RESEARCH ON TRANSLATIONS OF TESTS:

Spanish Translation and Validation of the Interpersonal Needs Questionnaire

Caroline Silva,
University of Rochester School of Medicine and Dentistry

Gabriela Hurtado,
University of Toledo

Chelsey Hartley,
Florida International University

José N. Rangel,
Universidad Nacional Autónoma de México

Joseph D. Hovey,
University of Texas Rio Grande Valley

Jeremy W. Pettit,
Florida International University

Paloma Chorot,
Universidad Nacional de Educación a Distancia, Spain

Rosa M. Valiente,
Universidad Nacional de Educación a Distancia, Spain

Bonifacio Sandín, and
Universidad Nacional de Educación a Distancia, Spain

Thomas E. Joiner
Florida State University

Abstract

The present study reports the multistage development and evaluation of a Spanish translation of the Interpersonal Needs Questionnaire (INQ). The INQ measures the constructs of thwarted belongingness and perceived burdensomeness, which the interpersonal theory of suicide proposes are proximal causes of suicidal desire. Participants were bilingual Hispanic college students in the United States ($n = 56$), heritage Spanish-speaking college students in the United States ($n = 281$),

Correspondence concerning this article should be addressed to Caroline Silva, Department of Psychiatry, University of Rochester School of Medicine and Dentistry, 300 Crittenden Boulevard, Box PSYCH, Rochester, NY 14642. caroline_silva@urmc.rochester.edu.
Caroline Silva, Department of Psychiatry, University of Rochester School of Medicine and Dentistry; Gabriela Hurtado, Department of Psychology, University of Toledo; Chelsey Hartley, Department of Psychology, Florida International University; José N. Rangel, Department of Psychology, Universidad Nacional Autónoma de México; Joseph D. Hovey, Department of Psychological Science, University of Texas Rio Grande Valley; Jeremy W. Pettit, Department of Psychology, Florida International University; Paloma Chorot, Rosa M. Valiente, and Bonifacio Sandín, Department of Psychology,
Gabriela Hurtado is now in private practice, Austin, Texas, and Chelsey Hartley is now at University of Rochester School of Medicine and Dentistry.

college students in Spain ($n = 1,016$), psychiatric inpatients in Mexico ($n = 181$), college students in Mexico ($n = 239$), and Spanish-speaking U.S. adults ($n = 104$). Results indicated that a 9-item 2-factor solution (INQ-S-9) provided good fit. Multiple group analyses were also consistent with measurement invariance across nationalities and clinical severity. Finally, both subscale scores demonstrated good internal consistency, test–retest reliability, convergent validity, and concurrent associations with scores on measures of suicide ideation. Cultural considerations and implications for use in clinical and research settings are discussed.

Keywords

Interpersonal Needs Questionnaire; Spanish; thwarted belongingness; perceived burdensomeness; suicide ideation

Historically, compared with other racial or ethnic groups, Hispanics in the United States have been at decreased risk for suicide ideation, attempts, and death (Centers for Disease Control and Prevention, 2017; Hoppe & Martin, 1986; Oquendo et al., 2001; Sorenson & Golding, 1988). However, research has indicated a drastic rise in suicide attempts and rates among Hispanics in the last decade (Centers for Disease Control and Prevention, 2017; Oquendo, Lizardi, Greenwald, Weissman, & Mann, 2004; Zayas, Lester, Cabassa, & Fortuna, 2005). Although studies examining suicidal behaviors among Hispanics are limited, Hispanics are more likely to experience depressive disorders (Limon, Lamson, Hodgson, Bowler, & Saeed, 2016) and tend to have more persistent psychiatric disorders (Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005) than non-Hispanic Whites. Importantly, depression and anxiety are related to increased risk for suicide ideation and attempts among Hispanics (Fortuna, Perez, Canino, Sribney, & Alegria, 2007).

