



Discrimination and Romani health: a validation study of discrimination scales among Romani women in Macedonia and Serbia

T. Janevic · D. Gundersen · K. Stojanovski · J. Jankovic ·
Z. Nikolic · B. Kasapinov

Received: 2 March 2015 / Revised: 26 June 2015 / Accepted: 1 July 2015 / Published online: 15 July 2015
© Swiss School of Public Health 2015

Abstract

Objectives Scales used to assess discrimination in public health research have rarely been validated outside of high income countries. Our objective was to validate the Experiences of Discrimination (EOD) scale and the Everyday Discrimination Scale (EDS) among 410 Romani women in Macedonia and Serbia.

Methods Romani female interviewers conducted interviews in 2012–2013. We used a multiple indicator multiple cause approach to test a one-factor model for each scale and to assess differential item functioning (DIF) by age, wealth, country, and education. We also measured

associations between the EOD and EDS with smoking in the past year and psychological distress.

Results Three items of the EOD were conceptually irrelevant. Two items of the EDS were not conditionally independent. DIF was found by country for one item in each scale. After excluding these items, all scales exhibited good model fit and were associated with smoking (EOD beta = 0.40, 95 % CI = 0.18, 0.63; EDS beta = 0.33, 95 % CI = 0.12, 0.54) and psychological distress (EOD beta = 0.26, 95 % CI = 0.15, 0.37; EDS beta = 0.26, 95 % CI = 0.04, 0.47).

Conclusions Discrimination scales can be adapted for use among Romani women and are associated with both smoking and psychological distress.

T. Janevic (✉)
Department of Epidemiology, Rutgers School of Public Health,
683 Hoes Lane West, Room 209, PO Box 9, Piscataway,
NJ 08854, USA
e-mail: Teresa.Janevic@Rutgers.edu

D. Gundersen
Department of Family Medicine and Community Health,
Rutgers University-Robert Wood Johnson Medical School,
New Brunswick, NJ, USA
e-mail: Daniel.gundersen@rwjms.rutgers.edu

K. Stojanovski
Studiorum, Skopje, Macedonia
e-mail: kristefers@gmail.com

J. Jankovic · Z. Nikolic
Faculty of Medicine, Institute of Social Medicine, University of
Belgrade, Belgrade, Serbia
e-mail: drjankojankovic@yahoo.com

B. Kasapinov
Institute of Social Medicine, Saints Cyril and Methodius
University, Skopje, Macedonia
e-mail: bkasapinov@gmail.com

Keywords Discrimination · Validity · Smoking ·
Mental health · Minority health

Introduction

Racial discrimination is increasingly recognized as an important social determinant of health (Krieger 2012; Pascoe and Smart Richman 2009; Williams and Mohammed 2009). Health researchers have developed scales to measure perceived discrimination, and these scales have been used in a growing volume of epidemiologic literature. However, existing scales, many developed for the African American population in the USA, have not been psychometrically evaluated in populations outside of the USA, particularly in low- or middle-income country settings, with a few exceptions (Bastos et al. 2012; Paradies and Cunningham 2008; Paradies 2006). Roma, as the largest ethnic minority group in Europe, experiences widespread social exclusion and poor health, yet few researchers have

measured discrimination in this population (Cook et al. 2013). A psychometric evaluation of measures of perceived discrimination is needed to advance research on Romani health and is a useful case study for methods to apply US-developed measures of discrimination internationally.

Serbia and Macedonia have the largest Romani population outside of the European Union countries, as well as in the former Yugoslavia (Crowe 2007). Roma have historically experienced discrimination and social exclusion in Southeastern Europe, although contextual influences have varied from victimization during the Holocaust, increased access to employment during the Communist era, a post-communist rise in hate crimes, and finally internal displacement during war and conflict in the 1990s (Crowe 2007). Recent survey research has shown that Roma are the most hated ethnic group in Europe (Pew Research Center 2009). These prejudices were studied in more detail among Serbian adolescents, who perceived Roma as a threat due to cultural differences (Ljubic et al. 2013). Romani women in particular are subject to discrimination due to a low level of education and patriarchal social norms (Bond 2004; Ravnol 2010) and have been found in Serbia to have poorer self-rated health than non-Roma and Romani men (Janevic et al. 2011a, b). Although the influence of discrimination on the health of Roma has been a topic of limited qualitative research and reports by non-governmental organizations (Abdikeeva et al. 2013), few studies have quantitatively assessed experiences of discrimination among Roma. Measuring and studying discrimination in diverse global contexts is vital to understanding global inequalities in health and has the potential to advance causal inference on how discrimination influences health.

