



## Racial discrimination and health: A systematic review of scales with a focus on their psychometric properties

Joao Luiz Bastos<sup>a,\*</sup>, Roger Keller Celeste<sup>b</sup>, Eduardo Faerstein<sup>b</sup>, Aluisio J.D. Barros<sup>a</sup>

<sup>a</sup> Post-graduate Program in Epidemiology, Federal University of Pelotas, Department of Social Medicine, Rua Marechal Deodoro, 1160, 96020-220 Pelotas, Rio Grande do Sul, Brazil

<sup>b</sup> Post-graduate Program in Public Health, Institute of Social Medicine, State University of Rio de Janeiro, Rio de Janeiro, Brazil

### ARTICLE INFO

#### Article history:

Available online 1 February 2010

#### Keywords:

Race relations  
Prejudice  
Causality  
Questionnaires  
Psychometrics  
Systematic review  
Racial discrimination

### ABSTRACT

The literature addressing the use of the race variable to study causes of racial inequities in health is characterized by a dense discussion on the pitfalls in interpreting statistical associations as causal relationships. In contrast, fewer studies have addressed the use of racial discrimination scales to estimate discrimination effects on health, and none of them provided a thorough assessment of the scales' psychometric properties. Our aim was to systematically review self-reported racial discrimination scales to describe their development processes and to provide a synthesis of their psychometric properties. A computer-based search in PubMed, LILACS, PsycInfo, Scielo, Scopus and Web of Science was conducted without any type of restriction, using search queries containing free and controlled vocabulary. After initially identifying 3060 references, 24 scales were included in the review. Despite the fact that discrimination stands as topic of international relevance, 23 (96%) scales were developed within the United States. Most studies (67%,  $N = 16$ ) were published in the last 12 years, documenting initial attempts at scale development, with a dearth of investigations on scale refinements or cross-cultural adaptations. Psychometric properties were acceptable; sixteen of all scales presented reliability scores above 0.7, 19 out of 20 instruments confirmed at least 75% of all previously stated hypotheses regarding the constructs under consideration, and conceptual dimensional structure was supported by means of any type of factor analysis in 17 of 21 scales. However, independent researchers, apart from the original scale developers, have rarely examined such scales. The use of racial terminology and how it may influence self-reported experiences of discrimination has not yet been thoroughly examined. The need to consider other types of unfair treatment as concurrently important health-damaging exposures, and the idea of a universal instrument which would permit cross-cultural adaptations, should be discussed among researchers in this emerging field of inquiry.

© 2010 Elsevier Ltd. All rights reserved.

### Introduction

Social inequalities in health are increasingly becoming part of the mainstream public health research agenda worldwide. As a result, close relationships among health inequalities and existing forms of social stratification have been evidenced in many societies. In particular, "racial inequalities in health" stands as an emerging topic of investigation based on the persistence of racial disparities over time and on their intrinsically unfair nature. The health literature is characterized by a considerable amount of evidence systematically indicating worse health outcomes across the life-course for some minority groups (Williams & Mohammed, 2009),

with less emphasis being given to outcomes where a majority group shows poorer health status, and still fewer studies on conditions showing no racial differences (Bhopal, 2007).

Debate over racial inequities in health has focused on innate genetic differences, disparities in the distribution of individual health-behaviors (cultural traits, such as diet, exercise, tobacco use), and on the over-representation of some racial groups in lower socioeconomic strata as major causes of racial inequities in morbidity and mortality (Dressler, Oths, & Gravlee, 2005; Krieger, 2005b). Alternative perspectives to explain this type of health inequity are the structural-constructivist and the psychosocial stress models (Dressler et al., 2005). The first model emphasizes the intersection of racially stratified social structures with goals and aspirations constructed within racial groups, while the latter focuses on experiences of racism and discrimination as important, but not single-contributors, for racial inequities in health. In fact,

\* Corresponding author. Tel./fax: +55 53 3284 1300.  
E-mail address: [joao.luiiz.epi@gmail.com](mailto:joao.luiiz.epi@gmail.com) (J.L. Bastos).

research in the area shows that interpersonal discrimination is one of a myriad factors that produce racial inequities in health (Williams & Mohammed, 2009).

Two main methodological approaches have been utilized to estimate health consequences of discrimination for individuals belonging to various racial groups (Krieger, 2000). The most commonly adopted strategy consists of examining the association between the race variable and a selected health outcome in the context of multivariable regression models. Increasingly, discrimination effects on health-related outcomes are also assessed with the use of scales of experiences of discrimination. Such scales are designed to inquire respondents about their experiences of racial discrimination, so that the relationship between discrimination and ill health may be examined.

Several authors have commented on the difficulties of using the race variable in epidemiologic studies (Bastos et al., 2009; Kaufman, 2008; Penner & Saperstein, 2008; Travassos & Williams, 2004), raising concerns about the validity and the reliability of this variable. In a recent paper, for example, Lee (2009) observed that authors deemed important to have race as a category of analysis in their investigations, but they rarely defined this concept or articulated how race operated in their theoretical models. Furthermore, when racial inequities were evidenced, most researchers did not provide an explanation for how and why such findings resulted or their medical significance.

On the other hand, the literature assessing scales of experiences of racial discrimination in health-related studies is scarce. Several reviews on the topic have recently been published (Brondolo, Rieppi, Kelly, & Gerin, 2003; Harrell, Hall, & Taliaferro, 2003; Krieger, 2000, 2005a; Paradies, 2006b; Pascoe & Smart Richman, 2009; Williams & Mohammed, 2009; Williams, Neighbors, & Jackson, 2003; Wyatt et al., 2003). However, despite the fact that many racial discrimination scales can be found in those reviews, none of them evaluated these instruments from a psychometric perspective, including validity and reliability assessments. In addition, previous reviews on discrimination scales (Kressin, Raymond, & Manze, 2008; Utsey, 1998) were neither systematic nor comprehensive, so as to include instruments developed outside the United States (US). The first study, published more than a decade ago by Utsey (1998), reviewed six instruments with regard to their development, format, psychometric properties, and was not restricted to instruments assessing only self-reports of experiences of racial discrimination. A recent literature review by Kressin et al. (2008) identified 34 different measures to assess self-reported discrimination. This study focused on discrimination perpetrated by health care providers against minority groups living in the US, and included instruments for which no psychometric evaluation was available. Additionally, even though this latter review summarized a few psychometric properties of some included measures, it provided neither a detailed synthesis of their processes of content, convergent, and discriminant validation, nor an exhaustive assessment of empirical data in support of the scales' dimensional structure.

