Original Research Article

Measuring Anxiety Among Latino Immigrant Populations: Within-Country and Between-Country Comparisons

Sumeyra Sahbaz[1] [ID]**, Ronald B. Cox, Jr.[2]** [ID]**, Pablo Montero-Zamora[1],
Christopher P. Salas-Wright[3], Mildred M. Maldonado-Molina[4],
Melissa M. Bates[4], Augusto Pérez-Gómez[5], Juliana Mejía-Trujillo[5],
Saskia R. Vos[6], Carolina Scaramutti[6], Patrizia A. Perazzo[1], Maria Duque[1],
Maria Fernanda Garcia[3], Eric C. Brown[6], and Seth J. Schwartz[1]**

Abstract

Anxiety is the most prevalent mental health disorder among adults worldwide. Given its increased prevalence among migrants due to their marginalized position in the societies where they reside, psychometric evaluations of anxiety measures such as the Generalized Anxiety Disorder–7 (GAD-7) are needed for use with migrants. The present study is the first attempt to compare the structure of GAD-7 scores for (a) different Latino groups in the same country and (b) the same Latino group in two different countries. Using three samples of Mexican and Venezuelan migrants (total *N* = 933), we provide reliability and validity evidence of the GAD-7 for use with adult Latino migrants. Utilizing confirmatory factor analysis and item response theory, we demonstrate that the GAD-7 is internally consistent, possesses a strong single-factor structure, and generates scores with equivalent psychometric properties. GAD-7 is appropriate for use with Mexican and Venezuelan migrants across differing gender groups and education levels.

Keywords

anxiety, measurement, cross-national, Venezuelan, Mexican

Anxiety disorders affect nearly 40 million U.S. adults every year (Anxiety and Depression Association of America [ADAA], 2022) and represent the most prevalent mental health disorder among adults worldwide (Dattani, 2021). Migrant populations may be particularly vulnerable to anxiety disorders due to often being marginalized in the societies where they reside, the harsh adversities they must overcome (e.g., fear of deportation, language barriers, and lack of resources) in their new homelands, the stressful circumstances that led them to migrate, and leaving behind significant relationships. However, for researchers and practitioners to adequately address mental health needs and evaluate interventions for diverse groups of migrants, they must first ensure that the scores generated by associated measures used are valid indicators of the construct of interest (Arafat et al., 2016).

A growing literature demonstrates that the psychometric properties of scores generated by psychosocial scales can be influenced by aspects of respondents'

cultural context, including regional deviations in language and dialect (Onodera et al., 2005), differences in beliefs and values (Benítez et al., 2016), how questions and response options are construed and understood (Davis et al., 2011), and acculturation levels or profiles (Greenfield, 1997; Tan et al., 2021). Importantly, when cultural context is not properly assessed and accounted for, it can introduce systematic bias into research findings (Choi & Pak, 2005). This need to account for

[1]The University of Texas at Austin, USA
[2]Oklahoma State University, Stillwater, USA
[3]Boston College, Chestnut Hill, MA, USA
[4]University of Florida, Gainesville, USA
[5]Corporación Nuevos Rumbos, Bogotá, Colombia
[6]University of Miami, Coral Gables, FL, USA

Corresponding Author:
Sumeyra Sahbaz, The University of Texas at Austin, 2109 San Jacinto Blvd., Stop D3700, Austin, TX 78712-1415, USA.
Email: sahbaz@utexas.edu

cultural influences suggests that studies assessing the validity of psychosocial measures should consider the inherent cultural differences between individuals from different countries of origin and the characteristics of the countries to which they are migrating. To this end, we test the validity of the Generalized Anxiety Disorder–7 (GAD-7), a widely used measure of anxiety, with three samples of Latino adults (Venezuelans in Colombia, Venezuelans in the United States, and Mexicans in the United States) that vary by national origin, migration experience, and migration destination. Specifically, we compare (a) different migrant groups migrating to the same destination country and (b) the same migrant group migrating to different destination countries.

Background

International migration has reached unprecedented levels, with more than 250 million people currently living in countries other than where they were born (McAuliffe & Triandafyllidou, 2022). The majority of these individuals have migrated from the Global South—Latin America, Asia, Africa, and the Middle East (Abel & Sander, 2014). Some Global South migrants have moved to countries in the Global North—the United States, Canada, Western Europe, Australia, and New Zealand—whereas others have moved to other Global South countries. There are often systemic differences in terms of which Global South migrants have the resources and desire to relocate to a Global North country and which Global South migrants relocate to neighboring Global South countries. For example, although many Syrian civil war survivors migrated to Western Europe, the United States, Canada, and Australia, many more resettled in other Middle Eastern nations such as Turkey and Jordan (Balcilar & Nugent, 2019; Yazgan et al., 2015). Syrians relocating to Germany, for example, may have been wealthier and/or more highly educated compared with those who relocated to neighboring Middle Eastern countries (Bogotch et al., 2020), due to the greater resources needed to migrate to Europe relative to relocating to a neighboring country. Similarly, Venezuelans moving to the United States, Italy, or Spain were generally wealthier and more highly educated compared with those who relocated to Colombia, Peru, or Brazil (López & Patten, 2015; Rueda, 2017).

Of course, not all Global South migrants relocate for the same reasons (Sam & Berry, 2010). For example, Mexicans have relocated to the United States for more than a century, primarily for economic reasons (Henderson, 2011). Mexicans have settled primarily in Texas, California, New Mexico, Arizona, and other southwestern U.S. states—although Mexican migration has begun to spread to states in the Southeast, Midwest, and Plains as well. On the contrary, Venezuelan migration is much more recent, having begun following the election of Hugo Chávez as president in 1998 and escalating after Chávez's death, and Nicolás Maduro's inauguration and his handpicked successor, in 2013 (Salas-Wright et al., 2022). Whereas the original Venezuelan arrivals were wealthy expatriates who settled primarily in South Florida, Venezuelans have now become the fastest-growing Latino group in the United States and are settling in various parts of the country (Moslimani et al., 2023). Venezuela has descended into a national crisis, with starvation, street violence, and government repression reaching epic proportions—and more than 20% of the country's population has emigrated since 2014 (United Nations High Commission for Refugees, 2023).

Largely because they are migrating to escape a crumbling society and an escalating crisis in their homeland, Venezuelans have relocated to multiple destination countries, including the United States, Colombia, and Peru. As noted earlier, Global South migrants who move to a Global North country often differ meaningfully from those who migrate to neighboring Global South countries. It is not known whether mental health problems, which are often elevated among people who migrate to escape natural or human disasters (such as Venezuelans), manifest similarly versus differently among individuals who move to different destinations—and are significantly higher in such "forced" migrants than in other types (Barbot et al., 2020). Such information is needed if we are to develop interventions to address mental health impacts of different types of migration in different destination countries (Vos et al., 2021). Because Mexicans and Venezuelans represent two large Latino groups that migrate for somewhat different reasons, it may be that mental health outcomes—such as anxiety—manifest differently across these two Latino migrant groups.

Accordingly, it is critical to ensure that measures of anxiety—as a primary mental health outcome among migrants—produce scores that are psychometrically equivalent across destination countries. In addition, if we are to examine whether anxiety differs across Latino migrant groups, then measures of anxiety must generate scores that are psychometrically equivalent across migrant groups within a given destination country. Without demonstrating such psychometric equivalence, we do not know the extent to which differences in observed scores represent true mean differences between groups versus representing psychometric differences in

measurement across the two groups (Putnick & Bornstein, 2016).

Within a given pan-ethnic group (e.g., Latinos), different subgroups may migrate for different reasons. For example, many Venezuelans have been displaced by the combination of government repression, street violence, and economic collapse (Salas-Wright et al., 2022). For this reason, Venezuelans have been provided with relief from deportation in multiple destinations such as the United States, Colombia, and Peru (Selee & Bolter, 2022). In contrast, the majority of Mexican migrants to the United States have relocated primarily for economic reasons (Henderson, 2011)—and most Mexicans in the United States are not protected from deportation. It might be expected, then, that Venezuelan and Mexican migrants within the United States (for example) would differ in terms of anxiety, with Venezuelans scoring significantly higher because of the crumbling society from which they have migrated. Provided that the psychometric structure of scores generated by anxiety measures were sufficiently equivalent across these groups, mean comparisons could then be conducted using latent variables that allow any psychometric nonequivalence (differences in factor loadings or item intercepts) to be taken into account (Hancock et al., 2000).