Two measures of discrimination most commonly utilized in health research are the Experiences of Discrimination Scale (Krieger et al. 2005) (EOD) and the Everyday Discrimination Scale (Williams et al. 1997) (EDS). The EOD and EDS have demonstrated high construct validity among African Americans, Latinos, Asians, and whites in the USA (Chan et al. 2012; Cunningham et al. 2011; Krieger et al. 2005; Lewis et al. 2012; Shariff-Marco et al. 2011; Stucky et al. 2011). The validity of these scales in diverse ethnic populations suggests that they may be appropriate for use in ethnic minority groups in other countries. In particular, the scales may be useful in the Romani population, which one sociologist has directly compared to the inner-city African American population as a structural underclass (Szelenyi and Ladanyi 2006).

The objective of our study was to assess the validity of the EOD and EDS in Romani women in Serbia and Macedonia. Specifically, we assessed the construct validity of both scales and assessed differential item functioning (DIF) using confirmatory factor analysis with covariates [also known as multiple indicator multiple cause (MIMIC)

models]. We also measured associations between discrimination scores and two health outcomes most frequently associated with self-reported discrimination in previous research, smoking and psychological distress (Paradies 2006; Williams and Mohammed 2009).

Methods

Study design

We collected data on 410 women in the countries of Serbia and Macedonia between October 2012 and April 2013. Eligible women had given birth in the 2 years preceding the interview date. We used a purposeful snowball sampling technique based on a previous qualitative study conducted in eight cities in Serbia and Macedonia (Janevic et al. 2011a, b). Romani women interviewers chose seed respondents in Romani settlements based on word of mouth. For example, they would enter the settlement and ask the first people they saw if there was a woman nearby who had a baby. Seed participants referred interviewers to other eligible women (gave birth in previous 2 years) in the settlement, thus creating a snowball sample. In the capital cities, we asked NGOs to choose settlements of low, middle and high level of development to ensure variation in individual and neighborhood socioeconomic status in the sample. In the smaller cities, where there are fewer settlements, most Romani neighborhoods were included in the sample. Face-to-face interviews were conducted in Serbian and Macedonian. The study was approved by the Rutgers Biomedical Health Sciences Institutional Review Board.

Assessment of perceived discrimination

University collaborators translated the EOD and EDS into Serbian and Macedonian. We then conducted in-person meetings in five cities with Romani women from five local NGOs to discuss the relevance of the instruments to their communities and to adapt the language to be appropriate for women of all educational levels. The order in which the EOD and EDS were presented in the survey was alternated randomly to prevent order bias.

The EOD is a series of nine items collecting information on self-reported occurrences of discrimination, including frequency of the events, and responses to unfair treatment. The respondent is asked, "Have you ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior in any of the following situations because of your race, ethnicity or color?", and then must both recall the occurrence and attribute it to their race/ethnicity at once (Shariff-Marco et al. 2011). For each situation to which the respondent

says yes, they are asked, “How many times did this happen?”. We made several changes to EOD items. First, the word “race” was removed from the question, because this word is not commonly used in the Romani community. Second, we removed “getting credit, bank loans or a mortgage” and modified “Getting service at a store or restaurant” to read “Getting service at a store” due to lack of relevance in Romani women’s lives. Following previous research, we used the EOD items both as a yes/no dichotomous variable (situation score) and a numeral representing the frequency of occurrences (frequency score) (Krieger et al. 2005).

The EDS measures differential treatment in daily life and conceptualizes discrimination as a chronic daily stressor (Williams et al. 1997). The nine items in the EDS ask more generally about unfair treatment, such as “You have been treated with less respect than other people”, and then asks the respondent to attribute the unfair treatment to a list of reasons. For item analysis of the EDS, we included all positive responses to discrimination regardless of attribution because “your ethnicity” was by far the most common attributions for all positive responses (range 93–100 %). Also, the Romani interviewers felt that the respondents could not distinguish easily between the items on the list of reasons due to a lack of distinct identities based on ethnicity, gender and religion. This is coherent with the theory of intersectionality, which suggests that the gender, ethnicity and poverty status of Romani women together form an identity making them vulnerable to discrimination (Mullings and Schulz 2005).

Assessment of covariates

We collected information on socio-demographic characteristics by self-report on the questionnaire. We created a wealth index by replicating the inventory of household items from the Serbia and Macedonia UNICEF Multiple Indicator Cluster Survey (UNICEF 2007). The index was created by performing a principal component analysis (PCA) of the household items with one factor a priori determined.

Health outcomes

We asked women about smoking in the past year (yes/no). Psychological distress was measured using the Kessler-6 (K6) (Kessler et al. 2003). The K6 is a common measure of psychological distress, asking, “During the past 30 days, about how often did you feel...” for six items: “nervous”, “hopeless”, “restless or fidgety”, “so depressed that nothing could cheer you up”, “that everything was an effort” and “worthless”. Response options were on

a five-point likert scale. A similar procedure of translation and community feedback was followed for the K6 as described previously for the discrimination scales. Items were summed to represent a single score of psychological distress.

