Received: 5 October 2019 | Revised: 6 March 2020 | Accepted: 15 May 2020

DOI: 10.1002/mpr.1844

ORIGINAL ARTICLE

WILEY

Refining acculturation measures for health research: Latina/o heterogeneity in the National Latino and Asian American Study

Kimberly B. Roth | Rashelle J. Musci[1] | William W. Eaton[1]

Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland

Correspondence
Kimberly B. Roth, Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway Suite 850, Baltimore, MD 21205.
Email: kroth@wustl.edu

Present address
Kimberly B. Roth, Brown School of Social Work, Center for Mental Health Services Research, Washington University in St. Louis, St. Louis, Missouri

Funding information
National Institute of Mental Health, Grant/Award Numbers: 2T32MH019960, 5T32MH014592-39

Abstract

Objectives: This study factor analyzes six scales relating to acculturation and related experiences among a nationally representative sample of United States-residing Latina/os (n = 2,541) from the National Latino and Asian American Study (NLAAS), using measurement invariance (MI) testing to explore differences in latent constructs by Latina/o subgroup.

Methods: Factor Analysis (FA) within an Exploratory Structural Equation Modeling framework was used to analyze the factor structure of six scales measuring acculturation and related experiences (i.e., acculturation [language use and preference], enculturation [ethnic identity], discrimination, neighborhood context, and family environment). We tested for MI by two important Latina/o subgroups: ethnic heritage and generational status.

Results: The underlying latent factors resulting from FA strongly aligned with the NLAAS subscales. No scale achieved full MI, yet the degree to which MI held varied greatly by scale and by subgroup.

Conclusions: Findings show that Latina/os are heterogeneous, but that this often depends on the construct and subgrouping of interest. Future research should use these scales in a latent framework, accounting for the lack of MI, to ensure that the underlying acculturative constructs of interest are validly measured when investigating their association with mental health outcomes in this population.

KEYWORDS

acculturation, factor analysis, Latinos, measurement invariance, mental health

1 | INTRODUCTION

Latina/os are the largest foreign-born and third fastest growing minority in the United States (Colby & Ortman, 2015). As such, they will increasingly shape the prevalence of mental and behavioral disorders in the country. Acculturative processes have been consistently implicated in underlying mental and behavioral health disparities among Latina/os (Lara, Gamboa, Kahramanian, Morales, & Bautista, 2005; Romero & Piña-Watson, 2017). However, Latino mental health research is constrained by three major limitations: (a) lack of ethnic subgroup comparisons, (b) not accounting for generational status, and (c) inadequate measures of acculturation. To understand and reduce health disparities for the United States Latino population, a more nuanced approach must be used to disentangle the mix of risk and protective factors contributing to mental and behavioral disorder among Latina/os.

Experts have noted that analyses treating Latina/os as a homogenous ethnic group obscure within-group disparities among Latina/os (Alegría et al., 2007). The rare epidemiologic research that takes into account ethnic subgroups has found significant differences in prevalence of psychiatric disorder (Alcántara, Chen, & Alegría, 2014; Alegría et al., 2007), suicidality (Fortuna, Perez, Canino, Sribney, & Alegría, 2007), and general distress (Torres, Driscoll, & Voell, 2012). Differences also occur by generational status and time spent in the United States (Almeida, Johnson, Matsumoto, & Godette, 2012; Borges et al., 2011; Cook, Alegría, Lin, & Guo, 2009). These findings point toward the importance of acculturation in Latina/o mental health research.

Acculturation, defined as "the multidimensional process of the adoption of US cultural norms, values, and lifestyles" (Alegría, 2009, p. 996), has been linked to multiple mental and behavioral disorders (Alcántara et al., 2014; Blanco et al., 2013; Ortega, Rosenheck, Alegría, & Desai, 2000; Rivera et al., 2008; Valencia-Garcia, Simoni, Alegría, & Takeuchi, 2012), varying by ethnic subgroup and generational status (Guarnaccia et al., 2007). Acculturation is a general concept which connects to the process of immigration around the globe. The more recent concept of enculturation, or "the process of preserving the norms of the native group, whereby individuals retain identification with their ethnic cultures of origin" (Guarnaccia et al., 2007, p. 513), was introduced to expand the more traditional unidimensional approach used by acculturation researchers, allowing classification of individuals using a bidimensional model (Berry, 2003). Experiences such as discrimination and family conflict also correlate with acculturation and mental and behavioral disorder (Cook et al., 2009; Mulvaney-Day, Alegría, & Sribney, 2007; Rivera et al., 2008; Torres et al., 2012), exhibiting similar variations by subgroup (Perez, Fortuna, & Alegría, 2008). Therefore, studies investigating the complex relationships between acculturation, psychiatric morbidity, and these related experiences need to account for differences by ethnic and generational subgroup.

Despite its importance, acculturation has been measured insufficiently and inconsistently. A systematic review of acculturation measures used in public health research with United States Latina/o populations (Thomson & Hoffman-Goetz, 2009) found that most research uses proxies or unidimensional measures in health research, despite their noted inadequacy. Therefore, the authors called for a more thoughtful approach to conceptualizing and measuring acculturation, including the refinement of existing measures. As acculturation is inherently unobservable, latent variable methods are one such approach that can help capture the nuances of this complex construct.

Along with acculturation, other factors associated with mental disorder—such as family environment and discrimination—are also complex, unobserved constructs. Treating scales as observed variables (e.g., by summing them) ignores the possibility of measurement error, introducing bias and/or decreased reliability into analyses. In the context of complex processes such as acculturation, failing to account for measurement error may lead to conflicting results. This may be one reason acculturation research has made little progress in teasing apart the complex mechanisms that promote disorder in immigrants. Latent variable models address these limitations by taking advantage of multiple indicators to reflect an individual's true underlying score.

Exploratory Factor Analysis (EFA) addresses these limitations while also allowing the assessment of dimensionality of a measurement instrument. It seeks to find the smallest number of underlying factors to best explain the correlations among a set of observed variables (Bartholomew, Knott, & Moustaki, 2011; Muthén & Muthén, 1998–2017; Spearman, 1904). It is by nature exploratory, meaning no structure is imposed on the relationships between the observed variables and the unobserved factors, including which factors influence which items. While in the past Confirmatory Factor Analysis (CFA) has been considered a more rigorous approach to factor analysis than EFA, recent statistical developments and the advent of Exploratory Structural Equation Modeling (ESEM) have bridged the gap between EFA and CFA. ESEM, essentially the incorporation of EFA into an SEM framework, allows prior advantages of CFA, such as measurement invariance (MI) testing, to be implemented within the flexibility of an EFA framework (Asparouhov & Muthén, 2009; Marsh et al., 2009). In fact, studies comparing the two modeling approaches have demonstrated that the CFA assumption of zero cross-loadings necessary is often untenable and leads to poor model fit (Marsh et al., 2009; Marsh, Nagengast, & Morin, 2013). Instead, allowing even small but significant cross-loadings can be important to fully measure complex constructs.