Similarly, it may be important to examine the *same* migrant group across multiple countries (Schwartz et al., 2018). Many groups fleeing natural and human disasters move to multiple countries, such as Venezuelans settling in the United States and in other Latin American countries (Salas-Wright et al., 2022), Syrians migrating to Turkey and Jordan as well as to Germany, and Ukrainians fleeing to Poland as well as to the United States, United Kingdom, and Ireland. In cases such as these, researchers may wish to compare anxiety levels among migrant groups that have moved to different destinations.

The GAD-7 Anxiety Measure in Latino Migrants

Although many measures of anxiety symptoms have been developed, the Generalized Anxiety Disorder–7 (GAD-7; Spitzer et al., 2006) is among the most commonly used and generates scores with high sensitivity and specificity (Williams, 2014). GAD-7 scores have been shown to be structured equivalently and to perform similarly between Latinos and Whites in the United States (Parkerson et al., 2015). Furthermore, the Spanish and English versions of the measure have been shown to yield structurally equivalent scores among U.S. Latino respondents (Mills et al., 2014). The GAD-7 has been used in Latin American samples, including samples of Venezuelan adults in the United States and Colombia (Schwartz et al., 2018) and Mexican adults in

the United States (Ryan et al., 2021). The measure has been used with various Latino groups in the United States, including day laborers (Acosta et al., 2020), primary care patients (Leathers et al., 2021), college students (Suárez-Orozco & López Hernández, 2020), and substance abusers (Levison et al., 2019).

However, to our knowledge, no published work has examined the psychometric properties of scores generated by the GAD-7 across multiple groups of Latino migrants, including those who have migrated to other Latin American countries. Such work is essential if researchers are to use this measure across Latino groups, either within a given country (such as the United States) or across countries. It is also essential to ascertain the extent to which the construct of anxiety carries the same meaning across various Latino groups—which might best be accomplished by comparing psychometric properties across these groups. Given that the GAD-7 is among the most prominent measures of anxiety symptoms among both community and clinical samples, it is both appropriate and important to conduct psychometric comparisons of GAD-7 scores across Latino migrant groups within and across receiving countries.

The Present Study

The purpose of the present study was to use confirmatory factor analysis (CFA) and item response theory (IRT) approaches to offer a comprehensive understanding of the psychometric properties of the GAD-7 scores for Latino migrants residing in the United States. CFA and IRT represent two different psychometric approaches used to analyze and evaluate the psychometric properties of self-report measures. Although CFA and IRT are associated with distinct methodologies and applications, they can complement each other in various ways.

Whereas CFA provides a broad understanding of the psychometric properties of scales, IRT offers specific information about the items that make up the scales. In other words, while the unit of analysis in CFA is the scale, the unit of analysis in IRT is the item. CFA provides initial insights offering reliability estimates and IRT provides information about the measurement precision of each item across the ability continuum, which helps identify the range of ability levels where each item is most informative. This level of information is not provided by CFA. Accordingly, in the present study we used CFA to examine the factor structure of GAD-7 scores and measurement invariance across different migrant groups. We then used IRT to examine item parameters (item discrimination, item difficulty) and item information functions (IFFs). By combining these two approaches, we aimed to offer scale-level and item-

level evaluations of the psychometric properties of scores generated by the GAD-7.

To provide within- and between-country comparisons, we have compared two groups (Venezuelans and Mexicans) residing in one country (the United States) and the same group (Venezuelans) in two different countries (Colombia and the United States). This was done to (a) compare psychological symptoms among individuals from the same country of origin living in different countries and (b) compare the psychological symptoms of individuals from different countries of origin living in the same country. Such comparisons are especially important given Marques et al.'s (2011) finding that individuals residing in non-Western countries may be more likely to experience anxiety physiologically (e.g., as stomach pains) than are individuals living in Western countries. Examining the same migrant group in two locations may help to ascertain the extent to which such variations may create differences in the factor structure of GAD-7 scores, as well as differences in the performance of specific GAD-7 items. To conduct these comparisons, we used measurement invariance tests within the multigroup confirmatory factor analysis (MGCFA) framework and examined equivalence of factor structure (configural invariance), factor loadings (metric invariance), and item intercepts (scalar invariance), on GAD-7 scores.

Specifically, we compared the structure of GAD-7 scores (a) between Mexican and Venezuelan migrants in the United States and (b) between Venezuelan migrants in the United States and Venezuelan migrants in Colombia. In addition to these between-group comparisons, we conducted within-group comparisons of the structure of the GAD-7 scores for males versus females and individuals with different educational levels. These comparisons also allow us to ascertain the extent to which the factor structure of GAD-7 scores (and perhaps of anxiety symptoms themselves) is equivalent between different migrant groups residing in the same country, across groups of migrants who have settled in different destination countries, and across migrant subgroups (males vs. females, individuals with college education vs. individuals with no college education). Next, we evaluated the psychometric properties of responses to the individual items contributing to the GAD-7 utilizing IRT and provided evidence for the rigor of the scale. IRT allowed us to study items' general performance and the extent to which the items were informative across the continuum of the latent trait.

We hypothesized that mean levels of anxiety would be greater among Venezuelans than among Mexicans in the United States, who generally migrate for economic reasons rather than to escape a crumbling society (Henderson, 2011; Salas-Wright et al., 2022). We also hypothesized that anxiety would be greater among Venezuelans in the United States than among their counterparts in Colombia because of the precarious legal status of U.S. Venezuelans prior to March 2021 (when Temporary Protected Status was granted to this group). Venezuelans migrating to the United States have also tended to be educated professionals who were targeted by the Maduro regime, whereas Venezuelans migrating to Colombia have been more likely to be less-educated economic migrants (Salas-Wright et al., 2022). The anxiety associated with opposing the government and having to flee one's country may be greater than the anxiety associated with economic migration (Vos et al., 2021). We did not anticipate any meaningful differences in the factor structure of GAD-7 scores across groups.

Method

Participants and Procedures

We investigated the measurement invariance and psychometric properties of the GAD-7 for use with different Latino migrant groups by analyzing survey responses from three samples. Sample 1 consisted of Venezuelans who migrated to Colombia (n = 305), Sample 2 was made up of Venezuelans who migrated to the United States (n = 342), and Sample 3 consisted of Mexicans who migrated to the United States (n = 286).

Venezuelan migrants were participants in the Colombia and Miami's Newest Arrivals (CAMINAR) study, which collected data from Venezuelan migrants in Bogotá, Colombia, and South Florida during the months of October and November 2017. Participants reported having one or more children below age 18, and the majority of participants in both the United States (82%) and Colombia (75%) had immigrated within the year leading up to the assessment. In both the United States and Colombia, participants were recruited by the study team in collaboration with community leaders. U.S. participants were given a US$40 gift card upon completing the survey and an additional US$15 gift card for each referral they provided as long as the referred participant joined the study (potentially yielding a total incentive of US$85). In Colombia, participants were given a gift card for 75,000 Colombian pesos (approximately US$25) for completing the survey. Referrals were not used in the Colombia sample. The disparity in payments between the United States and Colombia can be attributed to the differing exchange rates of the U.S. dollar in these two countries. The study received approval from the institutional review boards at the University of Miami and Boston University as well as by the ethics board at Corporación Nuevos Rumbos in Bogotá.

Mexican migrants were participants in a study on cultural stress, psychological distress, and alcohol use among Latino parents of children 8 to 16 years residing in the United States. Data were collected in July 2022. Participants were recruited through a web panel, which comprised individuals who had previously given their consent to participate in survey research via their personal accounts with a web panel company. Participants received US$5 for completing the survey. All data were collected using the Qualtrics survey software, and the survey took approximately 30 min to complete. Approval for this study was obtained from the associated university's institutional review board.