Statistical analysis

Evaluations of the discrimination scales were done in two phases. First, frequencies for each item in the EOD and EDS were inspected to look for high rates of item non-response. Items with high non-response were flagged and examined for conceptual relevance. Second, a confirmatory factor analysis with covariates—or multiple indicator multiple cause (MIMIC)—approach was used to test a one-factor model for each discrimination scale and assess for DIF by four respondent characteristics—age, wealth, country of respondent and education. A MIMIC model is a measurement model where a latent construct is estimated by several manifest variables (i.e., “indicators”) and is regressed on one or more covariates. DIF examines whether there are group differences in the respective indicators after adjusting for the latent trait. A substantial and significant DIF in an indicator may be due to different measurement properties, which in turn may bias subsequent analyses of the construct. As such, it can be advantageous to remove indicators that exhibit substantial DIF. All models were fit in MPlus version 7.2 using a nonlinear CFA with probit link and means and variance adjusted weighted least square (WLSMV) estimator. The WLSMV as implemented in MPlus applies a weighted least squares approach to a maximum-likelihood estimated polychoric correlation matrix. All available data was used for estimating each pairwise correlation, which results in WLSMV estimation that is consistent when missing values are present in the indicator variables. Model fit was assessed with Chi-square comparative fit index (CFI), and root mean square error of approximation (RMSEA) with conventional criteria for determining fit (Hooper et al. 2008).

DIF was assessed by examining modification indices for each discrimination indicator on each respondent characteristic. An indicator was considered to have significant DIF if the modification index was greater than 13. This corresponds to a Šidák corrected $\alpha = 0.05$ adjusting for the 55 DIF tests performed.

The above approach was used for the EOD situation and frequency scales. However, consistent with previous research, the EDS had poor model fit for a one-factor CFA which was due to substantial residual covariance (i.e., they were not conditionally independent) among the indicator variables (Stucky et al. 2011). As such, we followed the approach of Stucky et al. and performed an exploratory

factor analysis (EFA) with oblimin rotation to identify variables that were not conditionally independent in our sample (Stucky et al. 2011). We selected the indicator from each set of variables that exhibited residual covariance with the highest standardized factor loading for inclusion in the revised EDS. This resulted in a six-item EDS that included ‘Treated with less respect’, ‘You have been threatened or harassed’, ‘Received poorer service at stores’, ‘People...think you are not smart’, ‘People ...think you are dishonest,’ and ‘People...acted as if they’re better than you’. This six-item scale was then subjected to the MIMIC approach described above.

Lastly, we assessed the criterion validity of each discrimination scale by regressing the latent score for each on the K6 psychological distress scale and on smoking in the last year (yes vs. no). We present both unadjusted and adjusted coefficients because psychological distress may be on the causal pathway from discrimination to smoking, and a formal mediation analysis is beyond the scope of this paper (Purnell et al. 2012). We calculated the standardized coefficient for the K6, which can be interpreted as a standard deviation change in discrimination scores for a standard deviation increase in psychological distress. For smoking, we present the coefficient for the standardized latent discrimination scores, which can be interpreted as a standard deviation change in discrimination for those who smoked relative to those who did not smoke in the past year.

Results

Item analysis

In Table 1, we present the sample characteristics. The completion rate was not documented, but interviewers reported qualitatively near complete participation (>90 %), consistent with other survey research of Roma in Serbia and Macedonia (Multiple Indicator Cluster Survey Serbia 2011; Multiple Indicator Cluster Survey Macedonia 2012). The samples from Macedonia and Serbia were similar on all variables except for education—women from Macedonia had less education than those from Serbia (Table 1).

For the EOD, ‘at work’ (15.3 %), ‘getting hired or getting a job’ (15.1 %), and ‘getting housing’ (16.1 %) were outliers for item non-response (Table 2). Non-response rate for the other items ranged from 3.6 % for ‘Getting service in a store’ to 8.3 % for ‘From the police’. The situation for which the highest proportion of respondents reported experiencing discrimination was ‘On the street or in a public setting’ (42.8 %), followed by ‘Getting medical care’ (39.7 %) and ‘At school’ (35.7 %). No item in the EDS had unusually large rates of missing values. The events most frequently reported were ‘Being treated with less courtesy’ (54.5 %), ‘Being treated with less respect’ (54.3 %) and ‘Being threatened or harassed’ (49.1 %) (Table 2). Least reported was ‘People think that you are not smart’ (10.5 %).