In regard to culturally relevant factors, although family connectedness can ameliorate risk for Hispanics, family conflict can increase psychological distress (e.g., Rivera et al., 2008; Ryan, Huebner, Diaz, & Sanchez, 2009). As such, measures are needed to examine interpersonal *proximal* risk factors for suicide among Hispanics (especially Spanish speakers), which may better inform intervention efforts. The present study is comprised of a series of studies focused on the Spanish translation and validation of the Interpersonal Needs Questionnaire (INQ), a measure designed to assess thwarted belongingness and perceived burdensomeness, two well-established robust proximal suicide risk factors (Van Orden et al., 2010).

The paucity of studies related to suicide among Hispanic populations is a concern given that Hispanics are the largest ethnic or racial minority group in the United States (Ennis, Rio-Vargas, & Albert, 2011). Linguistic barriers contribute to this lack of research. Of Hispanics aged 5 years and older in the United States, approximately 74% speak Spanish at home (Ryan, 2013). Despite Spanish being the second most widely spoken language in the world, with 400 million native speakers and official status in 21 countries, there remains a lack of common psychological measures available in standardized Spanish.

In the United States alone, approximately 34% of Hispanics do not speak English proficiently (Ryan, 2013). In some samples, 42% to 50% of Mexican American adolescents

denied speaking "mostly only English" (Tortolero & Roberts, 2001). Yet studies examining suicide ideation among Hispanics have excluded participants who have not mastered English or presumed them to be bilingual (e.g., Rasmussen, Negy, Carlson, & Burns, 1997; Roberts & Chen, 1995). A number of studies also fail to consider language fluency or preference among Hispanic participants when assessing for suicide ideation or history of attempts using English language measures, potentially resulting in underreporting (Ungemack & Guarnaccia, 1998).

Differing rates of suicide across Hispanic populations (see Silva & Van Orden, 2018, for a review) highlight the importance of understanding culturally or ethnically based protective factors that may deter or increase risk for suicide. However, a viable measure to assess well-known theory-based suicide risk factors (i.e., thwarted belongingness and perceived burdensomeness) with Spanish-speaking individuals does not exist. Such a measure could promote cross-cultural comparison with studies conducted in foreign countries as well as improve suicide risk assessments among Spanish speakers, which then can inform treatment targets that alleviate risk.

Theories of Suicide

Identifying theory-based risk and protective factors for suicidal behavior can inform the development of effective suicide screening and prevention interventions. Zayas et al. (2005) provided a framework for understanding suicidal behaviors among Latina adolescents. This conceptual model emphasizes sociocultural factors that either increase or ameliorate suicidal behaviors. The authors noted that certain cultural factors, such as family functioning, are particularly relevant contributors to the experience of distress among Hispanic adolescents. Although knowledge of societal risk and protective factors can enhance preventative interventions at the community level, identifying mechanisms that contribute to immediate suicide risk at the individual level is necessary for effective suicide screening, assessment, and management. Given the emphasis on interpersonal relationships and the serious impact that conflict can cause in the well-being of Hispanic individuals, the interpersonal theory of suicide (Joiner, 2005; Van Orden et al., 2010) can be particularly relevant for identifying Hispanic individuals that are more immediately at risk for suicide.

The interpersonal theory of suicide posits that an individual will only engage in fatal (or near-fatal) suicidal behavior when they have both the desire to die by suicide and the capability to do so. The theory holds that active suicidal desire arises from the simultaneous presence of thwarted belongingness and perceived burdensomeness. Thwarted belongingness is characterized by feelings of loneliness and lack of reciprocal care. Perceived burdensomeness is characterized by self-hatred and feeling like a liability to others. The theory goes on to state that acquired capability for suicide (i.e., fearlessness of death and elevated physical pain tolerance) is necessary for lethal or near-lethal suicidal behavior to occur. Empirical examinations of the theory's hypotheses regarding suicidal desire have found associations between these constructs and suicide ideation and behavior (see Joiner, Ribeiro, & Silva, 2012). The joint experience of thwarted belongingness and perceived burdensomeness has been found to confer greater risk for suicide ideation, even

after controlling for strong covariates, including depressive symptom severity (Joiner et al., 2009; Van Orden, Witte, Gordon, Bender, & Joiner, 2008).