The majority of all three samples are female adults with 58%, 66%, and 61%, respectively. While the majority of Venezuelans in Colombia are aged 30 or younger (56%), the majority of Venezuelans and Mexicans in the United States are aged 30 or older (75% and 92%). Less than half of the Venezuelans in Colombia have a college or graduate school degree (45%), whereas this percentage has increased to 65% for Venezuelans in the United States. Mexicans in the United States consisted of individuals with the least amount of education with the majority (63%) not having a college degree. Furthermore, only about 25% of Venezuelans in Colombia were married, whereas the majority of Venezuelans and Mexicans in the United States were married (64% and 72%, respectively). The majority of Venezuelans and Mexicans in the United States (>65%) report that Spanish was their language of preference. This question was not asked to Venezuelans in Colombia for obvious reasons. Most Venezuelans in Colombia (70%) report that they moved to Colombia after 2017. The same percentage of Venezuelans in the United States report that they moved to the United States in 2014. The majority (60%) of Mexicans in the United States had lived in the United States for more than 10 years.

Measures

GAD-7 (Spitzer et al., 2006) was used to measure symptoms of generalized anxiety, and participants were asked to indicate the number of days within the past 2 weeks during which they experienced each of the seven symptoms. The present study used an established Spanish version of the GAD-7 provided by García-Campayo et al. (2010) and validated by Mills et al. (2014) for use with U.S. Latino individuals. Example items include (a) *feeling nervous, anxious, or on edge*; (b) *not being able to stop or control worrying*; and (c) *worrying too much about different things*. Responses were scored on a 4-point Likert-type scale (1 = *not at all*, 2 = *several days*, 3 = *over half the days*, 4 = *nearly every day*) with higher scores indicating greater symptoms. Internal consistency was greater than .80 for all samples (α = .82 for Venezuelans in Colombia, α = .91 for Venezuelans in the United States, and α = .94 for Mexicans in the United States).

Analytic Plan

Analyses for the present study proceeded in three primary steps—preliminary, CFA, and IRT analyses. CFA and IRT are two different statistical approaches used to analyze and evaluate the psychometric properties of scores generated by self-report measures. Although they involve distinct methodologies and applications, CFA and IRT can complement each other in various ways. Whereas CFA provides a broad understanding of the psychometric properties of scale scores, IRT offers specific information about the items that make up the scales. In other words, while the unit of analysis in CFA is the scale, the unit of analysis in IRT is the item. CFA provides initial insights offering reliability estimates, and IRT provides information about the measurement precision of each item across the ability continuum, which helps identify the range of ability levels where each item is most informative. This level of information is not provided by CFA. The present study used CFA to examine the factor structure of GAD-7 scores and measurement invariance across different migrant groups. Furthermore, we used IRT and studied the item parameters (item discrimination, item difficulty) and IFFs. By combining both approaches, we aimed to offer a more comprehensive understanding of the psychometric properties of scores generated by the GAD-7.

Preliminary Analyses. We examined the pattern of interitem correlations to detect possible redundancy among the GAD-7 items and tested the assumption of multivariate normality to decide whether we need to use robust standard errors, the estimator maximum likelihood mean-adjusted (MLM), and the MLM chi-square test statistic Satorra–Bentler χ^2 to adjust for nonnormality.

Confirmatory Factor Analysis. To examine the fit of the GAD-7 items to the proposed one-factor solution, we used CFA and examined the model fit using absolute fit indices such as the chi-square test (χ^2), root mean square error of approximation (RMSEA), and standardized root mean residual (SRMR), as well as relative fit indices such as the Tucker–Lewis index (TLI) and comparative fit index (CFI). SRMR values less than .05, TLI values greater than .90, and RMSEA values less

than .08 indicate acceptable model fit (Hu & Bentler, 1999).

IRT Item Properties

Item Discrimination and Item Difficulty. To further investigate the psychometric properties of the GAD-7, we used IRT. Because IRT is a model-based framework presupposing an underlying latent trait depending on both a person's responses and the item parameters, it allows us to simultaneously estimate item parameters and participants' trait levels (Embretson & Reise, 2000). IRT, also known as latent response theory, refers to a family of generalized linear models explaining the relationship between latent traits (unobservable characteristics or attributes) and their manifestations (observed outcomes or responses). These models establish a link among the properties of items on a measure, individuals responding to these items, and the underlying trait being measured. IRT assumes that the latent construct (e.g., anxiety) and items on a measure are organized on an unobservable continuum. Therefore, the main purpose of IRT is to establish a given individual's position on that continuum. As the IRT model's unit of analysis is the item, IRT can be used to compare item characteristics from one or multiple measures if they are measuring the same latent construct. When analyzing polytomous item responses that are ordered, a graded response model is used (Samejima, 1969, 1997, 2008). Graded response modeling (GRM) is an extension of the dichotomous two-parameter logistic (2pl) IRT model where each item in the scale is indexed as one discrimination parameter (a parameter) and one less than the number of response categories item difficulty parameters (b parameters). In other words, GRM assumes that response choices for the item are ordered, such that (a) higher response indicates a higher level of the trait being assessed, (b) item discrimination is not equal across all items of a scale, and (c) differences between adjacent response categories may not be equivalent across all items (Penfield, 2014).

Item discrimination—also called slope—is a measure of how well an item differentiates between individuals with low levels of the latent trait from individuals with higher levels of the latent trait (Baker, 2001). Item discrimination values theoretically range from $+\infty$ to $-\infty$ but usually do not exceed $+2$. Higher item discrimination values are associated with items that are better able to discriminate individuals based on their latent trait level. In other words, for items with higher discrimination parameters, the probability of choosing a response category will change more rapidly corresponding to changes in the latent trait levels (Hays et al., 2000). Items with negative discrimination values are considered problematic as they suggest that respondents with

increasing levels of the latent trait are less likely to endorse more severe response options. This could potentially happen if an item poorly discriminates between those with high and low levels of anxiety or if a coding error occurred (Yang, 2014). An item discrimination parameter (a) value of 0.65 is considered the minimum threshold for discriminating respondents; values greater than 1.34 are considered high; and values greater than 1.69 are considered very high (Baker, 2001).

Item difficulty—also called between category threshold—can be thought of as "cut points" on the latent trait continuum, where a participant with that level of the latent trait is equally likely to select response category j + 1 rather than category j. A J-point Likert-type scale item will have J − 1 difficulty parameters (b_1, b_2, and b_3 parameters in this study given the 4-point response scale). In other words, item difficulty for a response category represents the trait level necessary to select higher response options with .50 probability. Item difficulty estimates the appropriateness of response options for each item on a scale. They also help identify items for which respondents may not be selecting specific response categories or may be favoring extreme response categories. Most item difficulty estimates fall between -2 and $+2$, with higher values or more extreme responses indicating a higher trait level of selecting higher response options (Hays et al., 2000).

Item Information and Test Information Functions. In IRT, the item information is a function of the latent trait calculated using the participant's responses to the items. IFF is a measure of how much statistical information an item provides, and it is analogous to reliability of measurement indicating the precision of an item along the underlying trait continuum. In other words, the reliability of an item is evaluated by its information function representing the item's ability to differentiate participants at each level of the latent trait (Reise et al., 2005). Therefore, items with high information at a certain level of the latent trait provide precise estimates of person parameters at that level of the latent trait continuum (Baker, 2001). Information functions are summed together building a test information function (TIF) estimating how well the scale functions along the entire latent trait continuum.

Measurement Invariance. In our efforts to offer further validity evidence for the GAD-7 scores, we examined whether the construct measured was equivalent across different groups of individuals. The definition of anxiety or the symptoms associated with it might differ slightly across groups (Cheung & Rensvold, 2002), which then

could lead to a lack of measurement invariance (van de Vijver & Poortinga, 1997). In the absence of measurement invariance, we cannot make meaningful comparisons of different groups of individuals (Horn & McArdle, 1992), and between-group differences might be due to measurement artifacts rather than to real differences in anxiety symptoms across groups (Vandenberg & Lance, 2000). To ensure that GAD-7 scores yield equivalent interpretations across different migrant groups residing in the same country, migrant groups living in different countries, as well as across gender and educational levels within the same migrant group, we conducted measurement invariance tests. These tests were conducted using MGCFA as it represents the most powerful and versatile approach to evaluating measurement invariance (G. T. Brown et al., 2017; Jöreskog & Moustaki, 2001; Steenkamp & Baumgartner, 1998).