Table 1 Socio-demographic characteristics of Romani women, Macedonia and Serbia, 2012–2013

Characteristic	Overall		Macedonia		Serbia	
	Mean or %	SD or <i>n</i>	Mean or %	SD or <i>n</i>	Mean or %	SD or <i>n</i>
Age in years	24.4	5.4	25.0	5.7	23.8	5.0
Wealth quintile	2.0	1.4	1.9	1.4	2.0	1.4
Education						
No school	25.7	105	36.2	72	15.7	33
Primary	58.7	240	52.8	105	64.3	135
Secondary or higher	15.6	64	11.1	22	20.0	42
Marital status						
Married	47.9	196	48.2	96	47.6	100
Living with man	47.4	194	46.2	92	48.6	102
Single/do not know	4.7	19	5.5	11	3.8	8
City type						
Town/rural	41.5	170	40.0	80	42.9	90
Urban	58.5	240	60.0	120	57.1	120
Country						
Macedonia	48.8	200				
Serbia	51.2	210				

Table 2 Frequencies of Experiences of Discrimination Scale and Everyday Discrimination Scale among Romani women, Macedonia and Serbia, 2012–2013

Scale	Never (0) %	Once (1) %	Two or three (2) %	Four or more (3) %	Missing %
Experiences of discrimination					
Have you ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior in any of the following situations because of your ethnicity or color (because you are Roma)?					
At school ^{a,b}	58.2	15.3	8.5	11.9	6.1
At work	71.8	5.4	3.4	4.1	15.3
Getting medical care ^{a,b}	54.5	14.1	14.4	11.2	5.8
On the street or in a public setting ^{a,b}	51.6	11.9	15.8	15.1	5.6
Getting hired or getting a job	68.1	6.6	4.9	5.4	15.1
Getting housing	75.9	5.1	1.0	1.9	16.1
Getting service in a store ^{a,b}	65.9	12.9	9.5	8.0	3.6
From the police or in court ^b	83.9	5.8	1.2	0.7	8.3
Everyday discrimination					
In your day-to-day life, how often have any of the following things happened to you?					
Treated with less courtesy	44.8	11.2	22.1	21.2	0.7
Treated with less respect ^c	45.0	13.9	20.2	20.2	0.7
Received poorer service at stores ^c	72.5	6.3	8.3	11.4	1.5
People...think you are not smart ^c	89.5	1.5	3.4	4.9	0.7
People are afraid of you	63.3	11.4	12.9	11.7	0.7
People think you are dishonest	68.1	9.2	8.0	12.9	1.7
People acted as if they are better than you ^c	81.5	7.3	2.9	7.3	1.0
You have been called names or insulted	68.9	10.2	9.0	10.9	1.0
You have been threatened or harassed ^c	50.4	10.9	16.1	22.1	0.5

^a Item retained in final Experiences of Discrimination Scale scored by number of situations

^b Item retained in final Experiences of Discrimination Scale scored by frequency

^c Item retained in final Everyday Discrimination Scale

MIMIC model

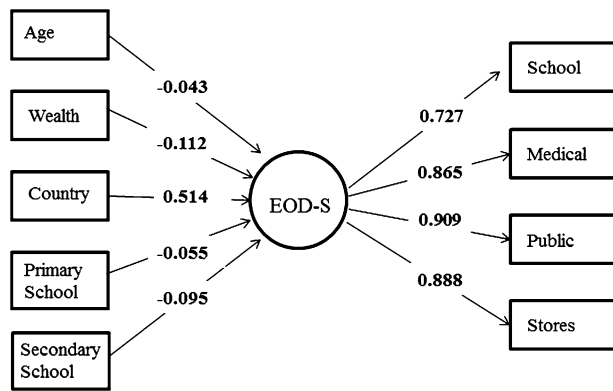
Figures 1 and 2 present the final models for the EOD situation and frequency scales. We excluded the three items with high non-response in the initial analyses of the EOD, with the rationale that the high non-response indicated an inappropriateness of these items in the study population due to the exclusion of Romani women from the formal labor and housing markets. The final model for the EOD situation score additionally excluded ‘from the police or in court’ due to significant DIF by country (modification index = 13.58) (Fig. 1). It had excellent model fit as indicated by the comparison of the fitted model to a saturated model [Chi-square (17) = 25.84, $P = 0.077$], as well as the root mean squared error of approximation (RMSEA = 0.036) and comparative fit index (CFI = 0.987). The standardized factor loadings ranged from 0.727 for ‘At school’ to 0.909 for ‘On the street or in a public setting’. For the EOD frequency scale, no modification indices

over 13 were observed and the standardized factor loadings ranged from 0.529 for police to 0.880 for public (Fig. 2). The model had good fit per the fit indices [Chi-square (25) = 43.97, $P = 0.011$; RMSEA = 0.043; CFI = 0.984].

Figure 3 presents the final model for the EDS. This model excluded ‘People...think you are dishonest’ as an indicator due to significant DIF by country (modification index = 16.37). The final model showed good model fit [Chi-square (25) = 41.09, $P = 0.022$; RMSEA = 0.044; CFI = 0.993] with factor loadings ranging from 0.796 for ‘You have been threatened or harassed’ to 0.896 for ‘Treated with less respect’.

