIF separately for each exogenous variable in contrast to a melting-pot DIF approach, which recognizes possible interactions among combinations of exogenous variables (9). For example, we did not examine whether our findings were due to black women's being different from black men, white women, or white men, since we examined race/ethnicity and gender separately. Although nonuniform DIF and interactions among exogenous variables may be relevant, a melting-pot approach is not readily available in JMetrik using the CMH procedure. One limitation of our approach to examining internal consistency reliability is that transient error and local dependence may inflate estimates of Cronbach's alpha. Lastly, while both versions of the EOD Index in this study reflect aspects of institutional and personally mediated discrimination, they lack items relevant to internalized discrimination. Though the EOD Index includes the response to unfair treatment items, it does not include items explicitly related to internalized discrimination like similar instruments do, such as the 31-item Measure of Indigenous Racism Experiences (5).

There are strengths of this study that should be considered as well. A noteworthy strength is our novel utilization of JMetrik—user-friendly, free, open-source statistical software. Another strength is the stratified random sampling design of the CARDIA Study, which resulted in large subgroups for the 4 exogenous variables (race/ethnicity, gender, age, and educational attainment). Additionally, we used numerous approaches that are appropriate for effect indicator models to evaluate item bias besides DIF analyses. Though unlikely, a causal indicator model would only be appropriate if there were one-to-one correspondence between actual institutional discrimination and personally mediated discrimination. In contrast to discrimination, socioeconomic position more appropriately fits a causal indicator model because education, occupation, income, and wealth are apparent causes. Lastly, our study is one of few assessing the psychometric properties of instruments used to measure self-reported experiences of discrimination in epidemiologic studies.

Given our observations that large DIF by race/ethnicity is present in both versions of the EOD Index, there are 3 options for DIF adjustment that have been identified: omitting the items, constructing separate measures, and retaining DIF items (36). The advantages and disadvantages of each of these alternatives have been discussed previously (36). Briefly, omitting DIF items permits comparison between groups, but it may adversely affect the reliability and content validity of a psychometric instrument by deleting important domains relevant to the latent construct. Constructing separate measures may inhibit comparisons between and across groups but may be most useful when examining 1 group only. In addition, although retaining DIF items could result in difficulties in interpretation and scoring, it may best represent the complexity of the latent constructs of interest. Users of the EOD Index should consider the advantages and disadvantages of each of 3 options for DIF adjustment when employing these instruments. Users of the EOD Index for gender discrimination

should also consider potential confounding with racial/ethnic discrimination.

This study should facilitate future examination of the psychometric instruments used in epidemiologic studies. As hypothesized, our observations indicated large DIF by race/ethnicity. Despite the fact that large DIF by race/ethnicity was present, its nature differed. The DIF observed in the EOD Index for racial/ethnic discrimination was probably due to item impact, while the DIF by race/ethnicity observed in the EOD Index for gender discrimination was more likely due to item bias. Nonetheless, as a whole, both versions of the EOD Index operate reasonably as instruments for assessing self-reported experiences of discrimination, and no items necessarily need to be deleted.

REFERENCES

1. Antonovsky A. The social meaning of discrimination. *Phylon*. 1960;21(1):81–95.
2. Jones CP. Confronting institutionalized racism. *Phylon*. 2002; 50(1/2):7–22.

3. Bastos JL, Celeste RK, Faerstein E, et al. Racial discrimination and health: a systematic review of scales with a focus on their psychometric properties. *Soc Sci Med*. 2010;70(7): 1091–1099.

4. Williams DR, Mohammed SA. Discrimination and racial disparities in health: evidence and needed research. *J Behav Med*. 2009;32(1):20–47.

5. Paradies YC, Cunningham J. Development and validation of the Measure of Indigenous Racism Experiences (MIRE). *Int J Equity Health*. 2008;7:9. (doi:10.1186/1475-9276-7-9).

6. Krieger N, Smith K, Naishadham D, et al. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Soc Sci Med*. 2005;61(7):1576–1596.

7. Landrine H, Klonoff EA. Racial discrimination and cigarette smoking among blacks: findings from two studies. *Ethn Dis*. 2000;10(2):195–202.

8. Krieger N. Racial and gender discrimination: risk factors for high blood pressure? *Soc Sci Med*. 1990;30(12): 1273–1281.

9. Krieger N, Sidney S. Racial discrimination and blood pressure: the CARDIA study of young black and white adults. *Am J Public Health*. 1996;86(10):1370–1378.