The interpersonal theory of suicide may elucidate the proximal mechanisms by which some Hispanics are at decreased versus increased risk for suicide. For example, acculturation has been associated with suicide ideation, attempts, and fatalities among Hispanics (Rasmussen et al., 1997; Wadsworth & Kubrin, 2007). Acculturative stress can lead to social group disintegration, role conflicts, and threats to stability and durability of social relationships (Ungemack & Guarnaccia, 1998). Such disintegration could lead to feelings of thwarted belongingness (e.g., because of resulting alienation, isolation, loneliness) as well as perceived burdensomeness (e.g., feeling like a burden to one's ethnic or cultural group), thus resulting in suicidal desire.[1]

Previous research has found that family conflict—a robust predictor of suicidal behaviors across U.S. groups—has been associated with an increased sense of burdensomeness (Van Orden et al., 2010). Perceived burdensomeness resulting from family conflict might particularly contribute to increased suicide risk among Hispanics. Of note, perceived burdensomeness has been a salient predictor of suicide ideation among Hispanics (Garza & Pettit, 2010). This is consistent with culturally focused theories (e.g., Zayas et al., 2005), which indicate that family dysfunction and conflict are central to suicide attempts among Hispanics. Family conflict and the acculturative process can also lead to a decreased sense of social support, which is associated with thwarted belongingness (Van Orden et al., 2010). Indeed, family has been identified as a main source of social support for Hispanics, which, when diminished, contributes to increased suicidal behaviors (e.g., Garcia, Skay, Sieving, Naughton, & Bearinger, 2008).

On the other hand, feeling close to one's ethnic and cultural group (e.g., strong ethnic identity), as well as the beliefs or values held by the group (e.g., religious beliefs; value of family or tradition, etc.), could increase feelings of connection and contribution to one's cultural group (Wadsworth & Kubrin, 2007). A stronger connection to Hispanic identity and culture can thereby protect against thwarted belongingness or perceived burdensomeness, and thus suicidal desire. Despite the prevalence of research on the interpersonal theory of suicide, however, there is a dearth of studies examining the theory among Hispanic populations.

Only one study has examined the association of an interpersonal theory construct with suicide ideation among Spanish speakers (Garza & Pettit, 2010). Specifically, Garza and Pettit (2010) found that perceived burdensomeness, but not familism, was associated with suicide ideation among Mexican women outpatients. Although this study used Spanish measures, only the Burdensomeness subscale from the original English 25-item INQ (as opposed to the more recently validated 15-item INQ; Van Orden, Cukrowicz, Witte, & Joiner, 2012) was translated and used ($\alpha = .79$). Detailed translation and psychometric analysis were not provided.

---

[1]Although considered more distal, painful and provocative experiences associated with migrant living, poverty, or discrimination could also increase the acquired capability for suicide.

The Interpersonal Needs Questionnaire

The 15-item INQ (Van Orden et al., 2012) was developed and validated in English to measure the extent to which individuals believe their need to belong is not met (i.e., thwarted belongingness) and perceive themselves to be a burden on others (i.e., perceived burdensomeness). Items on the original 25-item version of the INQ were rationally derived from hypotheses of the interpersonal theory (Joiner et al., 2009). Participants rate how true each statement is for them recently on a 7-point Likert scale (1 = *not at all true for me* to 7 = *very true for me*). Scores are totaled such that higher numbers reflect higher levels of thwarted belongingness and perceived burdensomeness. Studies using early versions of the INQ found thwarted belongingness and perceived burdensomeness to be associated with severity of concurrent suicide ideation in both young (Van Orden et al., 2008) and older (Cukrowicz, Cheavens, Van Orden, Ragain, & Cook, 2011) adults, thus providing initial evidence for the concurrent validity of the INQ. These constructs have also displayed high sensitivity and specificity for suicide ideation among active duty personnel using an abbreviated INQ (Bryan, 2011).