It is important to test for the presence of MI in measurement models such as factor analyses. Because a large part of epidemiologic research involves comparison of means or prevalences, failing to account for differences in measurement across meaningful subgroups may lead to biased inferences (Meredith, 1993; Meredith & Teresi, 2006). An implicit assumption of group comparisons is construct consistency across groups or time, but if this assumption is invalid, observed mean differences may instead be due to construct variation. Confirming the presence of MI in a factor analysis model allows latent or observed construct scores to be validly compared. The ESEM approach is flexible enough to allow imposition of increasingly stringent constraints on the various parts of the factor analysis model to determine to what extent indictors and their constructs have the same relationships and correlational structure across multiple groups.

The nationally representative National Latino and Asian American Study (NLAAS) provides rich data on acculturation as well as psychiatric disorder. Measures were carefully selected and adapted (Alegría, Vila, et al., 2004), but their latent structures remain largely unexplored. The current analysis explores the latent variable properties of six scales evaluating United States Latina/os' acculturative (language, ethnic identity) and immigration-related experiences (neighborhood context, family context, discrimination, acculturative stress), taking into consideration differences between ethnic heritage and generational subgroups. The objective is to evaluate the factor structure of these scales and test for MI across Latina/o subgroups to determine the most appropriate way to use these scales in future studies.

2 | METHODS

The NLAAS is a nationally representative, probability-based survey conducted between 2001 and 2003 as part of the Collaborative Psychiatric Epidemiology Surveys (Heeringa et al., 2004; Pennell et al., 2004). The NLAAS oversampled its target population of civilian, noninstitutionalized adults (18 years or older) of Latina/o or Asian origin in the contiguous United States using a stratified, multi-frame probability sampling strategy (Pennell et al., 2004). Lay interviewers conducted in-person, computer-assisted structured interviews at the respondent's home. The Latina/o sample final response rate was 75.5% (Heeringa et al., 2004). These analyses utilized data from the 2,554 NLAAS participants of Latina/o ethnicity.

All NLAAS study procedures were approved by the Institutional Review Board (IRB) Committees of Cambridge Health Alliance, the University of Washington, and the University of Michigan (Pennell et al., 2004). Additional details regarding the study sample and procedures are found elsewhere (Alegría, Takeuchi, et al., 2004; Heeringa et al., 2004; Pennell et al., 2004). The IRB Office at the Johns Hopkins Bloomberg School of Public Health approved the present study (IRB#00008615).

2.1 | Measures

All questionnaires were adapted, translated into Spanish, and back-translated to ensure linguistic equivalency (Alegría, Vila, et al., 2004). Respondents could complete the interview in the language of their choice. This study utilizes ethnic/heritage subgroup, generational status, and measures of acculturation and related acculturative experiences (language, ethnic identity, discrimination, acculturative stress, neighborhood context, and family context). All six scales described below can be reviewed in full in Data S1.

2.1.1 | Acculturation and enculturation

2.1.1.1 Language

Level of acculturation was assessed via two Spanish and English language domains: proficiency (six items) and preference (three items). Three Spanish language proficiency items were taken from the Cultural Identity Scales for Latino Adolescents (Felix-Ortiz, Newcomb, & Myers, 1994). Corresponding English items were created specifically for the NLAAS English language proficiency scale and mirrored the Spanish language items. When summed, higher language-specific scores indicate greater language proficiency. For this study, the four response categories were dichotomized into Poor/Fair and Good/Excellent. Three language preference items (Felix-Ortiz et al., 1994) assessed preference for using Spanish or English in speaking and thinking. Response categories ranged from: "Spanish All the Time" to "English All the Time." Higher scores indicate increased preference for English.

2.1.1.2 Ethnic identity

Ethnic identity is the degree to which individuals identify with their own ethnic group, and is often used as a proxy measure for enculturation (Guarnaccia et al., 2007). This four-item scale assessed respondents' closeness and identification with others in their own ethnic group (Guarnaccia et al., 2007). Responses were collapsed into Low ("not at all" or "not very"), Medium ("somewhat"), and High ("very").

2.1.2 | Acculturative experiences

2.1.2.1 Neighborhood context

This 7-item scale has two subscales: Neighborhood Social Cohesion (four items) and Neighborhood Safety (three items) and reflects the cohesiveness and safety of respondents' neighborhoods. Adapted from three instruments (Bearman, Jones, & Udry, 1997; National Institute of Mental Health, 1994; Sampson, Raudenbush, & Earls, 1997), higher Social Cohesion scores indicate less neighborhood cohesion. The Neighborhood Safety scale assessed the respondent's perception of neighborhood violence and safety, with higher scores representing greater safety. All responses were dichotomized: Not true ("not very true" and "not at all true"), and True ("somewhat true" and "very true").

2.1.2.2 Family context

This 15-item measure contains three subscales: Family Pride (seven items), Family Cohesion (three items), and Family Cultural Conflict (five items). Family Pride assessed the respondent's feelings of loyalty and respect toward family members while Family Cohesion assessed feelings of closeness. Responses on both subscales (Olson, 1986, 1989) were collapsed into Agree ("Strongly" and "Somewhat") and Disagree ("Strongly" and "Somewhat").

The Family Cultural Conflict subscale (Cervantes, Padilla, & Salgado de Snyder, 1991) addresses intergenerational cultural conflict between respondents and their families. Questions assessed respondents' familial cultural conflict views and experiences with three response options dichotimized into: No ("Hardly ever or never") and Yes ("Sometimes" and "Often"). Higher scores indicate greater conflict.

2.1.2.3 Discrimination

Two subscales assessed discrimination. A nine-item everyday discrimination scale (Jackson, Williams, & Torres, 1995; Williams, Yu, Jackson, & Anderson, 1997) asked about the frequency of discriminatory experiences in day-to-day life. Six response options were collapsed into: Never ("Never"), Rarely ("A Few Times a Year" and "Less Than Once a Year") and Often ("Almost Every Day", "At Least Once a Week", and "A Few Times a Month"). The three-item perceived discrimination subscale (Vega, Gil, Warheit, Zimmerman, & Apospori, 1993) asked how often respondents or their friends are disliked or treated unfairly due to their Latina/o descent. Four response options were collapsed into: Never, Rarely and Often (from "Sometimes" and

"Often"). For both subscales, lower scores represent higher discrimination frequency.

2.1.2.4 Acculturative stress

Acculturative stress was assessed in the foreign-born population only using a nine-item scale (Vega et al., 1998). Items asked about the presence ("Yes" or "No") of feelings or experiences regarding transition to the United States, both in leaving their country of origin and their United States experiences.

2.2 | Population subgroups

Self-reported ethnic heritage was collapsed into four Latina/o subgroups: Puerto Ricans (n = 495), Mexicans (n = 868), Cubans (n = 577), and All Others (n = 614). Four categories of generational status were created: first-generation (arriving in the United States at age 13 or older, n = 1,257), 1.5 generation (Rumbaut & Rumbaut, 2005; arriving when less than age 13, n = 365), second-generation (United States-born, at least one parent foreign-born, n = 522) and third-generation (United States-born, both parents United States-born, n = 397). Thirteen respondents were unable to be classified by generational status and were therefore excluded from generation-specific analyses.