The fact that considerable advances in the area have taken place since the publication of Utsey's work and no worldwide systematic review is available on racial discrimination scales warrant further exploration of the topic. Therefore, the objective of the present study is to update and expand upon these mentioned studies by addressing a specific research question: What is the current stage of development of available psychometric instruments in terms of their performance to validly and reliably assess race-based discriminatory experiences? Following Krieger's (2001) conceptual review, in this study discrimination refers to "the process by which a member, or members, of a socially defined group is, or are, treated differently (especially unfairly) because of his/her/their

membership of that group." In addition, this study dealt only with personally-mediated discriminatory experiences, and not with other analytic levels, such as institutional or internalized racism (Jones, 2000).

## Methods

A computer-based search was carried out in PubMed, LILACS, PsycInfo, Scielo, Scopus and Web of Science from the earliest records in these databases until the end of January 2009. Search strategies incorporated controlled vocabulary at the highest level of each tree hierarchy whenever a hierarchical database thesaurus was available. If an appropriate controlled term could not be identified, a search strategy was built with text words based on the authors' experience. The process of building search queries was specific to each of the abovementioned databases and no limits (e.g. language, sex, age, etc.) were set up during the course of identification of papers. Since the majority of studies in this area report empirical evidence on the association between self-reported racial discrimination and health, and some of them simultaneously describe the construction of a discrimination scale for that purpose, relatively sensitive search strategies were developed to retrieve all potentially eligible scales. The list of databases, date of each bibliographic search, queries used, and the total number of articles retrieved are described in Table 1.

This review required studies to be focused on, but not limited to, the development of racial discrimination scales including, at least, the following three methodological steps (DeVellis, 2003; Streiner & Norman, 1998):

- Item development, with sources of items ranging from focus groups to key informant interviews, including literature reviews, theoretical frameworks on the subject matter, empirical investigations or panels of experts;
- Reliability analysis, including internal consistency measures, test-retest reliability and/or inter-observer agreement; and
- Validity analysis, encompassing content, and/or construct assessments.

Therefore, the over 150 available sets of items that have not been subjected to this psychometric development/assessment – recently described elsewhere (Paradies, 2006b) – were excluded from the present review. Whenever information regarding scale development or psychometric properties was spread over more than one bibliographic reference, these were combined, so as to provide a complete description of item development, reliability analysis, and validity analysis of each included scale.

We used EndNote software version 8 to create data files from the databases' search results. The first author examined the title and, when necessary, the abstract of each retrieved article in order to exclude studies from the next phase of the revision, that is, full text examination. In the absence of abstracts, the complete texts of the retrieved articles were checked. Previous literature reviews (Kressin et al., 2008; Paradies, 2006b; Utsey, 1998) as well as a book (Blank, Dabady, & Citro, 2004), book chapters (Krieger, 2000, 2005a) and retrieved papers in full text had their reference lists searched for further scales. Additionally, some leading authors in the field were contacted through electronic messages asking for unpublished scales and documents difficult to obtain in full text.

Articles meeting the criteria for full text examination were independently submitted to data extraction by the first two authors, using a checklist pre-tested in a sample of articles. Then, data were entered in EpiData version 3.1 with automatic checks for consistency and range (spreadsheet available from the authors upon request), and analyzed with Stata, version 9. Disagreements were discussed

**Table 1**

Databases included in the literature review, their official languages, date of search, search queries used, and number of articles retrieved.

Database	Database language	Search date	Search query	Use of controlled vocabulary (e.g. MeSH terms in PubMed)?	Number of articles retrieved
LILACS	Portuguese	January 21, 2009	([MH]"GRUPOS ETNICOS" or [MH]"GRUPOS POPULACIONAIS" or [MH]"GRUPOS MINORITARIOS" or [MH]"GRUPOS DE POPULAÇÕES CONTINENTAIS")	Yes	667
PsycInfo	English	January 23, 2009	((racial and ethnic groups) or (population) AND (racial and ethnic attitudes) or (racial and ethnic relations) or (racial and ethnic discrimination) or (racism) or (prejudice) or (social discrimination) AND (questionnaires) or (psychometrics) or (measurement) or (test construction) or (test reliability) or (test validity) or (statistical validity)):Any Field	Yes	583
PubMed	English	January 21, 2009	("minority groups"[MeSH Terms] OR "ethnic groups"[MeSH Terms] OR "continental population groups"[MeSH Terms]) AND ("prejudice"[MeSH Terms] OR "race relations"[MeSH Terms] OR "discrimination"[Title/Abstract] OR "social perception"[MeSH Terms] OR "stereotyping"[MeSH Terms]) AND ("questionnaires"[MeSH Terms] OR "reproducibility of results"[MeSH Terms] OR "validation studies as topic"[MeSH Terms] OR "validation studies"[Publication Type] OR "factor analysis, statistical"[MeSH Terms] OR "psychometrics"[MeSH Terms])	Yes	580
Scielo	Portuguese, Spanish and English	January 22, 2009	Portuguese: (raça OR etnia) AND (preconceito OR discriminação OR relações raciais OR racismo) Spanish: (raza OR etnia) AND (prejuicio OR discriminación OR relaciones raciales OR racismo) English: (race OR ethnicity) AND (prejudice OR discrimination OR race relations OR racism)	No	130
Scopus	English	January 22, 2009	INDEXTERMS(racism OR discrimination OR prejudice) AND INDEXTERMS(psychometric OR validity OR reliability)	Yes	659
Web of Science	English	January 22, 2009	(minority groups OR ethnic group* OR population group OR race) AND (prejudice OR racism* OR discrimination* OR stereotyping) AND (questionnaire* OR reliability OR reproducibility OR validation OR validity OR psychometric* OR factor analys*)	No	441

until a final consensus could be achieved between the two authors. The reviewers were not blinded in relation to the authors of the original articles, and, due to the methodological nature of the studies reviewed, neither a formal meta-analysis nor an assessment of publication bias was conducted in the present study.

For each scale, data were extracted on year of publication, journal of publication, country of origin, sample size, participants' age range, sex, number of scale items, mean duration of scale completion, method of data collection (face-to-face, self-completed, etc.), reliability and validity scores. Participants' race or ethnicity were also extracted from the reviewed studies. By doing so, we acknowledge that there was a lack of precision in the usage of the terms race and ethnicity by scale authors; importantly, these concepts were eventually used interchangeably and, sometimes, without the needed theoretical clarification as to which meaning authors were attributing to these terms. Methods adopted during scale development were also described here. The number of conceptual dimensions assessed and the intent to identify the scale's respondent as a potential perpetrator of discriminatory acts were recorded. Availability of cross-culturally adapted versions (Berry, Poortinga, Segall, & Dasen, 2007; Herdman, Fox-Rushby, & Badia, 1998; Hunt & Bhopal, 2004) and the use of racial terminology – i.e. any mention to race, racism, discrimination and/or reference to racial groupings in scale items or instructions – were also scrutinized.