In response to psychometric concerns with the 25-item INQ (e.g., multicollinearity and administration length), Van Orden et al. (2012) examined the psychometric properties of the original 25-item INQ within multiple samples and developed a psychometrically sound 15-item version of the INQ for use in future research. The first six items measure perceived burdensomeness (e.g., "These days the people in my life would be better off if I were gone"); all items are directly worded. The next nine items measure thwarted belongingness (e.g., "These days, other people care about me" [reversed]); Items 7, 8, 10, 13, 14, and 15 are reverse-worded and reversed scored. Good fit was established for a model with two latent variables (i.e., thwarted belongingness and perceived burdensomeness). These latent variables displayed convergent associations with related interpersonal constructs (e.g., loneliness, social support, social worth, and death ideation) as well as prospective associations with suicide ideation. Multiple-group analyses further supported the relevance of these constructs to diverse populations (i.e., younger and older adults, clinical and nonclinical samples).

The INQ-15 (Van Orden et al., 2012) has some notable limitations. First, moderate to high intercorrelations have been reported for subscale scores. Potentially related, the Thwarted Belongingness subscale displayed low divergent validity (Van Orden et al., 2012). Second, there are a large number of reverse-worded items within the Thwarted Belongingness subscale (with only directly worded items for the Perceived Burdensomeness subscale). Although reverse-worded items are usually included to protect against acquiescent or response set behaviors, doing so tends to negatively affect internal consistency (Barnette, 2000). Accordingly, it has been suggested that reverse-worded items may not actually be interpreted as the opposite of directly worded items (e.g., reverse scoring "I feel like I belong" may not actually capture "I feel like I do not belong"; Segura & González-Romá, 2003). This may problematically impact translations by further complicating the interpretation of psychological constructs cross-culturally.

Hill et al. (2015) examined the factor structure, internal consistency, and concurrent predictive validity of all the different item-versions of the INQ used in previous research (i.e., 10, 12, 15, 18, and 25 items) across three samples (a psychiatric adolescent inpatient sample and two undergraduate samples). All versions had a two-factor solution, reflecting perceived burdensomeness and thwarted belongingness. The empirically derived 15-item version (Van Orden et al., 2012) and a 10-item version primarily used with military samples[2] (e.g., Bryan, Morrow, Anestis, & Joiner, 2010) were the only ones to display acceptable model fit among undergraduates and close to acceptable fit among clinical adolescents. Scores on both subscales of Thwarted Belongingness and Perceived Burdensomeness displayed acceptable to good internal consistency but were only associated with concurrent suicide ideation on the 10-item version.[3]

Importantly, the psychometric methods used to validate the 15-item INQ in English have not previously been applied to a Spanish translation of the INQ with diverse Spanish-speaking samples. The differing degree of fit displayed across English-speaking samples (partially because of item content; Hill et al., 2015) highlights the need to examine the factorial validity and internal consistency of the INQ cross-culturally and in different languages as well.

Current Study

The present study sought to (a) translate and validate the 15-item INQ into Spanish (INQ-S) and evaluate its psychometric properties, including the (b) factor structure, (c) measurement invariance, and (d) reliability, construct validity, and criterion validity of the translated measure. The INQ-S was developed with the goal of promoting research on interpersonal theory constructs and suicide risk,[4] as well as providing a means to measure therapeutic outcomes, among Spanish speakers. We did not examine the original (25-item) version of the INQ, as we aimed to create a Spanish version of the INQ that was comparable in content or structure with the current and most commonly used (across different populations) version of the INQ (i.e., the 15-item version). As the theory posits that thwarted belongingness and perceived burdensomeness are proximal causes of suicidal desire regardless of population, we hypothesized that the two-factor measurement model of the original INQ would maintain adequate fit for the INQ-S.