2.3 | Statistical analysis

All analyses were conducted in Mplus Version 8 (Muthén & Muthén, 1998–2017). EFA with Geomin rotation, a Weighted Least Squares Means and Variance (WLSMV) adjusted estimator, and full information maximum likelihood (FIML) was used. For each scale different factor models were compared using several absolute fit statistics, including a chi-square test, Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), and the Standardized Root Mean Square Residual (SRMR). Guidelines for good model fit are: CFI and TLI greater than 0.95, SRMR at or below 0.08, and RMSEA below 0.06 (Hu & Bentler, 1999). Final factor models were chosen based on scree plots, fit statistics, and factor interpretability based on item loadings.

MI was evaluated across heritage and generational subgroups using an ESEM framework (Asparouhov & Muthén, 2009) as described by Marsh et al. (2009) using a theta parameterization. Prior to evaluating MI, the factor structure within each subgroup was explored to verify equivalency to that of the overall sample ("configural invariance"). Marsh et al. (2009) put forth a 13-step process to fully evaluate MI in a scale with continuous indicators. Here, statistical limitations due to categorical indicators necessitated a 7-step process, excluding steps that test invariance of item uniquenesses/residuals (see Table 1). Models were named to harmonize with Marsh taxonomy, and steps testing for invariance of item intercepts had a secondary "partial" invariance step. To identify Models 1 through 4, factor means were necessarily constrained to be zero. Once item intercepts were constrained to be either fully or partially invariant as determined by Model 5, factor means were freed.

MI testing starts with Model 1 and proceeded through Model 12. Once the constraint of a model parameter significantly worsened model fit, invariance testing stopped. At this time, partial invariance of factor loadings is not allowed, nor is partial FVCV invariance. Models were compared using chi-squared difference testing at the p = .05 level and substantive evaluation of parameter estimates. Model 1 reflects configural invariance, meaning only the factor structure is the same across groups. With categorical indicators, Model 12 indicates full invariance.

3 | RESULTS

3.1 | Language

EFA fit statistics favored a 2-factor model (Table S1). Factor loadings for the total group analysis (Table S2) produced clear "Spanish" and "English" factors. Items 1 through 3 involving Spanish proficiency loaded strongly (all $\lambda > 0.850$) on the "Spanish" factor. English Proficiency items 4 through 6 had "English" loadings greater than 0.990. Language Preference subscale items loaded strongly (all $\lambda > 0.700$) on "English," with moderate inversely related "Spanish" cross-loadings (λ range:−0.304 to −0.402). Factors were significantly negatively correlated (r = −0.218).

Results from MI testing across subgroups are displayed in Table 2. When looking across both heritage and generational subgroups, constraining all factor loadings (Model 2) significantly worsened fit. Therefore, MI testing stopped and Model 1 (configural invariance) was chosen as the final model in both subgroups. Table S2 also presents standardized ESEM factor loadings, means, and variances for the chosen MI models. Although the pattern of ESEM-derived loadings was similar when compared to the EFA results, several differences highlight why loadings could not be constrained to equality. Among Mexicans, Spanish-speaking ability was less strongly related to the Spanish language construct as compared to other groups. While still small, the loadings of reading and writing in English among Puerto Ricans were two to three times the strength when relating to the Spanish language latent construct. All factors were modestly but significantly correlated, but less so among Puerto Ricans.

Differences were more striking across generational subgroups. For example, first-generation adult immigrants' language preference related less strongly with the Spanish language construct (all $\lambda < -0.19$), but had strong "English" loadings (all $\lambda > 0.76$). Among United States-born Latina/os, however, higher preference for thinking/communicating in English had a stronger negative association with the Spanish language factor. Moreover, factor correlations were more variable across subgroups, ranging from a significant positive correlation for adult migrants (r = 0.188) to an inverse but similar strength correlation among third-generation Latina/os (r = −0.219). The Spanish

TABLE 1 Overview of model taxonomy[a] for measurement invariance testing with categorical indicators

Model number	Invariant parameters	Description	Nested models
Model 1	None (FMn = 0)	Configural invariance	—
Model 2	FL (FMn = 0)	Weak factorial/measurement invariance	[1]
Model 3	FL, Uniq (FMn = 0)		[1,2]
Model 4	FL, FVCV (FMn = 0)		[1,2]
Model 5	FL, INT	Strong factorial/measurement invariance	[1,2,5p]
Model 5p	FL, INT(p)	Strong factorial/measurement invariance	[1,2]
Model 6	FL, Uniq, FVCV		[1,2,3,4]
Model 7	FL, Uniq, INT	Strict factorial/measurement invariance	[1,2,3,5]
Model 8	FL, FVCV, INT		[1,2,4,5]
Model 8p	FL, FVCV, INT(p)		[1,2,4,5p]
Model 9	FL, FVCV, INT, Uniq		[1–8]
Model 10	FL, INT, FMn	Latent mean invariance	[1,2,5]
Model 10p	FL, INT(p), FMn	Latent mean invariance	[1,2,5p]
Model 11	FL, Uniq, INT, FMn	Manifest mean invariance	[1,2,3,5,7,10]
Model 12	FL, FVCV, INT, FMn		[1,2,4,5,6,8,10]
Model 12p	FL, FVCV, INT(p), FMn		[1,2,4,5p,6,8p,10p]
Model 13	FL, FVCV, INT(p), Uniq, FMn	Complete factorial invariance	[1–12]

Note: Models in gray were unable to be tested with categorical factor indicators.
Abbreviations: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; p, partial; Uniq, item uniquenesses.
[a]Adapted from Marsh et al. (2009).

and English factors were almost unrelated among the 1.5 and second generations.

3.2 | Ethnic identity

EFA supported a 1-factor model (Table S3). Factor loadings (Table S4) ranged from 0.547 (identifying with others of similar descent) to 0.895 (marrying within one's racial/ethnic group). The latent "Identity" factor characterized identification with one's ethnic group, with higher scores indicating closer group identification. In MI testing across heritage subgroups, the model with the best fit (Model 5p; Table 3) had invariant factor loadings, partially invariant item intercepts, and free factor means and variances (strict factorial/MI). A similar pattern emerged when testing for invariance across generational subgroups, except factor variances were deemed invariant. Therefore, Model 8p was selected for best fit.

After accounting for measurement noninvariance, the standardized ESEM factor loadings differed somewhat from the EFA results (Table S4). In particular, the loading for item 4 decreased in strength by approximately two thirds, indicating that the importance of marrying someone of the same racial/ethnic descent is much less related to the latent construct of ethnic identity than what the EFA results portrayed. However, factor means and variances were not constant across subgroups. Compared to Puerto Ricans, Cubans had a significantly higher mean, whereas Mexicans and other Latina/os did not. All three groups had "Identity" factors with significantly more variation as compared to Puerto Ricans. Among the generational groups, first-generation immigrants were the only group with a mean significantly different than zero.