During scale evaluation, we observed whether or not authors adopted an explicit theoretical framework, and which efforts they employed to examine content validity, scale reliability, and construct validity, including assessment of dimensional structure. A satisfactory condition was positively rated, while an unsatisfactory one received a null score. We positively rated scales that:

- Stated clearly the underlying theory guiding scale construction;

- Described any effort employed towards content validation of scale items, such as applying procedures to select relevant items, regardless of specification of the full content domain that was relevant to the particular measurement situation. Content validity refers to the extent to which a specific set of items reflects a content domain (e.g. the discrimination construct or one of its sub-constructs) (DeVellis, 2003);
- Reported a (Cronbach's alpha) coefficient or other forms of reliability above 0.70 for the scale as a whole or its sub-scales. Pearson correlations were considered inadequate (Streiner & Norman, 1998). Reliability may be conceptualized as the extent to which a scale produces similar measurements for individuals under different circumstances (Streiner & Norman, 1998);
- Had at least 75% of the hypotheses specified in advance confirmed regarding the construct under consideration and its relationship with other construct(s) or variables. This included convergent, discriminant validity and extreme groups comparisons. Convergent validity consists in examining whether a scale's score is associated to other variables and other measures of the same construct to which it should be related. Discriminant validity is the extent to which the scale's score does not correlate with unrelated constructs. Extreme group comparisons refer to the capacity of a scale in producing different scores among two or more groups of participants, known to possess distinct amounts of the construct under consideration (Streiner & Norman, 1998); and
- Found support for the conceptual dimensional structure by means of any type of factor analysis. Conceptual dimensional structure refers to the theorized underlying dimensions of the construct under consideration (DeVellis, 2003). For example, if discrimination has been theorized as having two underlying dimensions (e.g. major and everyday discrimination), this must be empirically checked during scale development/refinement.

Data pertaining to studies meeting the inclusion criteria were analyzed with absolute and relative frequencies. Ratings attributed to scales during this evaluation process were displayed in a contingency table.

## Results

Although we were able to identify 27 scales, which were, in principle, eligible for inclusion in the present review, three of them (Harrell, Merchant, & Young, 1997; Lang, 2001, p. 101; Terrell & Miller, 1980) could not be analyzed (see Fig. 1). The Racism and Life Experiences Scales (Harrell et al., 1997), as well as the Racial Discrimination Index (Terrell & Miller, 1980) were not published in peer-reviewed journals, and we could not access them, even after contacting the authors and other researchers in the field. The Schedule of Racist Events – Generic (SRE-G) (Lang, 2001) was published as a PhD thesis in 2001, but was not available in full text as well. However, according to Dr Lang (personal communication, February 10, 2009), Landrine, Klonoff, Corral, Fernandez, and Roesch (2006) subsequently published an article utilizing the SRE-G, which was included in this review.

Table 2 shows that the majority of the 24 reviewed scales were published in the last 12 years, and 23 originated from the US. One fifth of the studies were published in the Journal of Applied Social Psychology, followed by Ethnicity & Disease and Journal of Counseling Psychology. All scales were published in English, but two of them were also available in Spanish. None of these were subjected to a complete process of cross-cultural adaptation. One-half of the reviewed studies recruited between 201 and 1000 participants, with only four including more than one thousand individuals (Table 2). The interviewees' age range was not available in four studies; for those with this information, the interviewees' age range was expressive: 12 of these studies recruited individuals who were, at least, as far as 31 years of age apart. Studies recruited participants from multiple racial or ethnic groupings, such as blacks, Asians, West Indians, and whites. In cases in which only one racial or ethnic group was considered, blacks were more frequently recruited. All scales considered the respondent solely as a potential victim of racial discrimination. Scales consisted frequently of less than 30 items (Table 2), all measured through the Likert-reponse format. The mean subject-to-item ratio was 23 (standard deviation = 30; range 3–119). Duration of scale completion was between 10 and 15 min for most scales, and the most frequent method of data collection was self-completion, followed by face-to-face interviews and self-completed methods.

With regard to the methods adopted during item development, literature review showed the highest frequency (Table 3). For content validation, some authors subjected the scales' items to a review by members of the target population through focus groups, pre-tests, pilot studies or interviews. Content validity was also assessed by panels of experts in three of the scales, and through literature review in only one case. Table 3 also shows that 22 scales had their internal consistency scores examined, while only nine were subjected to an assessment of test-retest reliability. Conceptual dimensional structure was checked by means of principal component analyses, exploratory factor analysis and confirmatory factor analysis in 11, 9, and 8 scales, respectively. Convergent and discriminant analyses were employed in 16 and 7 of the cases, in that order. Structural equation modeling was applied in three cases.

Only eight studies documented efforts towards content validation (Table 4). A third of the scales presented reliability scores under 0.70. There was a general trend over confirming at least 75% of all previously stated hypotheses regarding the constructs under consideration. Conceptual dimensional structure was supported by

means of any type of factor analysis in 17 out of 21 scales. Two out of 24 studies did not state clearly the theoretical framework adopted for development (Table 4). In the remaining 22 studies, theories adopted were diverse: acculturation theories, social/ethnic identity models, and theoretical formulations on how racism manifests itself in the US were cited. The most common theoretical framework was the stress-coping model. Accordingly, scales were rarely devised to assess only self-reported experiences of discrimination, such that closely related constructs (e.g. attitudes toward racial integration, backlash feelings, racial climate, ethnic identity, institutional/collective racism, emotional and behavioral coping responses to racism and discrimination, and appraisal of racists events) were important to be measured concurrently.

Many scales were designed to measure multi-dimensional constructs in which discrimination was only one of them. Therefore, the number of dimensions within the discrimination sub-construct was not always well reported or analyzed. The number of assessed dimensions varied expressively, ranging from 1 to 10, with the majority of scales being designed to assess three dimensions. The bulk of them were closely related to the interviewees' experiences of racial discrimination, assessed as personally mediated events. For that purpose, authors commonly used the word discrimination while labeling these dimensions, such as in the following examples: recent discrimination, past week discrimination, lifetime discrimination, perceived discrimination, frequency of everyday mistreatment, peer discrimination, and educational discrimination. The second most common group of examined dimensions referred to the respondents' responses, reactions to and appraisal of discriminatory experiences. Another relatively common but less frequent group of dimensions dealt with ethnic identity.

There was a general trend over making race salient in the assessment of experiences of discrimination, by asking repeated questions about "racial discrimination" or experienced discriminatory events "because of the respondents' own race or ethnicity." This is in contrast to only two scales that assessed the frequency of life events first (e.g. whether or not interviewees were unfairly treated in a restaurant, store, etc.), and, subsequently, queried respondents about whether these events were perceived as racially biased. Additionally, during the development processes of all reviewed scales, the potential impact of racial terminology on the scales' psychometric properties was neither assessed nor discussed.