To examine measurement invariance in all three samples, we followed the stepwise approach given its ability to identify the point at which invariance is no longer achieved between groups (T. A. Brown, 2015). In the stepwise approach, we impose an increasing number of equality constraints on item factor loadings and item intercepts and test whether each progressively constrained model varies significantly from the previous, less constrained model. At the configural invariance level, factor loadings and item thresholds are allowed to be freely estimated and the number of factors and patterns of indicator-factor are examined. Furthermore, we examined the degree to which the factor structure fits all groups equally well. If and only if, configural invariance is supported across groups, then the next level of measurement invariance may be examined. At the level of metric or weak factorial invariance, item factor loadings are constrained to be the same across groups and the degree to which the model is degraded is examined. This test examined whether the items were associated with comparable relationships to the underlying latent construct (anxiety) across groups. Metric invariance is supported if model fit does not degrade significantly relative to the fit of the previously tested model. If, and only if, metric invariance is supported across groups, then the next step of measurement invariance may be examined. For scalar or strong factorial invariance, item intercepts are constrained in addition to item factor loadings, and the degree to which model fit is degraded is examined. Scalar invariance test allows us to examine whether individuals across groups who fall on the same levels of the underlying latent trait (anxiety) will produce the same raw score. The lack of scalar invariance is known as differential item functioning (McDonald, 1999) and means that individuals in one group yield different raw scores than individuals in another group, despite individuals falling at the same level of the underlying trait. Mplus 8.4 was used to conduct CFA and measurement invariance analyses (Muthén & Muthén, 2017) and Stata 17 was used for all IRT analyses (StataCorp, 2021).

Results

Preliminary Analyses

We report observed means here and latent means later in the results section. Venezuelans in Colombia reported a higher anxiety score on average (1.72) compared with Venezuelans in the United States (1.64). And Mexicans in the United States reported higher anxiety scores on average (1.89) compared with Venezuelans in the United States. For all three samples, the item *worrying too much* had the highest mean response. For Venezuelans in Colombia and the United States, the item *feeling anxious* had the second highest mean response (Table 1). In addition, for all samples, interitem correlations among the GAD-7 items do not exceed ($r > .70$), suggesting that redundancy among GAD-7 items was not a concern. Preliminary analyses indicated a violation of the assumption of normality, so we proceeded to use robust standard errors and the scaled Satorra–Bentler χ^2.

Prior to conducting tests of invariance, we examined the prevalence of multivariate outliers in each of the samples. Multivariate outliers are often identified using Mahalanobis distances (REF), which represent the distance between a given score and the centroid of all scores in the sample. In the two Venezuelan samples, Mahalanobis distances did not identify any significant outliers. In the Mexican sample, Mahalanobis distances identified nine outliers out of 287 cases (3.1% of the sample). Zygmont and Smith (2014) suggest that factor-analytic procedures are generally robust to the presence of outliers as long as the representation of outliers in the sample is small and as long as the estimator used is able to handle deviations from normality. We, therefore, believe that our results are robust and valid.

Factor Structure of the GAD-7

A confirmatory factor model was estimated for the GAD-7 across all three samples. Following the scoring algorithm for the GAD (Spitzer et al., 2006), all seven items were specified as loading on one single factor. For Venezuelans in Colombia, the unifactorial structure provided a moderate fit to the data, CFI = .94, TLI = .92, RMSEA = .08 (90% confidence interval [CI] = [.04, .10]), SRMR = .04, Satorra–Bentler $\chi^2(14, N = 305) = 38.51$, $p = .001$. For Venezuelans in the United States, the unifactorial model provided a good fit to the data,

Table 1. Descriptive Statistics, Internal Consistency Coefficients, and Confirmatory Factor Analysis Parameter Estimates.

Generalized Anxiety Disorder-7	M (SD)	Estimate/factor loadings	SE
Venezuelans in Colombia			
Feeling anxious	1.84 (.85)	.63	.05
Not stopping to worry	1.57 (.85)	.66	.05
Worrying too much	2.45 (.93)	.52	.05
Trouble relaxing	1.51 (.78)	.74	.05
Hard to sit still	1.66 (.81)	.79	.04
Easily irritable	1.43 (.79)	.54	.06
Feeling afraid	1.59 (.79)	.56	.06
Venezuelans in the United States			
Feeling anxious	1.78 (.84)	.80	.03
Not stopping to worry	1.61 (.80)	.84	.03
Worrying too much	2.00 (.92)	.73	.03
Trouble relaxing	1.66 (.83)	.83	.02
Hard to sit still	1.45 (.74)	.79	.03
Easily irritable	1.56 (.77)	.63	.05
Feeling afraid	1.46 (.75)	.77	.04
Mexicans in the United States			
Feeling anxious	1.87 (1.04)	.87	.02
Not stopping to worry	1.85 (1.04)	.87	.02
Worrying too much	2.08 (1.02)	.84	.02
Trouble relaxing	1.94 (1.03)	.88	.02
Hard to sit still	1.82 (1.02)	.82	.03
Easily irritable	1.89 (1.00)	.84	.02
Feeling afraid	1.81 (1.02)	.78	.03

Confirmatory factor model fit	Venezuelans in Colombia	Venezuelans in the United States	Mexicans in the United States
χ^2 (df)	38.51* (14)	24.67* (14)	29.47* (14)
CFI/TLI	.94/.91	.98/.97	.98/.98
RMSEA	.08	.05	.06
RMSEA 90% CI	[.05, .11]	[.01, .08]	[.03, .09]
SRMR	.04	.03	.02
AIC	4,422.26	3,909.26	3,977.69

Note. Prompt of the GAD-7 is: Over the last 2 weeks, how often have you been bothered by the following problems? GAD-7 = Generalized Anxiety Disorder–7; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean residual; AIC = Akaike information criterion.
*p < .05. For all factor loadings, the p value is less than .001.

CFI = .98, TLI = .97, RMSEA = .05 (90% CI = [.01, .08]), SRMR = .03, Satorra–Bentler $\chi^2(14, N = 342) = 24.67$, $p = .04$. Finally, for Mexicans in the United States, the unifactorial solution provided a good fit to one latent factor, CFI = .98, TLI = .98, RMSEA = .06 (90% CI = [.03, .09]), SRMR = .02, Satorra–Bentler $\chi^2(14, N = 286) = 29.47$, $p = .001$ (Table 1).

IRT Item Properties of the GAD-7

Item Parameter Estimates. The discrimination (slope) and difficulty (between category threshold) estimates for the GAD-7 for all three samples appear in Table 2. As the table illustrates, all discrimination parameters were significantly different from 0 indicating that the GAD items are well suited for differentiating between participants who differ in their underlying levels of anxiety. In

addition, the similarity across the seven discrimination parameter estimates supports the use of an unweighted sum scoring procedure for the scale (Cheng et al., 2012). For Venezuelans in Colombia, the item discrimination values vary between 1.27 and 2.95; for Venezuelans in the United States, the range is 1.72 to 3.89. And finally, for Mexicans in the United States, the discrimination values vary between 2.78 and 4.48. For Venezuelans in Colombia, the items *trouble relaxing* and *hard to sit still* had the highest discrimination values; for Venezuelans and Mexicans in the United States, the items with the highest discrimination values were *not stopping to worry* and *trouble relaxing*. Thus, the item *trouble relaxing* was the common item that discriminated the most across all three samples. For Venezuelans in Colombia, the items with the lowest discrimination parameters were *worrying too much* and *feeling afraid*. For Venezuelans in the

Table 2. GAD-7 Item Difficulty and Discrimination Parameter Estimates for All Samples.