3.3 | Neighborhood context

A 2-factor EFA model was selected (Table S5). Items 1 through 4 loaded primarily on the first factor ("Community," range λ = 0.547 to 0.895; Table S6), while items 6 and 7 loaded heavily on the second factor ("Danger," λ = 0.896 and 0.712, respectively). Item 5 loaded significantly on both factors (λ = 0.400 and −0.469, respectively), which were significantly negatively correlated (r = −0.342). In MI testing (Table 4), Model 8p (invariant factor loadings, factor variances/covariances, partially invariant item intercepts, and free factor means) was the best-fitting model across heritage groups (strict factorial/MI). Across generational groups, factor loadings and means were deemed invariant but factor variances/covariances were not. Item intercepts were partially invariant. Therefore, Model 10p (latent mean invariance) was the best-fitting model.

The standardized ESEM factor loadings were similar to the EFA loadings (Table S6). Factor means varied significantly by heritage. Cubans had significantly higher levels of neighborhood community and danger than Puerto Ricans and other Latina/os. And while Mexicans were not different than Puerto Ricans and other Latina/os on levels of community, they reported significantly safer neighborhood environments, although not as safe as Cubans. Because the factor variance–covariance structure was allowed to vary across generational groups, standardized loadings varied slightly but were similar to

TABLE 2 Summary of goodness of fit statistics for all measurement invariance models for the 2-factor language scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	29	0.998	0.996	0.067 (0.060–0.075)	0.000	—	—	—	—	
Multiple group invariance models										**Invariant parameters**[a]
By ethnic heritage (four groups)										
MGI1	116	0.998	0.997	0.058 (0.049–0.066)	0.063	—	—	—	—	IN = none (FMn = 0)
MGI2	74	0.999	0.999	0.035 (0.027–0.043)	1.000	*68.412*	*42*	*[1]*	*0.0062*	IN = *FL* (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn
By generation (four groups)										
MGI1	116	0.996	0.992	0.066 (0.058–0.074)	0.001	—	—	—	—	IN = none (FMn = 0)
MGI2	74	0.996	0.996	0.051 (0.044–0.057)	0.436	*100.564*	*42*	*[1]*	*0.0000*	IN = *FL* (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn

Note: Gray highlight indicates chosen model. Italics indicate noninvariant parameters at $p < .05$ level.
Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; Params, parameters; RMSEA, root mean squared error of approximation; TG, total group; TLI, Tucker–Lewis Index.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

the EFA results. Factor correlations varied across groups, with the community and danger constructs for third-generation Latina/os being more inversely related ($r = -0.521$), but less so among first-generation adult migrants ($r = -0.220$). Compared to the third-generation, factor score distributions were more variable for other groups, with the exception of neighborhood danger among the first-generation.

3.4 | Family context

Based on the EFA results (Table S7), a 2-factor model was chosen. This model had clear loadings by subscale (Table S8): items 1 through 10 loaded heavily (all $\lambda > 0.75$) on the "Cohesion" factor, and items 11 through 15 on "Conflict" (all $\lambda > 0.7$). Other loadings were low, with several items having significant cross-loadings. Factors were significantly negatively correlated ($r = -0.603$). MI testing across heritage groups revealed that Model 8p (strict factorial/MI) was optimal (Table 5). This model had invariant factor loadings, variances/covariances, and partially invariant item intercepts. Factor means and the intercept for item 12 (arguments with family members) were unable to be constrained to equality. In contrast, only configural invariance was obtained (Model 1) during MI testing across generational subgroups.

Standardized ESEM-derived factor loadings by heritage subgroups closely resembled EFA loadings (Table S8). Compared to Puerto Ricans, only Cubans had significantly less conflict, yet all groups had significantly higher family cohesion. Conversely, the ESEM loadings were more heterogeneous across generational groups. In particular, the 1.5 Generation had more item cross-loadings as compared to other subgroups. For some items, these childhood immigrants looked more like their US-born counterparts with strong negative loadings of respect and shared values (items 1 and 2). Among other loadings, there was more of a gradient across the generations, for example in items 1 and 2, where the family cohesion loading was the strongest among the first-generation and the weakest among the third.

3.5 | Discrimination

EFA revealed a 2-factor model with satisfactory fit (Table S9). Although the RMSEA was larger than desirable, inspection of the eigenvalues and scree plot did not justify adding a third factor. Factor loadings (Table S10) mirrored the subscales, with the nine Everyday Discrimination items resulting in an "Observed" factor (λ range: 0.691 to 0.978), and the three Perceived Discrimination items loading on a

TABLE 3 Summary of goodness of fit statistics for all measurement invariance models for the 1-factor ethnic identity scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	12	0.996	0.989	0.061 (0.039–0.085)	.196	—	—	—	—	
Multiple group invariance models										**Invariant parameters[a]**
By ethnic heritage (four groups)										
MGI1	48	0.997	0.990	0.057 (0.032–0.084)	.286	—	—	—	—	IN = none (FMn = 0)
MGI2	39	0.996	0.994	0.044 (0.025–0.063)	.672	17.091	9	[1]	.0473	IN = FL (FMn = 0)
MGI4	36	0.994	0.993	0.049 (0.032–0.066)	.515	*11.160*	*3*	*[2]*	*.0109*	IN = FL, *FVCV* (FMn = 0)
MGI5	18	0.991	0.994	0.044 (0.031–0.056)	.780	*47.488*	*21*	*[2]*	*.0008*	IN = FL, *INT*
MGI5p	25	0.995	0.996	0.037 (0.022–0.051)	.929	18.664	14	[2]	.1782	IN = FL, INT(p)
MGI8p	22	0.992	0.995	0.041 (0.028–0.055)	.851	*10.286*	*3*	*[5p]*	*.0163*	IN = FL, *FVCV*, INT(p)
MGI10p	22	0.962	0.973	0.093 (0.082–0.105)	.000	*77.029*	*3*	*[5p]*	*.0000*	IN = FL, INT(p), *FMn*
MGI12p	N/A							[10p]		IN = FL, FVCV, INT(p), FMn
By generation (four groups)										
MGI1	48	0.997	0.991	0.052 (0.026–0.079)	.405	—	—	—	—	IN = none (FMn = 0)
MGI2	39	0.999	0.999	0.019 (0.000–0.043)	.988	5.217	9	[1]	.8150	IN = FL (FMn = 0)
MGI4	36	0.999	0.999	0.019 (0.000–0.041)	.993	3.849	3	[2]	.2782	IN = FL, FVCV (FMn = 0)
MGI5	18	0.969	0.980	0.079 (0.068–0.090)	.000	*177.874*	*21*	*[2]*	*.0000*	IN = FL, *INT*
MGI5p	29	0.998	0.998	0.022 (0.000–0.041)	.996	15.288	10	[2]	.1219	IN = FL, INT(p)
MGI8p	26	0.998	0.998	0.024 (0.000–0.041)	.996	5.025	3	[5p]	.1700	IN = FL, FVCV, INT(p)
MGI10p	26	0.988	0.991	0.055 (0.042–0.068)	.260	*28.624*	*3*	*[5p]*	*.0000*	IN = FL, INT(p), *FMn*
MGI12p	N/A							[10p]		IN = FL, FVCV, INT(p), FMn

Note: Gray highlight indicates chosen model. Italics indicate noninvariant parameters at $p < .05$ level.
Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; p, partial; Params, parameters; RMSEA, root mean squared error of approximation; TG, total group; TLI, Tucker–Lewis Index.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

"Perceived" factor (all $\lambda > 0.720$). There were several low cross-loadings of the Everyday items on the "Perceived" factor (λ range: 0.210 to 0.350), despite significant factor correlation ($r = 0.407$). Items 8 (insulted) and 9 (threatened/harassed) also had low inverse loadings on the "Perceived" factor. In MI testing, constraining all factor loadings significantly worsened fit, making Model 1 (configural invariance) the final model in both subgroupings (Table 6).