## Discussion

The fact that the development of discrimination scales has been concentrated in the US is remarkable. This finding probably reflects the central role race, racism and discrimination played and still play in the US history, and, accordingly, the strong traditions of race-based analysis in this country (Bhopal, 2007). In addition, it may be that racial discrimination scales are more frequent in the US due to the considerable scientific production observed in this country. Instruments assessing experiences of racial discrimination are also recent, with the majority of them published in the last 12 years. Additionally, scales mostly emphasized the African American experience, approaching only a few other stigmatized groups. The US prominence in the psychometric assessment of discrimination is paralleled by a growing but still recent body of evidence on discrimination effects on health in countries where this phenomenon is similarly important (Paradies, 2006b).

Most scales made use of racial terminology. The implications of this strategy are, at least, twofold. First, Gomez and Trierweiler (2001) recently demonstrated that making the words race, racism and discrimination salient in the assessment of discrimination has led to a higher frequency of self-reported experiences of

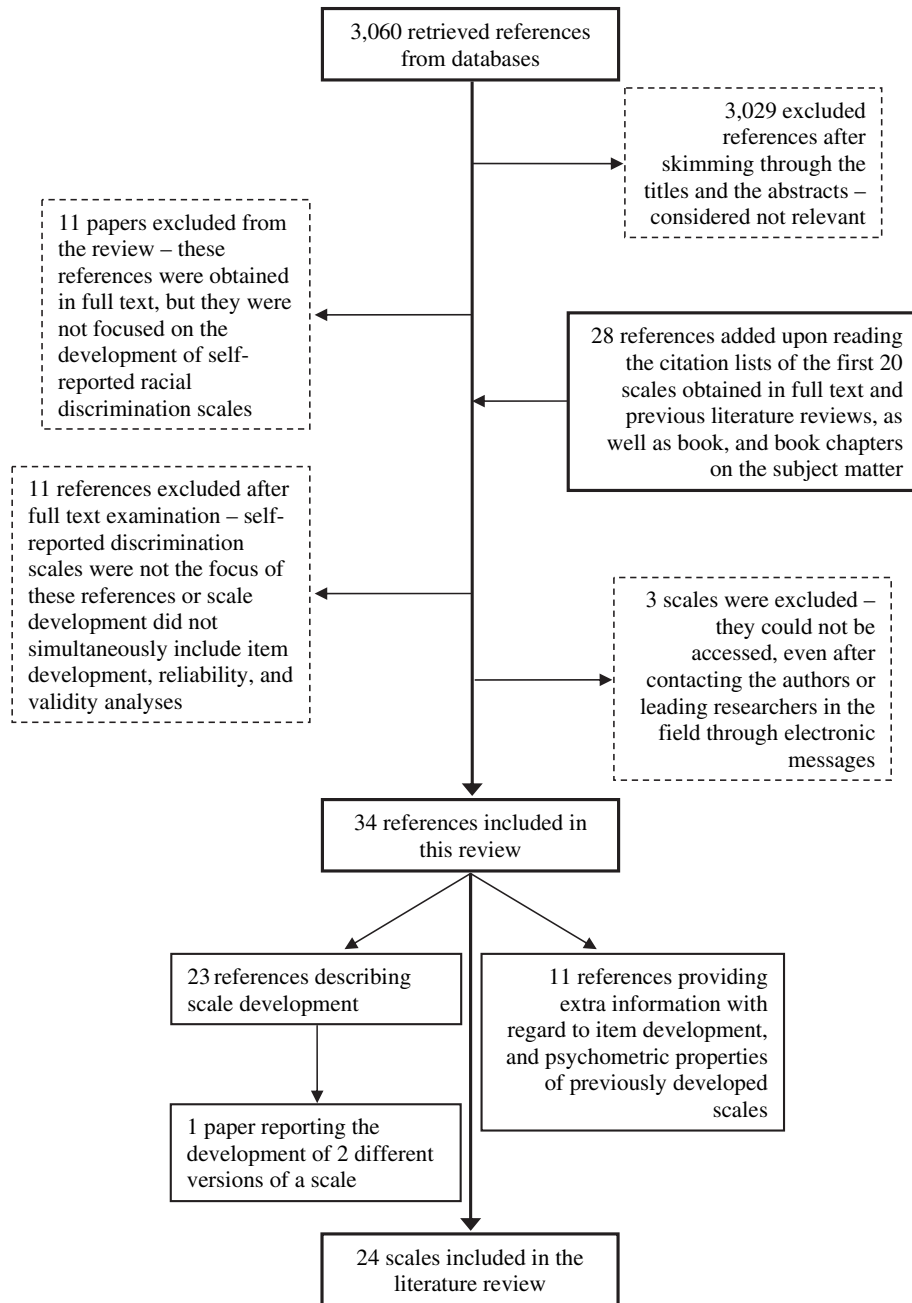


Fig. 1. Review flowchart.

discrimination, compared to the use of neutral terminology. Despite the fact that one of the instruments using neutral terminology, the Everyday Discrimination Scale (Clark, Coleman, & Novak, 2004; Williams, Yu, Jackson, & Anderson, 1997), is among the most widely used scales currently (Paradies, 2006b), the extent to which racial terminology affects validity and reliability of reviewed scales remains unexplored.

The second implication is that scale authors approached respondents as if they could disentangle the multifaceted nature of the discrimination experience. Often, they assumed that respondents clearly distinguished the reason why they were discriminated against. That is, they generally did not express concerns on attributional ambiguity, and asked respondents to report only on racial discrimination, as if this type of discriminatory experience were clearly distinguishable from or

unassociated with gender or age discrimination, to name a few examples. Attributional ambiguity stands as a debated issue in the field (Major, Quinton, & McCoy, 2002; Williams & Mohammed, 2009), such that it should have been discussed in the papers describing the development of the reviewed instruments. Consider, for instance, the case of a black, female, middle-aged respondent whose job opportunity has recently been denied. Assuming this event is interpreted as a consequence of discrimination, would she attribute it to her phenotypic appearance, gender or age? Which of these attributions would be “correct”? What impact a scale using racial terminology would have on her self-reports of experiences of discrimination? Would she have the opportunity to state that she was discriminated against because of her sex, race and age simultaneously? Evidence suggests that the pathogenic effect of discrimination results from