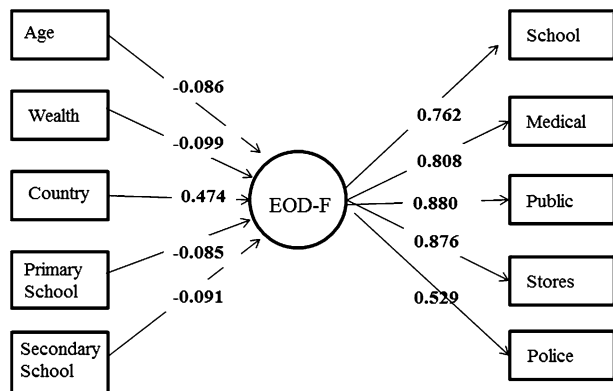
Criterion validity

Smoking prevalence and psychological distress score were somewhat higher in Serbia than in Macedonia (Table 3). The Kessler Stress Scale was positively associated with the EOD situation, EOD frequency and the EDS (Table 4). Women



Abbreviations: EOD-S = Experiences of Discrimination Situation Scale
All coefficients are standardized
Model Fit: Chi-squared=25.84, degrees of freedom=17, $p=0.077$;
Root Mean Square Error of Approximation=0.036; Comparative Fit Index=0.987

Fig. 1 Experiences of Discrimination Situation Scale, Final Multiple Indicator Multiple Cause Model, Serbia and Macedonia, 2012–2013



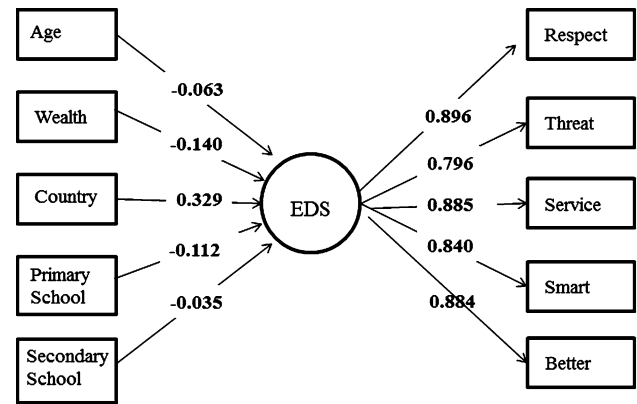
Abbreviations: EOD-F = Experiences of Discrimination Frequency Scale
All coefficients are standardized
Model Fit: Chi-squared=43.87, degrees of freedom=25, $p=0.011$;
Root Mean Square Error of Approximation=0.043; Comparative Fit Index=0.984

Fig. 2 Experiences of Discrimination Frequency Scale, Final Multiple Indicator Multiple Cause Model, Serbia and Macedonia, 2012–2013

who smoked in the past year had significantly higher scores on the EOD situation and EDS (Table 4). Smoking in the past year was not associated with EOD frequency.

Discussion

Our findings contribute new evidence regarding the use of discrimination measures designed in the USA among Romani women. We found that a reduced-item version of both the EOD and EDS provided a one-factor solution representing a single dimension of discrimination. DIF analyses demonstrated that the scale items performed similarly across



Abbreviations: EDS = Everyday Discrimination Scale
All coefficients are standardized
Model Fit: Chi-squared=41.09, degrees of freedom=25, $p=0.022$; Root Mean Square Error of Approximation=0.044; Comparative Fit Index=0.993

Fig. 3 Everyday Discrimination Scale, Final Multiple Indicator Multiple Cause Model, Serbia and Macedonia, 2012–2013

Table 3 Mean psychological distress score and smoking frequency, Macedonia and Serbia, 2012–2013

Country	Kessler-psychological distress		Smoked in past year (%)
	Mean	Standard deviation	
Macedonia	14.6	7.1	39.7
Serbia	15.1	6.9	47.6
Total	14.9	7.0	43.7

education, socioeconomic and age groups, but one item in each scale functioned differently between countries. In an assessment of criterion validity, higher scores on the EOD and EDS were associated with higher prevalence of smoking and psychological distress.

Ours is the first psychometric evaluation of measures of perceived discrimination among Roma, and the second in a middle-income country. Bastos et al. developed and evaluated a scale measuring differential treatment in Brazil (Bastos et al. 2012). A difference between the approach of Bastos et al. and our approach is that they developed a new scale, while we adapted existing scales. Both approaches resulted in measures with good validity and could be considered in international public health research.