The ESEM standardized loadings had a similar pattern to those from the EFA. And while consistent across subgroups, there was some variation in loading strength contributing to the rejection on invariant loadings during MI testing. For example, loadings for item 12 (seen friends treated unfairly due to being Latina/o) on the perceived discrimination factor ranged from 0.686 in other Latina/os to 0.836 in Cubans. In general, these differences were mild, and the moderate positive correlation between the factors was stable. Differences in standardized ESEM loadings by generational status were similarly mild, although again item 12 varied in loadings on the perceived construct from 0.644 in 1.5 Generation Latina/os to 0.797 in the First. Factor correlations, while similar, ranged from 0.294 in the 1.5 Generation to 0.407 in the first-generation, again putting the foreign-born groups at opposite ends of the spectrum.

3.6 | Acculturative stress

Acculturative stress items were not assessed in the United States-born population. In the overall sample, a 2-factor model fit the data best (Table S11). Items 1 through 6 loaded most heavily on the first factor ("Interpersonal") and items 8 and 9 on the second ("Legal"; see Table S12). Item 7 (questioned about legal status) loaded modestly on both factors, which were highly correlated ($r = 0.522$). Among subgroups, all except the 1.5 Generation supported 2 factors. The 1.5 Generation, however, required a 1-factor solution. Because configural invariance was not obtained, MI testing by generational status could not continue.

MI testing occurred only among three heritage groups; Puerto Ricans were excluded from testing as there was no variability in item 9 (avoided health services due to fear of immigration officials). When constraining factor loadings to be equal (Model 2), model convergence was not achieved (Table 7), indicating poor model fit. Therefore, Model 1 (configural invariance) was selected. The standardized ESEM-derived factor loadings among heritage subgroups were inconsistent in pattern (Table S12). Cubans and Mexicans resembled one another, but the loadings among other Latina/os followed no clear pattern.

TABLE 4 Summary of goodness of fit statistics for all measurement invariance models for the 2-factor neighborhood context scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	20	0.987	0.967	0.060 (0.049–0.072)	.071	—	—	—	—	
Multiple group invariance models										**Invariant parameters[a]**
By ethnic heritage (four groups)										
MGI1	80	0.984	0.958	0.068 (0.055–0.080)	.010	—	—	—	—	IN = none (FMn = 0)
MGI2	50	0.991	0.988	0.036 (0.025–0.046)	.987	23.575	30	[1]	.7909	IN = FL (FMn = 0)
MGI4	41	0.993	0.992	0.030 (0.018–0.040)	1.000	12.706	9	[2]	.1764	IN = FL, FVCV (FMn = 0)
MGI5	35	0.986	0.985	0.040 (0.031–0.049)	.963	*47.272*	*15*	*[2]*	*.0000*	IN = FL, *INT*
MGI5p	40	0.992	0.990	0.032 (0.022–0.042)	.999	4.329	10	[2]	.9313	IN = FL, INT(p)
MGI8p	31	0.993	0.993	0.028 (0.016–0.038)	1.000	12.921	9	[5p]	.1662	IN = FL, FVCV, INT(p)
MGI10p	34	0.983	0.982	0.045 (0.036–0.053)	.837	*36.579*	*6*	*[5p]*	*.0000*	IN = FL, INT(p), *FMn*
MGI12p	N/A							[10p]		IN = FL, FVCV, INT(p), FMn
By generation (four groups)										
MGI1	80	0.990	0.973	0.055 (0.043–0.069)	.230	—	—	—	—	IN = none (FMn = 0)
MGI2	50	0.994	0.992	0.030 (0.017–0.041)	.999	24.841	30	[1]	.7327	IN = FL (FMn = 0)
MGI4	41	0.993	0.992	0.031 (0.020–0.041)	.999	*18.991*	*9*	*[2]*	*.0253*	IN = FL, *FVCV* (FMn = 0)
MGI5	35	0.976	0.974	0.054 (0.046–0.063)	.182	*148.377*	*15*	*[2]*	*.0000*	IN = FL, *INT*
MGI5p	44	0.994	0.993	0.028 (0.016–0.039)	1.000	3.805	6	[2]	.7031	IN = FL, INT(p)
MGI8p	N/A							[5p]		IN = FL, FVCV, INT(p)
MGI10p	38	0.994	0.993	0.029 (0.017–0.039)	1.000	10.766	6	[5p]	.0959	IN = FL, INT(p), FMn
MGI12p	N/A							[10p]		IN = FL, FVCV, INT(p), FMn

Note: Gray highlight indicates chosen model. Italics indicate noninvariant parameters at $p < .05$ level.
Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; p, partial; Params, parameters; RMSEA, root mean squared error of approximation; TG, total group; TLI, Tucker–Lewis Index.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

Although a 2-factor solution was preferred in this group based on fit statistics (data not shown), the factors were not meaningfully interpretable. The two latent constructs of interpersonal and legal stress were most highly correlated among Mexicans ($r = 0.482$), and the least among Cubans ($r = 0.329$). All correlations were somewhat attenuated from those in the EFA. Because MI testing could not be pursued by generational group, stratified EFA results are presented instead (Table S12). First-generation immigrants had clearly distinguished "Interpersonal" and "Legal" factors, which were significantly correlated ($r = 0.447$). The single factor that emerged for the 1.5 Generation, labeled "Stress," had loadings that ranged in absolute strength from 0.402 (item 2) to 0.990 (item 9).

4 | DISCUSSION

The EFA results make substantive sense based on the design of the scales, corroborating the quality of theory and testing that went into the NLAAS development (Alegría, Vila, et al., 2004). A notable exception was Neighborhood Context. Item 5 from the Neighborhood Safety subscale loaded moderately on both the Community and Danger factors. This is unsurprising, as the statement "I feel safe being out alone in my neighborhood during the night" intuitively relates to both the community structure of one's neighborhood and the perception of safety.