**Table 2**  
Characteristics of the 24 self-reported discrimination scales included in the review.

Characteristic	No. of scales	% of total scales
Year of publication		
1973–1984	1	4.2
1985–1996	7	29.2
1997–2008	16	66.6
Country of origin		
United States of America	23	95.8
Australia	1	4.2
Journal of publication		
Journal of applied social psychology	5	20.8
Ethnicity & Disease	2	8.3
Journal of counseling psychology	2	8.3
PhD thesis	1	4.2
Remaining journals (each with only one published scale)	14	58.4
Sample size (participants)		
0–100	1	4.2
101–200	6	25.0
201–1000	13	54.2
≥1001	4	16.6
Age range of the study population (years) <sup>a</sup>		
0–30	8	40.0
31–45	1	5.0
46–60	8	40.0
≥61	3	15.0
Participants' sex		
Women only	2	8.3
Men only	2	8.3
Women and men	20	83.4
Participants' racial or ethnic group		
Blacks	7	29.2
Asians	2	8.2
Australian aboriginal and Torres Strait Islander	1	4.2
Latinos	1	4.2
Multiple racial/ethnic groups (blacks plus whites, Asians, etc.)	13	54.2
Number of scale items		
≤30	12	50.0
31–60	9	37.5
≥61	3	12.5
Duration of scale completion (min) <sup>a</sup>		
10–15	6	66.7
16–30	3	33.3
Method of data collection		
Self-completed	17	70.8
Face-to-face interview and self-completed	4	16.6
Audio-computer assisted self-interviewing	1	4.2
Mail	1	4.2
Telephone-administered	1	4.2

<sup>a</sup> Age range and duration of scale completion were not presented in four (Borus, Fiman, Stanton, & Dowd, 1973; James, Lovato, & Cropanzano, 1994; Loo et al., 2001; Smedley, Myers, & Harrell, 1993) and fifteen studies (Cardo, 1994; Clark et al., 2004; Collado-Proctor, 1999; Contrada et al., 2001; Fisher, Wallace, & Fenton, 2000; James et al., 1994; Krieger et al., 2005; Landrine & Klonoff, 1996b; Liang, Li, & Kim, 2004; Loo et al., 2001; McNeilly, Anderson, Armstead, et al., 1996; Paradies & Cunningham, 2008; Seaton, 2006; Smedley et al., 1993; Taylor, Kamarck, & Shiffman, 2004), respectively.

generic perceptions of unfair treatment, rather than the perceived reason(s) for such treatment (Kessler, Mickelson, & Williams, 1999).

Nevertheless, much controversy surrounds the issue of whether or not one should use racial terminology. Some authors (Krieger, Smith, Naishadham, Hartman, & Barbeau, 2005) argue that if a scale intends to measure perceived race-based discriminatory experiences, it should employ racial terminology. Evidence also supports the view that reported racial discrimination increases the tendency

**Table 3**  
Methods adopted in the development of the 24 self-reported racial discrimination scales included in the review.

Methods	No. of scales resorting to	% of scales resorting to
Item development		
Literature review	15	62.5
Focus group	6	25.0
Interview/key informant interview	5	20.8
Empirical investigation	5	20.8
Panel of experts	4	16.7
Authors' personal experiences	3	12.5
Other methods <sup>a</sup>	5	20.8
Content validity		
Review by target population through focus groups, pre-tests, pilot studies, and/or interviews	6	25.0
Panel of experts	3	12.5
Literature review	1	4.2
Reliability		
Internal consistency	22	91.7
Test-retest reliability	9	37.5
Split-half reliability	3	12.5
Item-total correlation/inter-item reliability	2	8.3
Inter-observer reliability <sup>b</sup>	–	–
Construct validity		
Convergent validity	16	66.7
Principal components analysis	11	45.8
Extreme group comparisons	9	37.5
Exploratory factor analysis	9	37.5
Confirmatory factor analysis	8	33.3
Discriminant validity	7	29.2
Structural equation modeling	3	12.5

<sup>a</sup> Other methods include analysis of data from previous studies on scale development (Brondolo et al., 2005; Utsey, 1999), assessment of information from the news media (Fisher et al., 2000), and pilot study (Contrada et al., 2001).

<sup>b</sup> Inter-observer reliability theoretically applies only to scales which were, at least partially, administered through face-to-face interviews (Brondolo et al., 2005; Collado-Proctor, 1999; Paradies & Cunningham, 2008) or by telephone (Vines et al., 2001).

to perceive ongoing social interactions as more negative (Broudy et al., 2007), meaning that other types of discrimination may actually be an outcome of race-based discriminatory experiences. In sum, it is still not clear that eliminating racial terminology allows researchers to measure self-reported discrimination more effectively.

Importantly, whether or not one is inclined to use racial terminology, scale developers must also consider that no instrument will be able to fully capture all instances of discriminatory experiences. Apart from experiences that fall outside conscious awareness (Krieger, 2000), Major et al. (2002) reviewed evidence on two phenomena that may affect the reporting of discriminatory experiences: the tendency of stigmatized groups to be highly sensitive to discrimination (vigilance bias), as well as their propensity to minimize the extent to which they are targets of unfair treatment (minimizing bias). Also, there is debate on the inherent subjectivity of experiences of racial discrimination, and the extent to which they reflect "real" experiences. Some authors claim that subjective experiences of discrimination may affect health, regardless of the objectivity of such reporting (Paradies, 2006a), and such distinctions may be more or less applicable to different racial or ethnic groups.

With regard to the development processes and characteristics of reviewed scales, construct mapping and content validation deserve further consideration. The production of a construct map and the formulation of items that represent it comprehensively (content validation) is an essential step during initial development of any instrument (Wilson, 2005). Whenever one intends to measure