The results of our adaption of the EOD and EDS can be compared and contrasted to evaluations of the same scales in the USA. Similar to two previous analyses of the EOD in the USA, we did not find DIF by age or education (Cunningham et al. 2011). However, we did find DIF by country for the item 'From the police'. One important difference between our findings regarding the EOD scale and previous research was high item non-response for three items: 'at

Table 4 Associations between perceived discrimination, psychological distress and smoking, Romani women, Macedonia and Serbia, 2012–2013*Adjusted mutually adjusted for psychological distress and smoking*^a *B* standardized regression coefficients^b *B* standard deviation change in discrimination for discrete change in smoking

Scale	Kessler-psychological distress ^a		Smoked in past year ^b	
	<i>B</i>	95 % CI	<i>B</i>	95 % CI
EOD-S				
Unadjusted	0.258	(0.147, 0.369)	0.403	(0.179, 0.627)
Adjusted	0.218	(0.103, 0.333)	0.295	(0.063, 0.527)
EOD-F				
Unadjusted	0.237	(0.126, 0.349)	0.287	(0.064, 0.510)
Adjusted	0.212	(0.098, 0.327)	0.187	(−0.043, 0.416)
EDS				
Unadjusted	0.257	(0.039, 0.474)	0.330	(0.118, 0.542)
Adjusted	0.158	(0.047, 0.270)	0.195	(0.087, 0.304)

work’, ‘getting a job’ and ‘getting housing’, most likely due to the inappropriateness of these items in this population. Romani women are excluded so severely from the employment and housing market and therefore might not have the opportunity to experience these discriminatory situations. It was for the same reason we had the item prior to administration of the survey, “getting credit, bank loans or a mortgage”. This demonstrates a characteristic of the EOD scale that the respondent must participate in a certain activity to experience a discriminatory event and may make it less appropriate for socially excluded populations. Instead, more specific measures related to work in the gray economy could be designed. Also, measures of institutional discrimination regarding exclusion from the employment and housing markets may be more appropriate to capture the true experience of marginalized women in these dimensions.

Our finding that a reduced-item version of the EDS performed better was similar to previous research in the USA (Chan et al. 2012; Kim et al. 2014; Reeve et al. 2011; Stucky et al. 2011). However, two other studies found a single-factor solution using the complete scale, although one of these two studies had deleted the ‘courtesy’ item a priori due to potential redundancy (Lewis et al. 2012; Shariff-Marco et al. 2011). Most studies that have performed DIF analysis on the EDS did so in diverse samples with the primary explanatory variable as race/ethnicity, to determine if the EDS items performed similarly between ethnic groups (Lewis et al. 2012; Reeve et al. 2011; Shariff-Marco et al. 2011). Our approach was similar to Stucky et al. which examined DIF of the EDS within African Americans by gender (Stucky et al. 2011). No previous studies to our knowledge have measured DIF of the EDS by age or socioeconomic status. We were particularly interested in socioeconomic status due to concerns that the items may not function well among illiterate women. However, the present analysis found no items exhibiting DIF by age, education or wealth, but ‘Dishonest’ did exhibit significant DIF by country. Thus, our finding builds on previous analyses of the EDS by demonstrating that

a reduced-item version functioned well among Romani women across socio-demographic groups.

The EOD and EDS represent two common approaches to the measurement of personally mediated discrimination: one stage and two stage (Lewis et al. 2015; Shariff-Marco et al. 2011). The one-stage approach used in the EOD includes in the question attribution to race/ethnicity (in this case “because of your race/ethnicity”, “because you are Roma”), whereas the two-stage approach used in the EDS asks first about an experience of unfair treatment and then in a follow-up question assesses attribution to race/ethnicity and other characteristics. Differences between these approaches are the subject of an ongoing debate (Lewis et al. 2015). In previous research, the one-stage approach most often returned higher frequencies of reports of discrimination specific to race/ethnicity, but the two-stage approach returned higher frequencies of discrimination due to any attribution (Lewis et al. 2015). The latter was also the case in our study (Table 1). Because of the marginal position of Romani women in society due to ethnicity, gender, poverty and in some cases refugee or internally displaced person status, the two-stage approach which captures all attributions of discrimination may be an advantage of use of the EDS among Romani women.

The associations we found between our reduced-item scales with psychological distress and smoking are consistent with previous literature conducted in ethnic minority groups, including outside of the USA (Harris et al. 2012; Ikram et al. 2014). A weakness of the analysis of the association of discrimination with mental health is that we did not perform psychometric analyses on the Kessler-6, which itself has not been validated in the Romani population. However, we would not expect measurement error in the K6 to be differential with respect to discrimination. As such, measurement error in the K6 would bias these associations toward the null. Moreover, the fact that such associations are constant across geography and populations is of interest to the literature on the health effects of discrimination and should be studied in further detail.

The design of our study had several strengths. First, we used a community-based approach which allowed us to incorporate the perspectives of Romani women in the design and conduct of the study, and to employ Romani interviewers who had access to Romani settlements. Second, our purposeful sampling method allowed us to include in the study sample women with varied socioeconomic status, enabling us to test DIF in different groups. Third, we tested both one-stage (EOD) and two-stage (EDS) assessment of ethnic discrimination. Finally, we administered the EDS and EOD in a random order, a limitation noted of previous validation studies of discrimination measures for health research (Shariff-Marco et al. 2011).