Complete factorial invariance was not achieved for any scale across the subgroups examined, although this varied significantly by scale and subgroup. The Neighborhood Context and Ethnic Identity scales were more alike across subgroupings, indicating that the underlying constructs are generally similar regardless of group membership. However, inability to achieve latent mean invariance (and thus manifest mean invariance) underscores the importance of treating these constructs as latent rather than observed. Conversely, Language, Discrimination, and Acculturative Stress attained only the loosest type of invariance (configural), demonstrating that certain constructs are extremely heterogeneous by Latina/o subgroup in how individual items relate to their underlying constructs. Finally, the Family Context scale was similar across ethnic heritage but extremely noninvariant by generation, indicating the need to consider multiple subgroupings. Further, using highly noninvariant scales to compare Latina/o generational groups is invalid without accounting for this fact; Ignoring it can lead to biased results when looking at associations between family conflict and mental disorder.

These results emphasize that Latina/os are heterogeneous across countries of origin in more than just observed endorsement of specific experiences. On average, Cubans had greater ethnic identity than other Latina/os, however within-group variability was also larger. Cubans also tended to have more favorable neighborhood

TABLE 5 Summary of goodness of fit statistics for all measurement invariance models for the 2-factor family context scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	44	0.992	0.989	0.031 (0.027–0.035)		—	—	—	—	
Multiple group invariance models										**Invariant parameters[a]**
By ethnic heritage (four groups)										
MGI1	176	0.994	0.991	0.027 (0.021–0.032)	1.000	—	—	—	—	IN = none (FMn = 0)
MGI2	98	0.995	0.994	0.022 (0.016–0.027)	1.000	*99.942*	*78*	*[1]*	*.0478*	IN = FL (FMn = 0)
MGI4	89	0.995	0.995	0.020 (0.014–0.026)	1.000	*15.365*	*9*	[2]	.0814	IN = FL, FVCV (FMn = 0)
MGI5	59	0.994	0.994	0.022 (0.016–0.027)	1.000	*62.512*	*39*	*[2]*	*.0098*	IN = FL, *INT*
MGI5p	62	0.995	0.994	0.021 (0.016–0.026)	1.000	43.212	36	[2]	.1904	IN = FL, INT(p)
MGI8p	53	0.995	0.995	0.020 (0.014–0.025)	1.000	15.157	*9*	[5p]	.0867	IN = FL, FVCV, INT(p)
MGI10p	56	0.992	0.992	0.025 (0.020–0.030)	1.000	*25.414*	*6*	*[5p]*	*.0003*	IN = FL, INT(p), *FMn*
MGI12p	N/A							[10p]		IN = FL, FVCV, INT(p), FMn
By generation (four groups)										
MGI1	176	0.995	0.993	0.024 (0.018–0.030)	1.000	—	—	—	—	IN = none (FMn = 0)
MGI2	98	0.994	0.993	0.023 (0.018–0.028)	1.000	*119.273*	*78*	*[1]*	*.0018*	IN = *FL* (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT(p)
MGI10	N/A							[5]		IN = FL, INT(p), FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT(p), FMn

Note: Gray highlight indicates chosen model. Italics indicate noninvariant parameters at $p < .05$ level.
Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; p, partial; Params, parameters; TG, total group; TLI, Tucker–Lewis Index; RMSEA, root mean squared error of approximation.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

TABLE 6 Summary of goodness of fit statistics for all measurement invariance models for the 2-factor discrimination scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	47	0.979	0.967	0.120 (0.115–0.125)	.000	—	—	—	—	
Multiple group invariance models										**Invariant parameters[a]**
By ethnic heritage (four groups)										
MGI1	188	0.979	0.968	0.117 (0.112–0.122)	.000	—	—	—	—	IN = none (FMn = 0)
MGI2	128	0.988	0.986	0.077 (0.072–0.081)	.000	*93.749*	*60*	*[1]*	*.0035*	IN = *FL* (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn
By generation (four groups)										
MGI1	188	0.978	0.966	0.116 (0.111–0.122)	.000	—	—	—	—	IN = none (FMn = 0)
MGI2	128	0.986	0.984	0.081 (0.076–0.085)	.000	*116.796*	*60*	*[1]*	*.0000*	IN = *FL* (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn

Note: Gray highlight indicates chosen model. Italics indicate noninvariant parameters at $p < .05$ level.
Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; Params, parameters; RMSEA, root mean squared error of approximation; TG, total group; TLI, Tucker–Lewis Index.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

TABLE 7 Summary of goodness of fit statistics for all measurement invariance models for the 2-factor acculturative stress scale

Model	# Free Params	CFI	TLI	RMSEA (95% CI)	*p*-value	χ2 DiffTest	*df*	Comparison model	*p*-value	
Total group models										
ESEM	26	0.959	0.922	0.067 (0.057–0.077)	.002	—	—	—	—	
Multiple group invariance models										**Invariant parameters[a]**
By ethnic heritage (three groups)										
MGI1	78	0.964	0.932	0.061 (0.050–0.073)	.054	—	—	—	—	IN = none (FMn = 0)
MGI2	Model did not converge							[1]		IN = FL (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn
By generation (two groups)										
MGI1	Configural invariance not obtained					—	—	—	—	IN = none (FMn = 0)
MGI2	N/A							[1]		IN = FL (FMn = 0)
MGI4	N/A							[2]		IN = FL, FVCV (FMn = 0)
MGI5	N/A							[2]		IN = FL, INT
MGI8	N/A							[5]		IN = FL, FVCV, INT
MGI10	N/A							[5]		IN = FL, INT, FMn
MGI12	N/A							[10]		IN = FL, FVCV, INT, FMn

Abbreviations: CFI, comparative fit index; df, degrees of freedom; DiffTest, difference test; MGI, multiple group invariance; Params, parameters; RMSEA, root mean squared error of approximation; TG, Total group; TLI, Tucker–Lewis Index.
[a]For multiple group invariance models, IN means the sets of parameters constrained to be invariant across the multiple groups: FL, factor loadings; FMn, factor means; FVCV, factor variance–covariances; INT, item intercepts; Uniq, item uniquenesses.

environments, less family conflict, and more cohesion. English and Spanish ability and preference were less correlated among Puerto Ricans. They also had the lowest family cohesion, with more conflict. Finally, the individual feelings and experiences that contribute to acculturative stress varied greatly by heritage group. These inconsistent patterns, along with relatively poor model fit statistics, call into question whether "acculturative stress" is a valid universal construct, or whether more nuanced and culturally relevant experiences need to be developed, particularly because the factor structure was uninterpretable among the still heterogeneous group of "Other Latina/os."

Latina/os are also heterogeneous across generational groups. It is especially clear that it is imperative distinguish the 1.5 Generation from their first-generation counterparts who arrived in the United States as teenagers or adults. For example, the 2-factor Family Context scale mirrored the Family Pride and Cohesion subscales and the Family Cultural Conflict subscale (loading on the Cohesion and Conflict factors, respectively) in most groups. However, for the 1.5 Generation, conflict items often heavily cross-loaded on both factors. Additionally, the acculturative stress factor structure was bidimensional in first-generation immigrants arriving as teenagers or adults, but unidimensional among Latina/os arriving in the United States as children. These findings highlight the understudied phenomenon that this generational group is qualitatively different than other first-generation immigrants (Rumbaut & Rumbaut, 2005; Zhou, 1997), and points toward the need to consider timing of developmentally relevant risk and protective factors across the lifespan.