**Table 4**  
Ratings for each of the 24 self-reported discrimination scales included in the review.

Scale name	Followed an explicit theoretical framework?	Reported efforts towards content validation?	Reliability scores above 0.70?	At least 75% of the hypotheses regarding relationships with the construct under consideration were confirmed?	Conceptual dimensional structure was supported by means of factor analysis?
Adolescent discrimination distress index (Fisher et al., 2000)	0	0	0	not assessed	+
Asian American racism-related stress inventory (Liang et al., 2004)	+	+	+	+	+
Detroit area study discrimination scale (Taylor et al., 2004; Williams et al., 1997)	+	0	0	0	not assessed
Everyday discrimination scale (Clark et al., 2004; Williams et al., 1997)	+	0	+	+	+
Experiences of discrimination (Krieger, 1990; Krieger et al., 2005)	+	0	+	+	+
General ethnic discrimination scale (Landrine et al., 2006)	+	0	+	+	+
Index of race-related stress – brief version (Utsey, 1999)	+	0	0	+	+
Index of race-related stress for African American adolescents (Seaton, 2003, 2006)	+	0	+	not assessed	+
Index of race-related stress (Utsey & Ponterotto, 1996; Utsey, Ponterotto, Reynolds, & Cancelli, 2000)	+	+	+	+	+
Measure of indigenous racism experiences (Paradies & Cunningham, 2008)	+	+	0	+	+
Minority student stress scale (Smedley et al., 1993)	+	0	+	not assessed	+
Perceived ethnic discrimination questionnaire/ brief version (Brondolo et al., 2005)	+	0	0	+	not assessed
Perceived ethnic discrimination questionnaire/community version (Brondolo et al., 2005)	+	0	+	+	+
Perceived ethnic discrimination questionnaire (Contrada et al., 2001)	+	0	+	+	+
Perceived racism scale for Latina/os (Collado-Proctor, 1999)	+	+	+	+	0
Perceived racism scale (McNeilly, Anderson, Armstead, et al., 1996, McNeilly, Anderson, Robinson, et al., 1996)	+	+	0	+	0
Perceptions of racism scale (Green, 1995)	+	+	+	+	+
Race-Related stressor scale for Asian American Vietnam Veterans (Loo et al., 2001)	+	+	0	+	+
Racial perceptions inventory (Borus et al., 1973)	0	0	+	+	+
Scale for the effects of ethnicity and discrimination (Cardo, 1994; Mirage, 1987, p. 223)	+	0	+	+	not assessed
Scale of ethnic experience (Malcarne, Chavira, Fernandez, & Liu, 2006)	+	0	+	+	+
Schedule of racist events (Klonoff & Landrine, 1999, 2000; Klonoff, Landrine, & Ullman, 1999; Landrine & Klonoff, 1996a, 1996b, 2000)	+	0	+	+	+
Telephone-administered perceived racism scale (Vines et al., 2001)	+	+	0	not assessed	0
Workplace prejudice/discrimination inventory (James et al., 1994)	+	0	+	+	+

+ means a positive rating, and zero, a negative one.

a specific construct, the first thing to do is to theoretically clarify the concept under consideration. Then, one should elaborate a construct map of it, and produce a set of items that is representative of such a map. Finally, the extent to which this set of items reflect the entirety of the mapped construct should be checked by panels of experts, members of the target population, etc., so that content validation can be achieved.

However, this was not made explicit in the majority of the reviewed studies. For instance, due to a lack of conceptual clarity, discrimination, racism and prejudice have been mixed unevenly and so some instruments may have not covered adequately the full discrimination construct. Relatedly, only a few scale authors have discussed personally mediated racial discrimination as to whether it may be conceptualized as a uni- or a multi-dimensional construct. Contrada et al. (2001) as well as Brondolo et al. (2005) have

proposed that discrimination could be assessed according to some of the following dimensions: exclusion/rejection, stigmatization/discrimination, discrimination at work/school, threat/aggression, etc. This is potentially useful and theoretically important, since health consequences of race-based discrimination may vary according to the way it is manifested.

Our analyses also showed that scale developers more likely stressed the need to develop new instruments, instead of adapting pre-existing ones. Among the 24 reviewed instruments, only seven (Brondolo et al., 2005; Clark et al., 2004; Krieger et al., 2005; Landrine et al., 2006; Seaton, 2006; Utsey, 1999) were attempts at developing brief versions of longer scales, refinements and/or editions of items to make them applicable to different population domains. The attempt to develop “universal” instruments that would permit cross-cultural fine-tuning to capture equivalent

aspects of experiences of discrimination among distinct socio-cultural contexts has not been as deeply debated as in other areas of health-related research (Herdman et al., 1998; Hunt & Bhopal, 2004). In fact, instead of adopting an absolutist approach, which assumes that culture has a negligible impact on the construct being measured, or the abovementioned universalist approach, in which adaptations of instruments to different cultural contexts is deemed possible, it appears that scale developers have implicitly followed a relativist approach (Berry et al., 2007). Thus, they probably assumed that the role of culture is substantial and that it is impossible to use standard instruments to assess experiences of racial discrimination across cultures.

In connection, only in a few instances, independent researchers apart from the original scale developers have psychometrically examined these instruments. The literature addressing properties of self-reported discrimination scales would benefit from further evidence produced by independent researchers, confirming or refuting preliminary psychometric data. For example, while all authors utilized the Likert-response format and a handful of them made use of factor analyses, data transformation in order to meet distributional requirements of factor analysis was not used at all. Thus, it remains to be tested whether or not factor analysis findings are consistent with future investigations carried out by independent researchers, incorporating more rigorous analytic techniques.

Of equal significance is the fact that, although some scales assessed experiences of discrimination in conjunction with behavioral coping responses to these experiences (Collado-Proctor, 1999, p. 184; McNeilly, Anderson, Armstead, et al., 1996; Paradies & Cunningham, 2008; Vines et al., 2001), such as getting violent and speaking up (McNeilly, Anderson, Armstead, et al., 1996), none of them aimed at identifying the respondent as a potential perpetrator of discriminatory acts. Violence perpetration and victimization are two constructs commonly assessed in studies on intimate partner and domestic violence. Discriminating against members of other groups may be frequent, and this type of behavior may have health-damaging effects as well. Scales could overcome the perspective of approaching their respondents as if they were only victims of discrimination.

With regard to the methods adopted for data collection, this review showed that most researchers favored the use of self-completed instruments, a trend probably reflecting efforts to avoid interviewer-effects on the process of eliciting sensitive information. Some scales combined self-completion with face-to-face interviews (Brondolo et al., 2005; Collado-Proctor, 1999; Paradies & Cunningham, 2008), but potential interviewer-effects on information disclosure were neither controlled nor discussed in these papers. According to Williams and Mohammed (2009), recent research on race-of-interviewer effects indicated that blacks were reluctant to reveal their true racial beliefs on race sensitive questions, when talking to white interviewers. This means that the social interaction characterized by the interviewer-respondent encounter may have implications for the validity and reliability of scales, such that this issue should be further investigated in future studies.

In conclusion, this study attempted to assist researchers in making a psychometrically informed choice as to which scale mostly fits their needs. The process of decision-making must also be guided by the etiological mechanisms of the outcome under study. Since reviewed scales tap the discrimination construct differently, including its intensity, duration and frequency of exposure, some may be useful for chronic health conditions and others for acute ones. It is clear that there is room for further development of racial discrimination scales. In spite of the acceptable initial psychometric results, existing scales should be put into further scrutiny by independent authors. The idea of

a universal instrument, cross-culturally adaptable to different contexts should be debated among researchers in the field.

## Acknowledgements

The first author was supported by a doctoral scholarship from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) between March 2007 and June 2008 and now holds a doctoral scholarship from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). The second author was supported by a scholarship from CAPES during his PhD studies in the Centre for Health Equity Studies in Stockholm University (Sweden) and now holds a doctoral scholarship from the CNPq. The third and the fourth authors hold grants for productivity in research from the CNPq. We would like to thank Ricardo Ventura Santos for his insightful comments on an earlier version of this manuscript.