The assessment of perceived discrimination is challenging in any setting, especially in a socially excluded group that is understudied such as Roma, and thus our study faced several limitations. One important limitation is that the scales we evaluated measure only personally mediated discrimination, not institutional or internalized discrimination, (Jones 2000) both of which are likely prevalent in Serbia and Macedonia. For example, qualitative research suggested that housing conditions, lack of access to education and employment are major forms of institutional discrimination faced by Roma in the region (Janevic et al. 2011a, b). Field studies may be a fruitful direction to better understand institutional discrimination (Rich 2014). Men were also not included in the study; considering that Romani men and women may face unique stereotypes and forms of social exclusion; future research should include men as well. Additionally, items measuring perceived discrimination which capture the intersectional experience of being both Roma and female, such as treatment due to stereotypes specific to “Gypsy” women, could be developed (Harnois and Ifatunji 2011). All of these aspects of discrimination should be considered in future research on how discrimination toward Roma influences their health.

A further limitation of our study is that because it is not drawn from a random sample of the Romani population, inference regarding the frequency of certain variables to the entire population of Romani women cannot be made. However, as a rough estimate of generalizability, demographics of women in our sample can be compared to UNICEF Multiple Indicator Cluster (MICS) surveys of Romani women aged 15–49 conducted in 2011 (Multiple Indicator Cluster Survey Serbia 2011; Multiple Indicator Cluster Survey Macedonia 2012). The educational distribution of women in Serbia in our sample was very similar to the MICS Serbia sample, but in Macedonia our sample had a somewhat higher proportion of women with no education (36.2 % in our sample vs. 16.8 % in MICS Macedonia). However, it should be noted that the extent to which the MICS sample are representative of the Romani

population is not clear, because the sampling frames were based on old census information in which it is known that Roma are severely underrepresented. Given the difficulty of establishing a representative sample of Roma, we feel that utilizing a community-based approach with purposeful snowball sampling to yield high internal validity of discrimination measures is a priority over generalizability in research on discrimination and health.

In conclusion, we found that modified version of EOD and EDS have good psychometric support for use in health studies of Romani women. Item non-response for the EOD scale suggested that situations included in the EOD may not be appropriate for all marginalized populations globally, and population-specific situations should be considered when adapting this or similar scales. Items of the EDS showed lack of conditional independence and a shortened version could be used in future research. Researchers wishing to study discrimination and health in global settings may use either the EOD or EDS, but should take care to use a community-based approach in modifying the scale for relevance of items, if needed, and perform a psychometric assessment.

Acknowledgments We thank the organizations Bibija, Nexus Vranje, Lil, Zar, and Roma SOS for their valuable contributions, and Dr. Kellee White for review of the manuscript. This work was supported by the New Jersey Health Foundation.

References

- Abdikeeva A, Covaci A, Ezer T (2013) Roma health rights in Macedonia, Romania and Serbia: a baseline for legal advocacy Alina Covaci. *European J Public Health* 23(ckt123):081
- Bastos JL, Faerstein E, Celeste RK, Barros AJ (2012) Explicit discrimination and health: development and psychometric properties of an assessment instrument. *Revista de Saúde Pública* 46:269–278
- Bond JE (2004) Intersecting identities and human rights: the example of Romani women’s reproductive rights. *Geo J Gend L* 5:897
- Chan KT, Tran TV, Nguyen T (2012) Cross-cultural equivalence of a measure of perceived discrimination between Chinese-Americans and Vietnamese-Americans. *J Ethnic Cultural Divers Social Work* 21:20–36
- Cook B, Wayne GF, Valentine A, Lessios A, Yeh E (2013) Revisiting the evidence on health and health care disparities among the Roma: a systematic review 2003–2012. *Int J Public Health* 58:885–911
- Crowe D (2007) A history of the gypsies of Eastern Europe and Russia. Palgrave Macmillan, New York
- Cunningham TJ, Berkman LF, Gortmaker SL, Kiefe CI, Jacobs DR Jr, Seeman TE, Kawachi I (2011) Assessment of differential item functioning in the experiences of discrimination index: the Coronary Artery Risk Development in Young Adults (CARDIA) study. *Am J Epidemiol* 174:1266–1274
- Harnois CE, Ifatunji M (2011) Gendered measures, gendered models: toward an intersectional analysis of interpersonal racial discrimination. *Ethnic Racial Stud* 34:1006–1028
- Harris R, Cormack D, Tobias M, Yeh L, Talamaivao N, Minster J, Timutimu R (2012) The pervasive effects of racism: experiences