10. Mustillo S, Krieger N, Gunderson EP, et al. Self-reported experiences of racial discrimination and black-white differences in preterm and low-birthweight deliveries: the CARDIA Study. *Am J Public Health*. 2004;94(12):2125–2131.

11. Yen IH, Ragland DR, Greiner BA, et al. Racial discrimination and alcohol-related behavior in urban transit operators: findings from the San Francisco Muni Health and Safety Study. *Public Health Rep*. 1999;114(5):448–458.

12. Peters RM. Racism and hypertension among African Americans. *West J Nurs Res*. 2004;26(6):612–631.

13. Borrell LN, Jacobs DR Jr, Williams DR, et al. Self-reported racial discrimination and substance use in the Coronary Artery Risk Development in Adults Study. *Am J Epidemiol*. 2007; 166(9):1068–1079.

14. Borrell LN, Kiefe CI, Williams DR, et al. Self-reported health, perceived racial discrimination, and skin color in African Americans in the CARDIA Study. *Soc Sci Med*. 2006;63(6): 1415–1427.

15. Holland PW, Wainer H, eds. *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1993.

16. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw-Hill; 1994.

17. DeVellis RF. *Scale Development: Theory and Applications*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc; 2003.

18. Osterlind SJ, Everson HT. *Differential Item Functioning*. 2nd ed. Thousand Oaks, CA: Sage Publications, Inc; 2009.

19. Cole SR, Kawachi I, Maller SJ, et al. Test of item-response bias in the CES-D Scale. Experience from the New Haven EPESE Study. *J Clin Epidemiol*. 2000;53(3):285–289.

20. Friedman GD, Cutter GR, Donahue RP, et al. CARDIA: study design, recruitment, and some characteristics of the examined subjects. *J Clin Epidemiol*. 1988;41(11):1105–1116.

21. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22(4):719–748.

22. Dorans NJ, Schmitt AP, Bleistein CA. The standardization approach to assessing comprehensive differential item functioning. *J Educ Meas*. 1992;29(4):309–319.

23. Zwick R, Thayer DT, Mazzeo J. *Describing and Categorizing DIF in Polytomous Items*. Princeton, NJ: Educational Testing Service; 1997.

24. Antecol H, Cobb-Clark DA. Racial and ethnic discrimination in local consumer markets: exploiting the Army's procedures for matching personnel to duty locations. *J Urban Econ*. 2008; 64(2):496–509.

25. Ladd HF. Evidence on discrimination in mortgage lending. *J Econ Perspect*. 1998;12(2):41–62.

26. Page M. Racial and ethnic discrimination in urban housing markets—evidence from a recent audit study. *J Urban Econ*. 1995;38(2):183–206.

27. Yinger J. Measuring racial discrimination with fair housing audits—caught in the act. *Am Econ Rev*. 1986;76(5):881–893.

28. Galster G, Godfrey E. By words and deeds—racial steering by real estate agents in the US in 2000. *J Am Plan Assoc*. 2005; 71(3):251–268.

29. Unzueta MM, Gutierrez AS, Ghavami N. How believing in affirmative action quotas affects white women's self-image. *J Exp Soc Psychol*. 2010;46(1):120–126.

30. Wilson DC, Moore DW, McKay PF, et al. Affirmative action programs for women and minorities—expressed support affected by question order. *Public Opin Q*. 2008;72(3):514–522.

31. Unzueta MM, Lowery BS, Knowles ED. How believing in affirmative action quotas protects white men's self-esteem. *Organ Behav Hum Decis Process*. 2008;105(1):1–13.

32. Haley H, Sidanius J. The positive and negative framing of affirmative action: a group dominance perspective. *Pers Soc Psychol Bull*. 2006;32(5):656–668.

33. Kravitz DA, Platania J. Attitudes and beliefs about affirmative action—effects of target and of respondent sex and ethnicity. *J Appl Psychol*. 1993;78(6):928–938.

34. Zumbo BD. *A Handbook on the Theory and Methods of Differential Item Functioning (DIF): Logistic Regression Modeling as a Unitary Framework for Binary and Likert-Type (Ordinal) Item Scores*. Ottawa, Ontario, Canada: Directorate of Human Resources Research and Evaluation, Department of National Defense; 1999.

35. Rogers HJ, Swaminathan H. A comparison of logistic-regression and Mantel-Haenszel procedures for detecting differential item functioning. *Appl Psychol Meas*. 1993;17(2):105–116.

36. Smith LL. On the usefulness of item bias analysis to personality psychology. *Pers Soc Psychol Bull*. 2002;28(6):754–763.