Items	Venezuelans in Colombia				Venezuelans in the United States				Mexicans in the United States			
	b_1	b_2	b_3	A	b_1	b_2	b_3	a	b_1	b_2	b_3	a
Over the last 2 weeks, how often have you been bothered by the following problems?												
Feeling anxious	−0.34	1.27	2.34	1.68	−0.16	1.09	1.93	3.23	−0.03	0.78	1.27	3.79
	(.11)	(.16)	(.26)	(.22)	(.08)	(.10)	(.16)	(.37)	(.07)	(.09)	(.11)	(.44)
Not stopping to worry	0.37	1.48	2.17	2.06	0.15	1.23	2.04	3.89	0.03	0.71	1.30	4.41
	(.10)	(.16)	(.23)	(.28)	(.07)	(.10)	(.16)	(.50)	(.07)	(.08)	(.11)	(.53)
Worrying too much	−1.67	0.25	1.65	1.27	−0.54	0.90	1.67	2.31	−0.39	0.59	1.25	3.75
	(.23)	(.12)	(.23)	(.17)	(.10)	(.10)	(.15)	(.25)	(.08)	(.08)	(.11)	(.42)
Trouble relaxing	0.41	1.35	2.35	2.71	0.09	1.07	2.04	3.76	−0.13	0.61	1.35	4.48
	(.90)	(.14)	(.23)	(.40)	(.07)	(.10)	(.17)	(.47)	(.07)	(.08)	(.11)	(.55)
Hard to sit still	0.06	1.28	2.04	2.95	0.50	1.43	2.29	3.36	0.06	0.81	1.43	3.20
	(.08)	(.13)	(.19)	(.42)	(.08)	(.12)	(.21)	(.42)	(.08)	(.09)	(.12)	(.36)
Easily irritable	0.94	1.93	3.08	1.34	0.27	1.76	2.62	1.72	−0.09	0.72	1.48	3.40
	(.15)	(.26)	(.43)	(.21)	(.10)	(.19)	(.29)	(.21)	(.08)	(.09)	(.13)	(.38)
Feeling afraid	0.31	1.73	3.29	1.32	0.47	1.40	2.27	3.00	0.08	0.82	1.48	2.78
	(.12)	(.24)	(.46)	(.20)	(.08)	(.13)	(.21)	(.37)	(.08)	(.10)	(.14)	(.31)

Note. b indicates difficulty parameter, and a indicates discrimination. Standard error estimates appear in parentheses. GAD-7 = Generalized Anxiety Disorder–7.

United States, the items *easily irritable* and *worrying too much* were least discriminative. And finally, for Mexicans in the United States, items *feeling afraid* and *hard to sit still* had the lowest discrimination parameters. Overall, all items have discrimination values much higher than the minimum threshold (.65) for discriminating respondents and the majority of them are considered very highly discriminative with discrimination values greater than 1.69 (Baker, 2001). It is important to note that the discrimination parameters for Venezuelans in Colombia are smaller compared with the two other U.S.-based groups. This diminished discriminative ability of the GAD-7 items for the Venezuelans in Colombia could have occurred because the scale was initially developed for U.S.-based individuals.

Difficulty parameters (or the between category threshold) at b_1 ranged from −1.67 to 0.94 for Venezuelans in Colombia, from −0.54 to 0.50 for Venezuelans in the United States, and from −0.39 to 0.08 for Mexicans in the United States. The item *worrying too much* had the lowest difficulty parameter b_1 for all three samples, meaning that this item requires the lowest anxiety level to select any other option rather than the first response option (not at all) with .50 probability. The difficulty parameters b_2 and b_3 were higher than b_1 for all items across all three samples indicating that individuals with higher anxiety levels selected the more severe response options. In addition, the gradually increasing difficulty parameters for each of the response options across all three samples indicates an absence of redundant response options. The standard errors for item difficulty parameters ranged from .07 to .55 across the three samples, suggesting precision in the estimates

of the difficulty parameters and indicating that participants used the full Likert-type scale range.

Information Functions

Figure 1 illustrates the IFFs as well as the TIFs. The x-axis represents the latent trait level, and the y-axis represents the amount of information provided by the item or the scale. IFFs inform us about items' ability to differentiate participants at each level of the latent trait. Therefore, the higher the item information, the more an item contributes to the overall scale. Figure 1 illustrates that, across all three samples, the items *hard to sit still* and *trouble relaxing* represent the most informative symptoms for individuals across all anxiety levels. However, for Venezuelans in the United States, the item *not stopping to worry* seems to be even more informative than the items *hard to sit still* and *trouble relaxing*. A commonality across all samples is that items in the scale GAD-7 are most informative for individuals with anxiety levels between the mean and 2 standard deviations above the mean. The items that are least informative are different for each sample. For Mexicans in the United States, the item *feeling afraid* is the least informative item, whereas for Venezuelans in Colombia and in the United States, items *easily irritable* and *worry too much* are least informative compared with the other scale items.

TIF is computed as the sum of the IFFs and informs us about how well the scale functions along the anxiety levels. Therefore, the greater the information provided by individual items of a scale, the higher the extent of information provided by the scale. The scale GAD-7 provides the highest amount of information for Mexicans in the
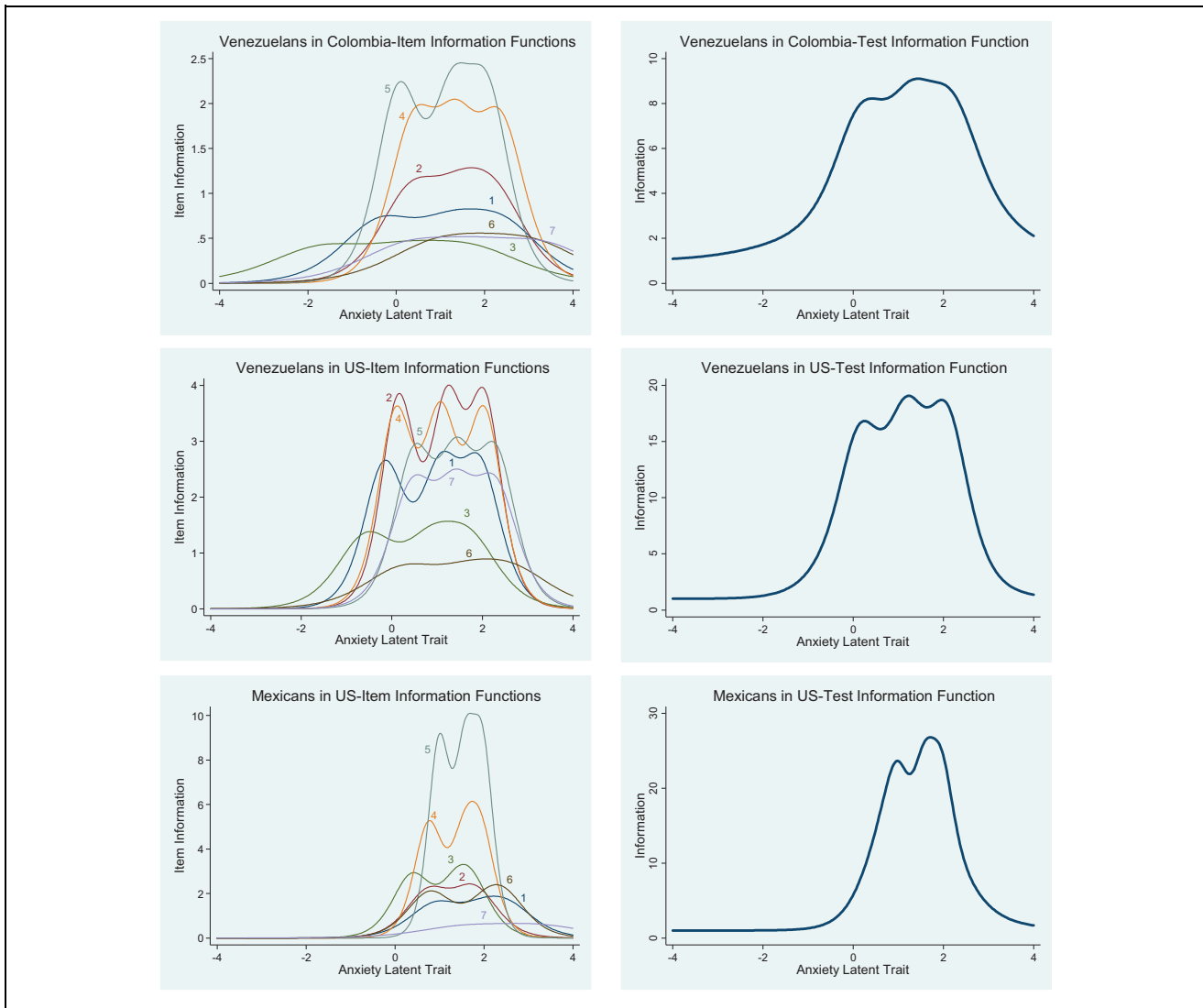
Figure 1. Item Information and Test Information Functions for All Samples.

United States, which is followed by Venezuelans in the United States and Venezuelans in Colombia. Common across all three samples is that the scale GAD-7 is most informative for individuals with anxiety levels between the mean and 2 standard deviations above the mean. Furthermore, the scale provides only a little information for individuals with anxiety levels below the mean or for individuals with very high anxiety levels.