## References

- Bastos, J. L., Dumith, S. C., Santos, R. V., Barros, A. J. D., Del Duca, G. F., Gonçalves, H., et al. (2009). Does the way I see you affect the way I see myself? Associations between interviewers' and interviewees' "color/race" in southern Brazil. *Cadernos de Saude Publica*, 25(10), 2111–2124.
- Berry, J. W., Poortinga, Y. H., Segall, M. H., & Dasen, P. R. (2007). *Cross-cultural psychology: Research and applications*. New York: Cambridge University Press.
- Bhopal, R. (2007). *Ethnicity, race, and health in multicultural societies: Foundations for better epidemiology, public health, and health care*. New York: Oxford University Press.
- Blank, R. M., Dabady, M., & Citro, C. F. (2004). *Measuring racial discrimination: Panel on methods for assessing discrimination*. Washington: The National Academies Press.
- Borus, J. F., Fiman, B. G., Stanton, M. D., & Dowd, A. F. (1973). The racial perceptions inventory. *Archives of General Psychiatry*, 29(2), 270–275.
- Brondolo, E., Kelly, K. P., Coakley, V., Gordon, T., Thompson, S., Levy, E., et al. (2005). The perceived ethnic discrimination questionnaire: development and preliminary validation of a community version. *Journal of Applied Social Psychology*, 35(2), 335–365.
- Brondolo, E., Rieppi, R., Kelly, K. P., & Gerin, W. (2003). Perceived racism and blood pressure: a review of the literature and conceptual and methodological critique. *Annals of Behavioral Medicine*, 25(1), 55–65.
- Broudy, R., Brondolo, E., Coakley, V., Brady, N., Cassells, A., Tobin, J. N., et al. (2007). Perceived ethnic discrimination in relation to daily moods and negative social interactions. *Journal of Behavioral Medicine*, 30(1), 31–43.
- Cardo, L. M. (1994). Development of an instrument measuring valence of ethnicity and perception of discrimination. *Journal of Multicultural Counseling & Development*, 22(1), 49–59.
- Clark, R., Coleman, A. P., & Novak, J. D. (2004). Brief report: initial psychometric properties of the everyday discrimination scale in black adolescents. *Journal of Adolescence*, 27(3), 363–368.
- Collado-Proctor, S. M. (1999). *The perceived racism scale for Latina/os: A multidimensional assessment of the experience of racism among Latina/os*. Durham: Duke University.
- Contrada, R. J., Ashmore, R. D., Gary, M. L., Coups, E., Egeth, J. D., Sewell, A., et al. (2001). Measures of ethnicity-related stress: psychometric properties, ethnic group differences, and associations with well-being. *Journal of Applied Social Psychology*, 31(9), 1775–1820.
- DeVellis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks: Sage Publications.
- Dressler, W. W., Oths, K. S., & Gravlee, C. C. (2005). Race and ethnicity in public health research: models to explain health disparities. *Annual Review of Anthropology*, 34, 231–252.
- Fisher, C. B., Wallace, S. A., & Fenton, R. E. (2000). Discrimination distress during adolescence. *Journal of Youth and Adolescence*, 29(6), 679–695.
- Gomez, J. P., & Trierweiler, S. J. (2001). Does discrimination terminology create response bias in questionnaire studies of discrimination? *Personality and Social Psychology Bulletin*, 27(5), 630–638.
- Green, N. L. (1995). Development of the perceptions of racism scale. *Image Journal of Nursing Scholarship*, 27(2), 141–146.
- Harrell, J. P., Hall, S., & Taliaferro, J. (2003). Physiological responses to racism and discrimination: an assessment of the evidence. *American Journal of Public Health*, 93(2), 243–248.
- Harrell, S. P., Merchant, M. A., & Young, S. A. (1997). Psychometric properties of the racism and life experiences scales (RaLES). *Annual convention of the American Psychological Association*, Chicago, IL.
- Herdman, M., Fox-Rushby, J., & Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of Life Research*, 7(4), 323–335.