Aim 1: Multistage Translation of the INQ

Although Hispanics are not a monolithic group, Hispanic subgroups are united via a common language and shared cultural traditions and values. As with other languages, including English, regional and national linguistic differences and dialects exist across Spanish speakers. In line with the goals of the Asociación de Academias de la Lengua Española[5] [Association of Academies of the Spanish Language], we aimed to achieve a

---

[2]The INQ-10 was not psychometrically derived; to balance brevity with psychometric utility, Bryan (2011) selected five perceived burdensomeness and five thwarted belongingness items from an unpublished version of Van Orden et al.'s (2012) factor analysis.
[3]The INQ-15 displayed concurrent predictive validity for suicide ideation for perceived burdensomeness only.
[4]Although the interpersonal theory of suicide is a prominent theory, its universal claims and cross-cultural relevancy to ethnic and linguistic minorities is understudied in part because of a lack of viable measures. The applicability of the theory, and how it fits with known cultural risk factors, is an important empirical question for future research.

standard Spanish INQ translation that could be understood across Spanish-speaking subgroups and adapted as needed. Figure 1 provides a flowchart of the full translation and validation process presented in this article.

Stage 1

Copyright permission was obtained to reproduce the INQ items in Spanish. Following procedures laid out by Chapman and Carter (1979), the INQ was initially translated (instructions and items) into Spanish by the first author of this article (native Spanish speaker of Colombian descent). The INQ-S was developed with a sixth-grade reading level to facilitate comprehension. Three independent native Spanish speakers (of Cuban, Mexican, and Spanish descent) separately reviewed this initial draft for grammar and fluency. A consensus meeting (with the first author and three reviewers) was held to discuss proposed edits and content validity for each item. Changes were made based on consensus agreement between all four individuals. The resulting second draft of the instrument was then translated back into English by a fifth individual (native bilingual English–Spanish speaker of Colombian descent) blind to the English version. A second consensus meeting was then held between the first author and initial three reviewers to discuss apparent item discrepancies between the original INQ (English version) and the back translation. Content validity for each item was extensively discussed, particularly in regard to retaining construct meaning while satisfying cultural relevancy and understanding for Hispanics subgroups.

It is important to note that translation methodologies often involve a procedure called *decentering* (Brislin, Lonner, & Thorndike, 1973), which involves modification of both the original and the second language version of a measure to attain equivalency. This method allows for the modification of words or concepts that are not equivalent between languages. Once all modifications are completed, both versions of the measure would then need to be psychometrically tested. Decentering was not used in this study because (a) the measure already has a wide history of use within the original language (English), and (b) keeping items intact is preferred so as to allow comparisons with previous research using the INQ English version (Chapman & Carter, 1979). For these reasons, the translation group did not drastically revise the content of existing items or create new items. A final draft was achieved following consensus agreement on word choice, grammar, and content. As intended, using Legibilidad software (Muñoz Fernández, 2016), this final draft produced a Fernández-Huerta Spanish readability score of 79.97 (Fernández Huerta, 1959), which corresponds to "somewhat easy" or a sixth-grade reading level.

Stage 2

To assess potential language effects, both versions of the instrument were administered to a pilot sample of bilingual (i.e., English and Spanish) college students. Participants were asked to complete the INQ online on two occasions, separated by a 1-week delay. The 1-week interval is consistent with previous research examining cross-language translations of commonly used measures (e.g., Wiebe & Penley, 2005). The theory also identifies perceived

---

[5]The Asociación de Academias de la Lengua Española (ASALE) is tasked with maintaining the integrity and growth of the Spanish language. As with other linguistic regulatory academies, the ASALE sets common norms (linguistic prescriptivism) aimed at fostering language unity across the Spanish-speaking world.

burdensomeness and thwarted belongingness as dynamic states that are proximal risk factors of suicide ideation; thus, a longer delay may reflect changes in these states as opposed to translational differences.

Participants were randomly assigned to either take the INQ in the same language both times or in a different language on each occasion, with language order counterbalanced. Participants were randomized to control for individual differences in English and Spanish language proficiency. Participants completed a demographics form in English at Time 2 assessing gender, age, race, ethnicity, country of birth, socioeconomic status, religion, education, and language proficiency. Participants received course credit for participating. Procedures were approved by the Florida State University (FSU) Institutional Review Board (IRB).