- of racial discrimination in New Zealand over time and associations with multiple health domains. *Soc Sci Med* 74:408–415
- Hooper D, Coughlan J, Mullen M (2008) Structural equation modelling: guidelines for determining model fit. *Articles*:2
- Ikram UZ, Snijder MB, Fassaert TJ, Schene AH, Kunst AE, Stronks K (2014) The contribution of perceived ethnic discrimination to the prevalence of depression. *Eur J Public Health* 25(2):243–248
- Janevic T, Jankovic J, Bradley E (2011a) Socioeconomic position, gender, and inequalities in self-rated health between Roma and non-Roma in Serbia. *Int J Public Health* 57:49–55
- Janevic T, Sripad P, Bradley E, Dimitrievska V (2011) “There’s no kind of respect here” A qualitative study of racism and access to maternal health care among Romani women in the Balkans. *Int J Equity Health* 10:53. doi:10.1186/1475-9276-10-53
- Jones CP (2000) Levels of racism: a theoretic framework and a gardener’s tale. *Am J Public Health* 90:1212–1215
- Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand ST, Manderscheid RW, Walters EE (2003) Screening for serious mental illness in the general population. *Arch Gen Psychiatry* 60:184–189
- Kim G, Sellbom M, Ford K (2014) Race/ethnicity and measurement equivalence of the everyday discrimination scale. *Psychol Assess* 26(3):892–900
- Krieger N (2012) Methods for the scientific study of discrimination and health: an ecosocial approach. *Am J Public Health* 102:936–944
- Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM (2005) Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. *Soc Sci Med* 61:1576–1596
- Lewis TT, Yang FM, Jacobs EA, Fitchett G (2012) Racial/ethnic differences in responses to the everyday discrimination scale: a differential item functioning analysis. *Am J Epidemiol* 175:391–401
- Lewis TT, Cogburn CD, Williams DR (2015) Self-reported experiences of discrimination and health: scientific advances, ongoing controversies, and emerging issues. *Annual review of clinical psychology*
- Ljubic V, Vedder K, Dekker H, Geel M (2013) Romaphobia among Serbian and Dutch adolescents: the role of perceived threat, nationalistic feelings, and integrative orientations. *Int J Psychol* 48:352–362
- Mullings L, Schulz A (2005) Intersectionality and health: an introduction. In: Schulz A, Mullings L (eds) *Gender, Race, Class & Health: Intersectional approaches*, pp 3–17
- Paradies Y (2006) A systematic review of empirical research on self-reported racism and health. *Int J Epidemiol* 35:888–901
- Paradies YC, Cunningham J (2008) International journal for equity in health. *Int J Equity Health* 7:9
- Pascoe EA, Smart Richman L (2009) Perceived discrimination and health: a meta-analytic review. *Psychol Bull* 135:531
- Pew Research Center (2009) Two decades after the wall’s fall: end of communism cheered but now with more reservations. Pew Research Center, Washington, DC
- Purnell JQ, Peppone LJ, Alcaraz K, McQueen A, Guido JJ, Carroll JK, Shacham E, Morrow GR (2012) Perceived discrimination, psychological distress, and current smoking status: results from the Behavioral Risk Factor Surveillance System Reactions to Race module, 2004–2008. *Am J Public Health* 102:844–851
- Ravnhol CI (2010) The human rights of minority women: Romani women’s rights from a perspective on international human rights law and politics. *Int J Minor Group Rights* 17:1–45
- Reeve BB, Willis G, Shariff-Marco SN, Breen N, Williams DR, Gee GC, Alegría M, Takeuchi DT, Stapleton M, Levin KY (2011) Comparing cognitive interviewing and psychometric methods to evaluate a racial/ethnic discrimination scale. *Field Methods* 23:397–419
- Republic of Macedonia multiple indicator cluster survey 2011 (2012)
- Rich J (2014) What do field experiments of discrimination in markets tell us? A meta analysis of studies conducted since 2000
- Shariff-Marco S, Breen N, Landrine H, Reeve BB, Krieger N, Gee GC, Williams DR, Mays VM, Ponce NA, Alegría M (2011) Measuring everyday racial/ethnic discrimination in health surveys. *Du Bois Rev Soc Sci Res Race* 8:159–177
- Statistical Office of the Republic of Serbia (2011) Multiple indicator cluster survey Serbia 2011. Statistical Office of the Republic of Serbia, Belgrade
- Stucky BD, Gottfredson NC, Panter A, Daye CE, Allen WR, Wightman LF (2011) An item factor analysis and item response theory-based revision of the Everyday Discrimination Scale. *Cultural Divers Ethn Minor Psychol* 17:175
- Szelenyi I, Ladanyi J (2006) Patterns of exclusion: constructing gypsy ethnicity and the making of an underclass in transitional societies of Europe. *East European Monographs*
- UNICEF (2007) Multiple indicator cluster survey. UNICEF, Serbia, Belgrade
- Williams DR, Mohammed SA (2009) Discrimination and racial disparities in health: evidence and needed research. *J Behav Med* 32:20–47
- Williams DR, Yan Y, Jackson JS, Anderson NB (1997) racial differences in physical and mental health: socio-economic status, stress and discrimination. *J Health Psychol* 2:335–351