Differences were also seen in varying factor mean across groups for multiple latent constructs. For example, first-generation immigrants had a significantly higher "Identity" factor mean, whereas the 1.5 Generation did not differ from their United States-born counterparts. This suggests that Latina/os migrating at older ages identify more strongly with their country of origin, whereas immigrants spending their formative years in the United States identify with their ethnic roots no more strongly than those born in the States. The 1.5 Generation also had fairly independent English and Spanish abilities and preference, much like their second-generation counterparts. In many ways, the underlying language constructs resembled those of United States-born Latina/os, except, understandably, the "preference" of speaking a certain language with family members; This more likely speaks to the necessity of speaking Spanish to family who are not proficient in English. Family Context loadings were also more reminiscent of Latina/os born in the United States, although the 1.5 Generation tended to have higher item cross-loadings, suggesting a blurring of the line between conflict and cohesion.

Despite its important findings that underscore Latina/o heterogeneity, this study is not without limitations. The NLAAS data are self-report and therefore subject to bias. Although the NLAAS has a large Latina/o sample size, some subgroups were relatively small, reducing power and contributing to some model nonconvergence. Due to small cell size within subgroups, some item response options were collapsed to address this issue. This results in a loss of information and requires assumptions about meaningful cut points in how responses were grouped. Participants were allowed to respond to different sections of the interview in either Spanish or English, depending on their preference. The data on language response by section is not available, prohibiting MI testing by language in this study. In the future it would be important to collect this level of detail in order to examine MI by language, as differences in each construct by language would be possible. Finally, the "Other Latina/o" category still represents a subgroup with considerable heterogeneity.

5 | CONCLUSION

The NLAAS is one of the largest, nationally representative samples of United States Latina/os with rich data on acculturation, ethnic identity, and other important contextual factors, allowing the first testing of MI across both Latina/o heritage and generational subgroups. The separation of child immigrants from their first-generation counterparts is an important distinction not often accounted for. To our knowledge, this is the first in-depth exploration of the factor structure of all scales, including assessment of MI. The findings underscore the need to account for Latina/o heterogeneity, not simply at the manifest level, but at the latent construct level. In the future, investigators can appropriately model these constructs when investigating associations with health outcomes in the NLAAS Latina/o sample without having to formally test for MI as the exact level of MI found in each scale across specific subgroups is provided. The resulting factor scores are calibrated in regard to either heritage or generational group, depending on the MI testing. These scores can then be used to make more valid comparisons when estimating associations with mental and behavioral disorder across subgroups.

The findings in these analyses underscore the need to consider acculturation carefully when conducting research on immigrant populations. Acculturation researchers should apply these methods when investigating mental health disparities in other samples and datasets.

ACKNOWLEDGEMENTS

We gratefully acknowledge the contribution of NLAAS participants and the contribution of Dr M. Claire Green and Dr Ramin Mojtabai for their comments on the manuscript throughout its development. This work was supported by the National Institute of Mental Health (Grant Numbers: 5T32MH014592-39, 2T32MH019960).

DECLARATION OF INTEREST STATEMENT

The authors declare no conflicts of interest.

ORCID

Kimberly B. Roth https://orcid.org/0000-0001-5006-9362

REFERENCES

Alcántara, C., Chen, C. N., & Alegría, M. (2014). Do post-migration perceptions of social mobility matter for Latino immigrant health? *Social Science and Medicine*, *101*, 94–106. https://doi.org/10.1016/j.socscimed.2013.11.024

Alegría, M. (2009). The challenge of acculturation measures: What are we missing? A commentary on Thomson & Hoffman-Goetz. *Social Science & Medicine*, *69*(7), 996–998. https://doi.org/10.1016/j.socscimed.2009.07.006

Alegría, M., Vila, D., Woo, M., Canino, G., Takeuchi, D., Vera, M., … Shrout, P. (2004). Cultural relevance and equivalence in the NLAAS instrument: Integrating etic and emic in the development of cross-cultural measures for a psychiatric epidemiology and services study of Latinos. *International Journal of Methods in Psychiatric Research*, *13*(4), 270–288.

Alegría, M., Mulvaney-Day, N., Torres, M., Polo, A., Cao, Z., & Canino, G. (2007). Prevalence of psychiatric disorders across Latino subgroups in the United States. *American Journal of Public Health*, *97*(1), 68–75. https://doi.org/10.2105/AJPH.2006.087205

Alegría, M., Takeuchi, D., Canino, G., Duan, N., Shrout, P., Meng, X.-L., … Gong, F. (2004). Considering context, place and culture: The National Latino and Asian American study. *International Journal of Methods in Psychiatric Research*, *13*(4), 208–220. https://doi.org/10.1002/mpr.178

Almeida, J., Johnson, R. M., Matsumoto, A., & Godette, D. C. (2012). Substance use, generation and time in the United States: The modifying role of gender for immigrant urban adolescents. *Social Science & Medicine*, *75*(12), 2069–2075. https://doi.org/10.1016/j.socscimed.2012.05.016

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Bartholomew, D., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach* (3rd ed.). West Sussex, England: John Wiley & Sons.

Bearman, J., Jones, J., & Udry, R. (1997). *The national longitudinal study of adolescent health*. Chapel Hill, NC: Carolina Population Center.

Berry, J. W. (2003). Conceptual approaches to acculturation). In K. M. Chun, P. Balls Organista, & G. Marin (Eds.), *Acculturation: Advances in theory, measurement, and applied research* (Vols. 1–Book, Section) (pp. 17–37). Washington, DC: American Psychological Association. https://doi.org/10.1037/10472-004

Blanco, C., Morcillo, C., Alegría, M., Dedios, M. C., Fernández-Navarro, P., Regincos, R., & Wang, S. (2013). Acculturation and drug use disorders among Hispanics in the U.S. *Journal of Psychiatric Research*, *47*(2), 226–232. https://doi.org/10.1016/j.jpsychires.2012.09.019

Borges, G., Azrael, D., Almeida, J., Johnson, R. M., Molnar, B. E., Hemenway, D., & Miller, M. (2011). Immigration, suicidal ideation and deliberate self-injury in the Boston youth survey 2006. *Suicide & Life-Threatening Behavior*, *41*(2), 193–202. https://doi.org/10.1111/j.1943-278X.2010.00016.x

Cervantes, R. C., Padilla, A. M., & Salgado de Snyder, N. (1991). The Hispanic stress inventory: A culturally relevant approach to psychosocial assessment. *Psychological Assessment*, *3*(3), 438–447. https://doi.org/10.1037/1040-3590.3.3.438

Colby, S. L., & Ortman, J. M. (2015). *Projections of the size and composition of the U.S. population: 2014 to 2060* (No. P25-1143; p. 13). U.S. Census Bureau.