Measurement Invariance of the GAD-7 for Venezuelan Versus Mexican Migrants

Measurement invariance tests were conducted to study between-country and within-country equivalence.

For between-country comparisons, we examined the measurement equivalence of GAD-7 scores between Venezuelans in Colombia and Venezuelans in the

United States using the stepwise approach (configural, metric, and scalar). And in terms of within-country comparisons, we investigated the measurement equivalence of GAD-7 scores between Venezuelans and Mexicans in the United States using the very same stepwise approach. Thus, we have imposed an increasing number of equality constraints on item factor loadings and item intercepts for both between- and within-country comparisons. We have assessed the discrepancies in the goodness of fit of the various models by examining the differences between chi-square statistics and other fit indices such as CFI, RMSEA, and SRMR. Following the guidelines of Chen (2007), we regarded invariance as established if (a) the difference between the chi-square statistics was not significant as indicated by the likelihood ratio test and if (b) the differences in CFI (ΔCFI)

Table 3. GAD-7 Between-Country and Within-Country Tests of Measurement Invariance.

GAD-7	χ^2 (df)	CFI/TLI	RMSEA	RMSEA 90% CI	SRMR	p value (χ^2-diff test)
Between-country						
Configural invariance	60.14* (28)	.97/.96	.06	[.04, .08]	.04	—
Metric invariance	70.65* (34)	.97/.96	.06	[.04, .08]	.05	.12
Scalar invariance	140.07* (40)	.91/.91	.09	[.08, .11]	.07	.00
Within-country						
Configural invariance	48.77* (28)	.99/.98	.05	[.03, .07]	.03	—
Metric invariance	60.08* (34)	.98/.98	.06	[.04, .08]	.05	.06
Scalar invariance	84.97* (40)	.97/.97	.07	[.05, .09]	.06	.00

Note. GAD-7 = Generalized Anxiety Disorder–7; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; *p < .001; between-country refers to the comparison of Venezuelans in the United States versus Colombia; within-country refers to the comparison of two Hispanic samples (Venezuelans vs. Mexicans) in the United States.

between models were smaller than .015, ΔRMSEA was smaller than .024, and ΔSRMR was smaller than .036.

Results of the between-country and within-country measurement invariance tests are presented in Table 3. For both comparisons, results indicated that we have configural invariance, meaning that the number of factors and the patterns of item-factor relationships are consistent for Venezuelans in Colombia versus Venezuelans in the United States as well as for Venezuelans versus Mexicans in the United States. After constraining item factor loadings (metric invariance test), we observed that there are no significant differences in item factor loadings for both between- and within-country comparisons as indicated by the insignificant chi-square difference test results and the nonsignificant differences between the other fit indices (ΔCFI < .015, ΔRMSEA < .024, and ΔSRMR < .036). Thus, regression slopes are consistent, and we achieve full metric invariance both between Venezuelan migrants in the United States and Colombia and across the two migrant groups in the United States. However, when we constrained the item intercepts, the model fit was significantly different from the previous model. Thus, the results do not support scalar invariance for between- or within-country comparisons. In other words, we found partial measurement equivalence for the GAD-7 when used with Venezuelan migrants living in two different countries and when used with Mexican versus Venezuelan migrants living in the same country.

Because the fit of the GAD single-factor algorithm was somewhat poorer for Venezuelans in Colombia than for the two groups in the United States, we examined modification indices to identify pairs of items for which residual terms remained significantly correlated even after common variance associated with the latent GAD factor was removed from consideration. Such residual correlations are useful to note because they represent pairs of items for which participants responded especially similarly—perhaps reflecting cultural differences

in how people may respond to the GAD items. Two residual correlations were identified within the sample of Venezuelans in Colombia: (a) *feeling anxious* with *worry too much* and (b) *trouble relaxing* and *hard to sit still*.

To examine population heterogeneity, we conducted latent mean comparisons using the scalar invariance model. Following Hancock et al., (2000), the latent mean for a reference group is constrained to zero, and the latent means for each of the other groups are freely estimated. The latent mean and variance for each group can be used to conduct pairwise *t* tests and to compute Cohen's *d* effect sizes.

Using this method, we found that Venezuelans in the United States versus Colombia did not differ significantly on anxiety, t(592) = 0.82, p = .41, d = .07. However, Venezuelans in both countries were significantly less anxious compared with Mexicans in the United States: Venezuelans in Colombia versus Mexicans in the United States, t(574) = 3.73, p < .001, d = .31; Venezuelans in the US versus Mexicans in the United States, t(590) = 4.19, p < .001, d = .34.

Next, we examined within-group invariance across gender and educational levels for each of the three samples. For gender, chi-square difference tests and differences in the other fit indices revealed no significant differences in model fit after constraining the number of factors, factor loadings, or item intercepts, indicating scalar variance for gender across all three samples. This suggests that comparisons between male and female participants are appropriate for all samples (see Table 4). Similarly for educational levels, we compared participants with less than a college degree with those with college or graduate degrees. Due to the lower overall educational level among Mexicans in the United States, the educational levels compared within this group were less than college versus some college education or more. Measurement invariance tests across educational levels indicated that we have scalar invariance for both Venezuelans in Colombia and in the United States while we have metric invariance for Mexicans in the United States.

Table 4. Tests of Measurement Invariance of the GAD-7 Across Gender Groups and Education Levels.

Tests of measurement invariance across gender groups

GAD-7	χ^2 (df)	CFI/TLI	RMSEA	RMSEA 90% CI	SRMR	p value (χ^2-diff test)
Venezuelans in Colombia						
Configural invariance	56.65* (28)	.93/.90	.08	[.05, .12]	.05	—
Metric invariance	63.93* (34)	.93/.91	.08	[.05, .11]	.07	.24
Scalar invariance	69.64* (40)	.93/.92	.07	[.04, .09]	.07	.63
Venezuelans in the United States						
Configural invariance	41.23 (28)	.98/.97	.06	[.00, .09]	.04	—
Metric invariance	48.75 (34)	.98/.97	.05	[.00, .08]	.06	.31
Scalar invariance	58.90 (40)	.97/.97	.06	[.02, .08]	.06	.08
Mexicans in the United States						
Configural invariance	48.42* (28)	.98/.97	.07	[.04, .11]	.03	—
Metric invariance	57.19 (34)	.98/.97	.07	[.04, .10]	.05	.24
Scalar invariance	66.05 (40)	.98/.98	.07	[.04, .10]	.05	.21

Tests of measurement invariance across education levels

GAD-7	χ^2 (df)	CFI/TLI	RMSEA	RMSEA 90% CI	SRMR	p value (χ^2-diff test)
Venezuelans in Colombia						
Configural invariance	55.96* (28)	.93/.90	.08	[.05, .11]	.05	—
Metric invariance	58.54* (34)	.94/.93	.07	[.04, .10]	.06	.78
Scalar invariance	66.44* (40)	.94/.94	.07	[.04, .10]	.06	.30
Venezuelans in the United States						
Configural invariance	37.59 (28)	.98/.98	.05	[.00, .08]	.03	—
Metric invariance	44.38 (34)	.98/.98	.04	[.00, .08]	.06	.37
Scalar invariance	49.45 (40)	.98/.98	.04	[.00, .07]	.06	.80
Mexicans in the United States						
Configural invariance	43.98* (28)	.99/.98	.06	[.02, .09]	.02	—
Metric invariance	52.46* (34)	.98/.98	.06	[.02, .09]	.05	.24
Scalar invariance	65.25* (40)	.98/.98	.07	[.03, .10]	.05	.02

Note. GAD-7 = Generalized Anxiety Disorder–7; CFI = comparative fit index; TLI = Tucker–Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; *$p < .05$; Venezuelans in Colombia (n_{male} = 123, n_{female} = 166), Venezuelans in the United States (n_{male} = 105, n_{female} = 199); Venezuelans in Colombia ($n_{less\ than\ college}$ = 152, $n_{college\ or\ more}$ = 137), Venezuelans in the United States ($n_{less\ than\ college}$ = 107, $n_{college\ or\ more}$ = 192), Mexicans in the United States ($n_{less\ than\ college}$ = 178, $n_{some\ college\ or\ more}$ = 99).