- Hunt, S. M., & Bhopal, R. (2004). Self report in clinical and epidemiological studies with non-English speakers: the challenge of language and culture. *Journal of Epidemiology and Community Health*, 58(7), 618–622.
- James, K., Lovato, C., & Cropanzano, R. (1994). Correlational and known-group comparison: validation of a workplace prejudice/discrimination inventory. *Journal of Applied Social Psychology*, 24, 1573–1592.
- Jones, C. P. (2000). Levels of racism: a theoretic framework and a gardener's tale. *American Journal of Public Health*, 90(8), 1212–1215.
- Kaufman, J. S. (2008). Epidemiologic analysis of racial/ethnic disparities: some fundamental issues and a cautionary example. *Social Science & Medicine*, 66(8), 1659–1669.
- Kessler, R. C., Mickelson, K. D., & Williams, D. R. (1999). The prevalence, distribution, and mental health correlates of perceived discrimination in the United States. *Journal of Health and Social Behavior*, 40(3), 208–230.
- Klonoff, E. A., & Landrine, H. (1999). Cross-validation of the schedule of racist events. *Journal of Black Psychology*, 25(2), 231–254.
- Klonoff, E. A., & Landrine, H. (2000). Is skin color a marker for racial discrimination? Explaining the skin color-hypertension relationship. *Journal of Behavioral Medicine*, 23(4), 329–338.
- Klonoff, E. A., Landrine, H., & Ullman, J. B. (1999). Racial discrimination and psychiatric symptoms among blacks. *Cultural Diversity and Ethnic Minority Psychology*, 5(4), 329–339.
- Kressin, N. R., Raymond, K. L., & Manze, M. (2008). Perceptions of race/ethnicity-based discrimination: a review of measures and evaluation of their usefulness for the health care setting. *Journal of Health Care for the Poor and Underserved*, 19(3), 697–730.
- Krieger, N. (1990). Racial and gender discrimination: risk factors for high blood pressure? *Social Science & Medicine*, 30(12), 1273–1281.
- Krieger, N. (2000). Discrimination and health. In L. F. Berkman, & I. Kawachi (Eds.), *Social epidemiology* (pp. 36–75). New York: Oxford University Press.
- Krieger, N. (2001). A glossary for social epidemiology. *Journal of Epidemiology and Community Health*, 55(10), 693–700.
- Krieger, N. (2005a). Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination. In N. Krieger (Ed.), *Embodying inequality: Epidemiologic perspectives* (pp. 101–158). Amityville: Baywood Publishing Company.
- Krieger, N. (2005b). Stormy weather: race, gene expression, and the science of health disparities. *American Journal of Public Health*, 95(12), 2155–2160.
- Krieger, N., Smith, K., Naishadham, D., Hartman, C., & Barbeau, E. M. (2005). Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Social Science & Medicine*, 61(7), 1576–1596.
- Landrine, H., & Klonoff, E. A. (1996a). Acculturation and physical health. In H. Landrine, & E. A. Klonoff (Eds.), *African American acculturation: Deconstructing race and reviving culture* (pp. 89–125). Thousand Oaks: SAGE.
- Landrine, H., & Klonoff, E. A. (1996b). The schedule of racist events: a measure of racial discrimination and a study of its negative physical and mental health consequences. *Journal of Black Psychology*, 22(2), 144–168.
- Landrine, H., & Klonoff, E. A. (2000). Racial discrimination and cigarette smoking among blacks: findings from two studies. *Ethnicity & Disease*, 10(2), 195–202.
- Landrine, H., Klonoff, E. A., Corral, I., Fernandez, S., & Roesch, S. (2006). Conceptualizing and measuring ethnic discrimination in health research. *Journal of Behavioral Medicine*, 29(1), 79–94.
- Lang, D. L. (2001). *The schedule of racist events-generic: A measure of perceived racism across ethnic groups*. Loma Linda: Loma Linda University.
- Lee, C. (2009). "Race" and "ethnicity" in biomedical research: how do scientists construct and explain differences in health? *Social Science & Medicine*, 68(6), 1183–1190.
- Liang, C. T. H., Li, L. C., & Kim, B. S. K. (2004). The Asian American racism-related stress inventory: development, factor analysis, reliability, and validity. *Journal of Counseling Psychology*, 51(1), 103–114.
- Loo, C. M., Fairbank, J. A., Scurfield, R. M., Ruch, L. O., King, D. W., Adams, L. J., et al. (2001). Measuring exposure to racism: development and validation of a race-related Stressor scale (RRSS) for Asian American Vietnam veterans. *Psychological Assessment*, 13(4), 503–520.
- Major, B., Quinton, W. J., & McCoy, S. K. (2002). Antecedents and consequences of attributions to discrimination: theoretical and empirical advances. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 251–330). San Diego: Academic Press.
- Malcarne, V. L., Chavira, D. A., Fernandez, S., & Liu, P. J. (2006). The scale of ethnic experience: development and psychometric properties. *Journal of Personality Assessment*, 86(2), 150–161.
- McNeilly, M. D., Anderson, N. B., Armstead, C. A., Clark, R., Corbett, M., Robinson, E. L., et al. (1996a). The perceived racism scale: a multidimensional assessment of the experience of white racism among African Americans. *Ethnicity & Disease*, 6(1–2), 154–166.
- McNeilly, M. D., Anderson, N. B., Robinson, E. L., McManus, C. H., Armstead, C. A., Clark, R., et al. (1996b). Convergent, discriminant, and concurrent validity of the perceived racism scale: a multidimensional assessment of the experience of racism among African Americans. In R. L. Jones (Ed.), *Handbook of tests and measurements for Black populations* (pp. 359–374). Hampton: Cobb & Henry Publishers.
- Mirage, L. W. (1987). *Valence of ethnicity, perception of discrimination and self-esteem in high risk minority college students*. New York: Fordham University.
- Paradies, Y. (2006a). Defining, conceptualizing and characterizing racism in health research. *Critical Public Health*, 16(2), 143–157.
- Paradies, Y. (2006b). A systematic review of empirical research on self-reported racism and health. *International Journal of Epidemiology*, 35(4), 888–901.
- Paradies, Y., & Cunningham, J. (2008). Development and validation of the measure of Indigenous racism experiences (MIRE). *International Journal for Equity in Health*, 7, 9.
- Pascoe, E. A., & Smart Richman, L. (2009). Perceived discrimination and health: a meta-analytic review. *Psychological Bulletin*, 135(4), 531–554.
- Penner, A. M., & Saperstein, A. (2008). How social status shapes race. *Proceedings of National Academy of Sciences U S A*, 105(50), 19628–19630.
- Seaton, E. K. (2003). An examination of the factor structure of the index of race-related stress among a sample of African American adolescents. *Journal of Black Psychology*, 29(3), 292–307.
- Seaton, E. K. (2006). Examination of a measure of racial discrimination among African American adolescents. *Journal of Applied Social Psychology*, 36(6), 1414–1429.
- Smedley, B. D., Myers, H. F., & Harrell, S. P. (1993). Minority-status stresses and the college adjustment of ethnic minority freshmen. *Journal of Higher Education*, 64(4), 434–452.
- Streiner, D. L., & Norman, G. R. (1998). *Health measurement scales: A practical guide to their development and use*. New York: Oxford Medical Publications.
- Taylor, T. R., Kamarck, T. W., & Shiffman, S. (2004). Validation of the Detroit area study discrimination scale in a community sample of older African American adults: the Pittsburgh healthy heart project. *International Journal of Behavioral Medicine*, 11(2), 88–94.
- Terrell, F., & Miller, F. S. (1980). The development of an inventory to measure experience with racialistic incidents among Blacks, unpublished manuscript.
- Travassos, C., & Williams, D. R. (2004). The concept and measurement of race and their relationship to public health: a review focused on Brazil and the United States. *Cadernos de Saude Publica*, 20(3), 660–678.
- Utsey, S. O. (1998). Assessing the stressful effects of racism: a review of instrumentation. *Journal of Black Psychology*, 24(3), 269–288.
- Utsey, S. O. (1999). Development and validation of a short form of the Index of Race-Related Stress (IRRS) – brief version. *Measurement and Evaluation in Counseling and Development*, 32(3), 149–167.
- Utsey, S. O., & Ponterotto, J. G. (1996). Development and validation of the index of race-related stress (IRRS). *Journal of Counseling Psychology*, 43(4), 490–501.
- Utsey, S. O., Ponterotto, J. G., Reynolds, A. L., & Cancelli, A. A. (2000). Racial discrimination, coping, life satisfaction, and self-esteem among African Americans. *Journal of Counseling and Development*, 78(1), 72–80.
- Vines, A. I., McNeilly, M. D., Stevens, J., Hertz-Picciotto, I., Baird, M., & Baird, D. D. (2001). Development and reliability of a telephone-administered perceived racism scale (TPRS): a tool for epidemiological use. *Ethnicity & Disease*, 11(2), 251–262.
- Williams, D. R., & Mohammed, S. A. (2009). Discrimination and racial disparities in health: evidence and needed research. *Journal of Behavioral Medicine*, 32(1), 20–47.
- Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: findings from community studies. *American Journal of Public Health*, 93(2), 200–208.
- Williams, D. R., Yu, Y., Jackson, J. S., & Anderson, N. B. (1997). Racial differences in physical and mental health: socio-economic status, stress and discrimination. *Journal of Health Psychology*, 2, 335–351.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah: Lawrence Erlbaum Associates.
- Wyatt, S. B., Williams, D. R., Calvin, R., Henderson, F. C., Walker, E. R., & Winters, K. (2003). Racism and cardiovascular disease in African Americans. *American Journal of Medical Sciences*, 325(6), 315–331.