Cook, B., Alegría, M., Lin, J. Y., & Guo, J. (2009). Pathways and correlates connecting Latinos' mental health with exposure to the United States. *American Journal of Public Health*, *99*(12), 2247–2254. https://doi.org/10.2105/AJPH.2008.137091

Felix-Ortiz, M., Newcomb, M. D., & Myers, H. (1994). A multidimensional measure of cultural identity for Latino and Latina adolescents. *Hispanic Journal of Behavioral Sciences*, *16*(2), 99–115. https://doi.org/10.1177/07399863940162001

Fortuna, L. R., Perez, D. J., Canino, G., Sribney, W., & Alegría, M. (2007). Prevalence and correlates of lifetime suicidal ideation and suicide attempts among Latino subgroups in the United States. *The Journal of Clinical Psychiatry*, *68*(4), 572–581. https://doi.org/10.4088/JCP.v68n0413

Guarnaccia, P. J., Pincay, I. M., Alegría, M., Shrout, P. E., Lewis-Fernández, R., & Canino, G. J. (2007). Assessing diversity among Latinos: Results from the NLAAS. *Hispanic Journal of Behavioral Sciences*, *29*(4), 510–534. https://doi.org/10.1177/0739986307308110

Heeringa, S. G., Wagner, J., Torres, M., Duan, N., Adams, T., & Berglund, P. (2004). Sample designs and sampling methods for the collaborative psychiatric epidemiology studies (CPES). *International Journal of Methods in Psychiatric Research*, *13*(4), 221–240.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jackson, J. S., & Williams, D. R. (1995). In Detroit Area Study, 1995: Social Influences on Health: Stress, Racism, and Health Protective Resources.

Lara, M., Gamboa, C., Kahramanian, M. I., Morales, L. S., & Bautista, D. E. H. (2005). Acculturation and Latino health in the United States: A review of the literature and its sociopolitical context. *Annual Review of Public Health*, *26*(Journal Article), 367–397. https://doi.org/10.1146/annurev.publhealth.26.021304.144615

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 439–476. https://doi.org/10.1080/10705510903008220

Marsh, H.W., Nagengast, B., & Morin, A.J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*(6), 1194–1218. https://doi.org/10.1037/a0026913.

Meredith, W. (1993). easurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*(4), 523–543. https://doi.org/10.1007/BF02294825.

Meredith, W., & Teresi, J. A . (2006). An essay on measurement and factorial invariance. *Medical Care*, *444*(suppl 3), S69–77. https://doi.org/10.1097/01.mlr.0000245438.73837.89.

Mulvaney-Day, N. E., Alegría, M., & Sribney, W. (2007). Social cohesion, social support, and health among Latinos in the United States. *Social Science & Medicine*, *64*(2), 477–495. https://doi.org/10.1016/j.socscimed.2006.08.030

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide (Eighth)*. Los Angeles, CA: Muthén & Muthén.

National Institute of Mental Health (1994). Cooperative agreement for a multi-site study of mental health service use, need, outcomes, and costs in child and adolescent populations (UNO-CAP). In *Catalog of Federal Domestic Assistance 93.242-request for applications* (Vols. 1–Book, Section). Rockville, MD: NIMH.

Olson, D. H. (1986). Circumplex model VII: Validation studies and FACES III. *Family Process*, *25*(3), 337–351.

Olson, D. H. (1989). Circumplex model of family systems VIII: Family assessment and intervention. In D. H. Olson, C. S. Russell, & D. H. Sprenkle (Eds.), *Circumplex model: Systemic assessment and treatment of families* (Vols. 1–Book, Section). New York, NY: Haworth Press.

Ortega, A. N., Rosenheck, R. M. D., Alegría, M., & Desai, R. A. (2000). Acculturation and the lifetime risk of psychiatric and substance use disorders among Hispanics. *Journal of Nervous*, *188*(11), 728–735.

Pennell, B. E., Bowers, A., Carr, D., Chardoul, S., Cheung, G. Q., Dinkelmann, K., … Torres, M. (2004). The development and implementation of the National Comorbidity Survey Replication, the National Survey of American life, and the National Latino and Asian American survey. *International Journal of Methods in Psychiatric Research*, *13*(4), 241–269.

Perez, D. J., Fortuna, L., & Alegría, M. (2008). Prevalence and correlates of everyday discrimination among U.S. Latinos. *Journal of Community Psychology*, *36*(4), 421–433. https://doi.org/10.1002/jcop.20221

Rivera, F. I., Guarnaccia, P. J., Mulvaney-Day, N., Lin, J. Y., Torres, M., & Alegría, M. (2008). Family cohesion and its relationship to psychological distress among Latino groups. *Hispanic Journal of Behavioral Sciences*, *30*(3), 357–378.

Romero, A., & Piña-Watson, B. (2017). In S. J. Schwartz & J. B. Unger (Eds.), *The Oxford handbook of acculturation and health Acculturative stress and bicultural stress: Psychological measurement and mental health* (pp. 119–133). New York, NY: Oxford University Press.

Rumbaut, R. D., & Rumbaut, R. G. (2005). Self and circumstance: Journeys and visions of exile. In *The dispossessed: An anatomy of exile* (pp. 331–355). Amherst, MA: University of Massachusetts Press.

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, *277*(5328), 918–924.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201–292. https://doi.org/10.2307/1412107

Thomson, M. D., & Hoffman-Goetz, L. (2009). Defining and measuring acculturation: A systematic review of public health studies with Hispanic populations in the United States. *Social Science & Medicine*, *69*(7), 983–991. https://doi.org/10.1016/j.socscimed.2009.05.011

Torres, L., Driscoll, M. W., & Voell, M. (2012). Discrimination, acculturation, acculturative stress, and Latino psychological distress: A moderated mediational model. *Cultural Diversity and Ethnic Minority Psychology*, *18*(1), 17–25. https://doi.org/10.1037/a0026710

Valencia-Garcia, D., Simoni, J. M., Alegría, M., & Takeuchi, D. T. (2012). Social capital, acculturation, mental health, and perceived access to services among Mexican American women. *Journal of Consulting and Clinical Psychology*, *80*(2), 177–185. https://doi.org/10.1037/a0027207

Vega, W. A., Gil, A. G., Warheit, G. J., Zimmerman, R. S., & Apospori, E. (1993). Acculturation and delinquent behavior among Cuban American adolescents: Toward an empirical model. *American Journal of Community Psychology*, *21*(1), 113–125. https://doi.org/10.1007/BF00938210

Vega, W. A., Kolody, B., Aguilar-Gaxiola, S., Alderete, E., Catalano, R., & Caraveo-Anduaga, J. (1998). Lifetime prevalence of DSM-III-R psychiatric disorders among urban and rural Mexican Americans in California. *Archives of General Psychiatry*, *55*(9), 771–778. https://doi.org/10.1001/archpsyc.55.9.771

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*(3), 335–351. https://doi.org/10.1177/135910539700200305

Zhou, M. (1997). Growing up American: The challenge confronting immigrant children and children of immigrants. *Annual Review of Sociology*, *23*(1), 63–95. https://doi.org/10.1146/annurev.soc.23.1.63

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Roth KB, Musci RJ, Eaton WW. Refining acculturation measures for health research: Latina/o heterogeneity in the National Latino and Asian American Study. *Int J Methods Psychiatr Res*. 2020;29:e1844. https://doi.org/10.1002/mpr.1844