Discussion

The GAD-7 scale is one of the most widely used measures to assess anxiety among both clinical and general populations. Recently, researchers have started using this measure to assess migrants' anxiety levels (Parkerson et al., 2015; Schwartz et al., 2018). The aim of the current project, which has not been addressed in other studies to our knowledge, was to validate the GAD-7 for use with different Latino migrant groups by systematically examining reliability, validity, item characteristics, and measurement invariance across gender and education level groups.

In the present study, we evaluated the psychometric properties and measurement equivalence of GAD-7 scores using three samples of Latino migrants residing in Colombia and the United States. This design allowed us to compare two different Latino migrant groups (Venezuelans and Mexicans) in the same

country, as well as the same Latino migrant group (Venezuelans) in two different countries (United States and Colombia).

GAD-7 scores fit well into a single-factor structure with excellent psychometric properties and evidence for metric invariance across different Latino migrant groups—although the model fit for Venezuelans in Colombia was slightly poorer than that for the two U.S.-based groups. We also found scalar invariance across gender and educational levels, with the exception of the comparison between Mexicans and Venezuelans residing in the United States that only rose to the level of metric invariance.

IRT results indicated that, across all three samples, all seven GAD-7 items yielded positive discrimination values, indicating that all the items are well suited for differentiating between participants who differ in their underlying levels of anxiety. For Venezuelans and

Mexicans in the United States, the same items (*not stopping to worry* and *trouble relaxing*) yielded the highest discrimination values. In addition to *trouble relaxing*, the item *hard to sit still* appeared to provide the greatest extent of discrimination for Venezuelans in Colombia. In addition, for all three samples and for all GAD-7 items, item difficulty parameters increased for more severe response options, indicating that severe response options are selected by participants with higher levels of anxiety. IFFs indicated that the item *hard to sit still* was the most informative item for Venezuelans in Colombia and for Mexicans in the United States, whereas the item *not stopping to worry* was most informative for Venezuelans in the United States. As Marques et al. (2011) have found, some anxiety symptoms, such as excessive worrying and trouble relaxing, tend to be prominent across cultural groups. On the contrary, non-Western groups may be more likely to experience physiological sensations related to anxiety, such as stomach pains. The fact that the Mexican sample had lived in the United States longer may have been at least somewhat responsible for the lack of invariance across samples.

The present study contributes to the literature by providing considerable evidence that the GAD-7 captures the same meaning across various Latin American migrant groups. Studies such as ours are crucial if researchers and practitioners are to have maximal confidence in using the GAD-7 across Latino groups, either within a given country (such as the United States) or across countries. Without evidence of measurement equivalence across groups, researchers and practitioners run the risk of making meaningless comparisons across different groups which could lead to inaccurate evaluations of prevention and treatment interventions. Or they may be assumed ineffective when in reality they are effective, which runs the risk of wasting valuable resources and time.

Limitations

The present findings, vis-à-vis examining the GAD-7 for use with different groups of Latino adult migrants, should be understood in light of several limitations. First, Latino migrants are a heterogeneous population. Our study was conducted using two samples of Venezuelan migrants and one sample of Mexican migrants, which may not generalize to other subsets of Latino migrants. Second, our study was cross-sectional, and as such, studies tracking patterns of interindividual variability and intraindividual change using longitudinal methods may represent an important future advance using the GAD-7 with Latino migrant samples. Third, the majority of the Mexican sample had been in the United States for more than 10 years, whereas most of the Venezuelan participants in the United States

and in Colombia had migrated within the 3 years prior to assessment. These differences in length of stay, along with associated experiences that may have occurred during longer or shorter stays, may have confounded the present results. Fourth, we did not include comparison variables, such as depressive symptoms, self-esteem, or coping strategies. Such comparison variables would have been important in gauging the equivalence of convergent and construct validity across the three samples. Fourth, data collection between our Venezuelan and Mexican samples has a 5-year difference. Geopolitical changes after November 2017, including the Covid-19 pandemic and the 2020 U.S. presidential elections, might have affected the perception of U.S. migrants' anxiety in a way only captured by our Mexican participants. Fifth, participants were not asked about their reasons for migration, resources they were able to access during the migration transition, or their migration status. Sixth, the results of the present study are limited by its reliance on a single measure (i.e., mono-method bias). It is important to utilize multiple data sources to reduce mono-method bias (Donaldson, 1995; Donaldson & Grant-Vallone, 2002; Shadish, 1993). Finally, the present study is limited to measurement invariance tests across gender and educational levels. Although some Mexican participants completed the survey in English, all Venezuelan participants completed the survey in Spanish. We were, therefore, unable to conduct measurement invariance tests across language of assessment.

In conclusion, the GAD-7 scale provides a robust assessment of migrants' anxiety symptoms and allows meaningful comparisons of anxiety across groups of Latino migrants that can be used in research and practice settings. Unlike other measures, GAD-7 is brief, simple to administer, and is appropriate for use with Latino adult migrants across differing educational levels and gender groups. We hope that the present results will inspire more work in this direction.

ORCID iDs

Sumeyra Sahbaz (iD) https://orcid.org/0000-0002-3265-5023
Ronald B. Cox, Jr. (iD) https://orcid.org/0000-0001-7389-4218

References

Abel, G. J., & Sander, N. (2014). Quantifying global international migration flows. *Science*, *343*(6178), 1520–1522.

Acosta, L. M., Andrews, A. R. III, Acosta Canchila, M. N., & Ramos, A. K. (2020). Testing traditional machismo and the gender role strain theory with Mexican migrant farmworkers. *Hispanic Journal of Behavioral Sciences*, *42*(2), 215–234.

Anxiety and Depression Association of America. (2022). *Facts & statistics*. https://adaa.org/understanding-anxiety/facts-statistics#Facts%20and%20Statistics

Arafat, S., Chowdhury, H. R., Qusar, M., & Hafez, M. (2016). Cross cultural adaptation & psychometric validation of research instruments: A methodological review. *Journal of Behavioral Health*, *5*(3), 129–136.

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse.

Balcilar, M., & Nugent, J. B. (2019). The migration of fear: An analysis of migration choices of Syrian refugees. *The Quarterly Review of Economics and Finance*, *73*, 95–110.

Barbot, B., Hein, S., Trentacosta, C., Beckmann, J. F., Bick, J., Crocetti, E., Liu, Y., Rao, S. F., Liew, J., & Overbeek, G. (2020). Manifesto for new directions in developmental science. *New Directions for Child and Adolescent Development*, *2020*(172), 135–149.

Benítez, I., He, J., Van de Vijver, F. J., & Padilla, J. L. (2016). Linking extreme response style to response processes: A cross-cultural mixed methods approach. *International Journal of Psychology*, *51*(6), 464–473.

Bogotch, I., Pappas, D., Kervin, C., & Silliman, E. (2020). Towards a socially just system of newcomer school integration: Syrians in Canada and Germany. In R. Papa (Ed.), *Handbook on promoting social justice in education* (pp. 2243–2263). Springer.

Brown, G. T., Harris, L. R., O'Quin, C., & Lane, K. E. (2017). Using multi-group confirmatory factor analysis to evaluate cross-cultural research: Identifying and understanding non-invariance. *International Journal of Research & Method in Education*, *40*(1), 66–90.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504.

Cheng, Y., Yuan, K.-H., & Liu, C. (2012). Comparison of reliability measures under factor analysis and item response theory. *Educational and Psychological Measurement*, *72*(1), 52–67.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233–255.

Choi, B. C., & Pak, A. W. (2005). Peer reviewed: A catalog of biases in questionnaires. *Preventing Chronic Disease*, *2*(1), 1–13.

Dattani, S., Ritchie, H., & Roser, M. (2021). *Mental health*. Our World in Data. https://ourworldindata.org/mental-health

Davis, R. E., Resnicow, K., & Couper, M. P. (2011). Survey response styles, acculturation, and culture among a sample of Mexican American adults. *Journal of Cross-Cultural Psychology*, *42*(7), 1219–1236.

Donaldson, S. I. (1995). Worksite health promotion: A theory-driven, empirically based perspective. In L. R. Murphy, J. J. Hurrel, S. L. Sauter, & G. P. Keita (Eds.), *Job stress interventions* (pp. 73–90). American Psychological Association.

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of Business and Psychology*, *17*, 245–260.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory*. Psychology Press.

García-Campayo, J., Zamorano, E., Ruiz, M. A., Pardo, A., Pérez-Páramo, M., López-Gómez, V., Freire, O., & Rejas, J. (2010). Cultural adaptation into Spanish of the Generalized Anxiety Disorder–7 (GAD-7) scale as a screening tool. *Health and Quality of Life Outcomes*, *8*(1), 1–11.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist*, *52*(10), 1115–1124.

Hancock, G. R., Lawrence, F. R., & Nevitt, J. (2000). Type I error and power of latent mean methods and MANOVA in factorially invariant and noninvariant latent variable systems. *Structural Equation Modeling*, *7*(4), 534–556.

Hays, R. D., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurement in the 21st century. *Medical Care*, *38*(9 Suppl.), II28–II42.

Henderson, T. J. (2011). *Beyond borders: A history of Mexican migration to the United States* (*Vol. 17*). Wiley.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*(3), 117–144.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55.

Jöreskog, K. G., & Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, *36*(3), 347–387.

Leathers, C., Kroenke, K., Flanagan, M., Diaz, S., Gruber, R., Tran, G., & Driver, D. (2021). Somatic, anxiety, and depressive (SAD) symptoms in young adult Latinx immigrants: Prevalence and predictors. *Journal of Immigrant and Minority Health*, *23*(5), 956–964.

Levison, J. H., Alegría, M., Wang, Y., Markle, S. L., Fuentes, L., Mejia, D. L., Tarbox, A., Albarracín García, L., Cellerino, L., & El-Bassel, N. (2019). High HIV/STI test acceptance through a behavioral health encounter in Latino immigrants with substance use and mental health problems. *AIDS and Behavior*, *23*(4), 835–846.

López, G., & Patten, E. (2015, September 15). *The impact of slowing immigration: Foreign-born share falls among 14 largest US Hispanic origin groups*. https://policycommons.net/artifacts/618888/the-impact-of-slowing-immigration/1599928/

Marques, L., Robinaugh, D. J., LeBlanc, N. J., & Hinton, D. (2011). Cross-cultural variations in the prevalence and presentation of anxiety disorders. *Expert Review of Neurotherapeutics*, *11*(2), 313–322.

McAuliffe, M., & Triandafyllidou, A. (2022). 1 Report Overview: Technological, Geopolitical and Environmental Transformations Shaping Our Migration and Mobility Futures. *World Migration Report*, *2022*(1), e00022.

McDonald, R. (1999). *Test theory: A unified treatment*. Psychology Press.

Mills, S. D., Fox, R. S., Malcarne, V. L., Roesch, S. C., Champagne, B. R., & Sadler, G. R. (2014). The psychometric properties of the Generalized Anxiety Disorder–7 scale in Hispanic Americans with English or Spanish language preference. *Cultural Diversity and Ethnic Minority Psychology*, *20*(3), 463–468.

Moslimani, M., Lopez, M., & Noe-Bustamante, L., (2023). 11 facts about Hispanic origin groups in the U.S., Pew Research Center. *United States of America*. Retrieved January 5, 2024, from https://policycommons.net/artifacts/4752726/11-facts-about-hispanic-origin-groups-in-the-us/5577372/

Muthén, B., & Muthén, L. (2017). Mplus. In W. J. van der Linden (Ed.), *Handbook of item response theory* (pp. 507–518). Chapman and Hall/CRC.

Onodera, N., Smith, T., Harkness, J., & Mohler, P. (2005). Methods for assessing and calibrating response scales across countries and languages. *Comparative Sociology*, *4*(3–4), 365–415.

Parkerson, H. A., Thibodeau, M. A., Brandt, C. P., Zvolensky, M. J., & Asmundson, G. J. (2015). Cultural-based biases of the GAD-7. *Journal of Anxiety Disorders*, *31*, 38–42.

Penfield, R. D. (2014). An NCME instructional module on polytomous item response theory models. *Educational Measurement: Issues and Practice*, *33*(1), 36–48.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90.

Reise, S. P., Ainsworth, A. T., & Haviland, M. G. (2005). Item response theory: Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*, *14*(2), 95–101.

Rueda, M. (2017). *Colombia's Venezuela problem: The relentless influx of desperate Venezuelans is placing a huge strain on its western neighbor*. https://www.thenewhumanitarian.org/feature/2017/12/01/colombia-s-venezuela-problem

Ryan, D., Tornberg-Belanger, S. N., Perez, G., Maurer, S., Price, C., Rao, D., Chan, K. C., & Ornelas, I. J. (2021). Stress, social support and their relationship to depression and anxiety among Latina immigrant women. *Journal of Psychosomatic Research*, *149*, 110588.

Salas-Wright, C. P., Maldonado-Molina, M. M., Pérez-Gómez, A., Trujillo, J. M., & Schwartz, S. J. (2022). The Venezuelan diaspora: Migration-related experiences and mental health. *Current Opinion in Psychology*, *47*, 101430.

Sam, D. L., & Berry, J. W. (2010). Acculturation: When individuals and groups of different cultural backgrounds meet. *Perspectives on Psychological Science*, *5*(4), 472–481. https://doi.org/10.1177/1745691610373075

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, *34*(4 Part 2), 100.

Samejima, F. (1997). Graded response model. In W. J. van der Linden (Ed.), *Handbook of modern item response theory* (pp. 85–100). Springer.

Samejima, F. (2008). Graded response model based on the logistic positive exponent family of models for dichotomous responses. *Psychometrika*, *73*, 561–578.

Schwartz, S. J., Salas-Wright, C. P., Pérez-Gómez, A., Mejía-Trujillo, J., Brown, E. C., Montero-Zamora, P., Meca, A., Scaramutti, C., Soares, M. H., Vos, S. R., Javakhishvili, N., & Dickson-Gomez, J. (2018). Cultural stress and psychological symptoms in recent Venezuelan immigrants to the United States and Colombia. *International Journal of Intercultural Relations*, *67*, 25–34. https://doi.org/10.1016/j.ijintrel.2018.09.001

Selee, A., & Bolter, J. (2022). Colombia's open-door policy: An innovative approach to displacement? *International Migration*, *60*(1), 113–131.

Shadish, W. R. (1993). Critical multiplism: A research strategy and its attendant tactics. *New Directions for Program Evaluation*, *60*, 13–57.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097.

StataCorp LLC. (2021). *Stata Statistical Software: Release 17*.

Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, *25*(1), 78–90.

Suárez-Orozco, C., & López Hernández, G. (2020). "Waking up every day with the worry": A mixed-methods study of anxiety in undocumented Latinx college students. *Frontiers in Psychiatry*, *11*, 568167. https://doi.org/10.3389/fpsyt.2020.568167

Tan, Y. W., Burgess, G. H., & Green, R. J. (2021). The effects of acculturation on neuropsychological test performance: A systematic literature review. *The Clinical Neuropsychologist*, *35*(3), 541–571.

United Nations High Commissioner for Refugees (UNHCR), (2023). *Venezuela situation*. Retrieved January 5, 2024, from https:// www.unhcr.org/venezuela-emergency.html

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70.

Van de Vijver, F. J., & Poortinga, Y. H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, *13*(1), 29–37.

Vos, S. R., Clark-Ginsberg, A., Puente-Duran, S., Salas-Wright, C. P., Duque, M. C., Herrera, I. C., Maldonado-Molina, M. M., Castillo, M. N., Lee, T. K., Garcia, M. F., Fernandez, C. A., Hanson, M., Scaramutti, C., & Schwartz, S. J. (2021). The family crisis migration stress framework: A framework

to understand the mental health effects of crisis migration on children and families caused by disasters. *New Directions for Child and Adolescent Development*, *2021*(176), 41–59.

Williams, N. (2014). The GAD-7 questionnaire. *Occupational Medicine*, *64*(3), 224–224.

Yang, F. M. (2014). Item response theory for measurement validity. *Shanghai Archives of Psychiatry*, *26*(3), 171–177.

Yazgan, P., Utku, D. E., & Sirkeci, I. (2015). Syrian crisis and migration. *Migration Letters*, *12*(3), 181–192.

Zygmont, C., & Smith, M. R. (2014). Robust factor analysis in the presence of normality violations, missing data, and outliers: Empirical questions and possible solutions. *Quantitative Methods for Psychology*, *10*(1), 40–55.