Participants.—Participants were 56 bilingual undergraduates at a large university in the Southeastern United States recruited from the psychology department.[6] Participants were randomized into one of four conditions: English-Spanish ($n = 13$), Spanish-English ($n = 16$), English-English ($n = 14$), and Spanish-Spanish ($n = 13$). Mean age was 22.57 years ($SD = 5.17$ years; range = 18– 47 years). Most participants were female (82.1%), single or never married (83.9%), Catholic (55.4%), and had a 2-year college degree (53.6%). Race was as follows: 80.4% White/Caucasian, 3.6% Black, and 16.1% Other. Of participants who selected Other, race was specified as Hispanic ($n = 4$), Latino ($n = 2$), or "mixed" ($n = 2$). All participants identified their ethnicity as Hispanic and reported fluent or native competency in English and Spanish. All participants learned both languages before Age 6, most frequently reporting learning English (33.9%) and Spanish (91.1%) since birth (see Supplementary Table 1 of the online supplemental materials for self-reported language fluency).

Results.—On average, participants completed the second survey 7.19 days ($SD = .46$ days; range = 6.49–8.97 days) after the first. Randomization was successful; participants did not differ across conditions on age, $F(3, 52) = 1.10$, $p = .36$, time between surveys, $F(3, 52) = 1.08$, $p = .37$, sex, $\chi^2(3) = 2.83$, $p = .42$, race, $\chi^2(6) = 3.66$, $p = .72$, marital status, $\chi^2(15) = 13.71$, $p = .55$, education, $\chi^2(9) = 6.47$, $p = .69$, religion, $\chi^2(12) = 11.62$, $p = .48$, or language proficiency—reading, $\chi^2(9) = 15.09$, $p = .09$; writing, $\chi^2(12) = 13.41$, $p = .34$; speaking, $\chi^2(12) = 14.09$, $p = .30$; and listening, $\chi^2(6) = 6.84$, $p = .34$.

Correlations.: Corresponding correlations for items completed in Spanish at both time points were all large and significant (see Supplementary Table 3 of the online supplemental materials). Items completed in English at both time points also displayed large and significant correlations, except for Bel3 ("I rarely interact with people who care about me"), Bel4 ("I am fortunate to have many caring and supportive friends"), and Bel8 ("I am close other people"), which were not significantly associated across time.

---

[6]Six other participants completed Part 1 (three in English and three in Spanish) but not Part 2, and were thus excluded from all analyses (attrition = 9.78%). Demographics were not available for these participants, as demographic data were collected at the end of Part 2 to avoid any priming effects.

Corresponding correlations for items completed in different languages at Time 1 and 2 were large and significant, except for Bel1 ("Other people care about me") and Bel3 ("I rarely interact with people who care about me") in both the English-Spanish and Spanish-English conditions. Correlations for all burdensomeness items for the English-Spanish condition and Bur4 ("I think my death would be a relief to the people in my life") for the Spanish-English could not be calculated because responses were constant at Time 1 (i.e., all respondents selected "1" for each item). Even so, responses stayed the same at Time 2 for almost all participants (93%) on each burdensomeness item. For the English-Spanish condition only, Bel5 ("I feel disconnected from others") was also not significant across time points. For the Spanish-English condition only, Bur5 ("I think the people in my life wish they could be rid of me"), Bel2 ("I feel like I belong"), Bel8 ("I am close to other people"), and Bel9 ("I have at least one satisfying interaction every day") were also not significant across time points. When the counterbalanced conditions (English-Spanish and Spanish-English) were combined to form a cross-language condition (Wiebe & Penley, 2005), only Bur5, Bel1, Bel3, Bel8, and Bel9 were not significantly associated across time. The correlations between corresponding burdensomeness ($r = .97$–1.00) and belongingness ($r = .59$–$.92$) total scores were all large and significant regardless of whether the same or different language was given across time points.

Reliability.: Scale reliabilities were good to excellent and similar in both languages for Perceived Burdensomeness and Thwarted Belongingness at baseline (English Perceived Burdensomeness, $\alpha = .95$, and Thwarted Belongingness, $\alpha = .88$; Spanish Perceived Burdensomeness, $\alpha = .96$, and Thwarted Belongingness, $\alpha = .78$). English average interitem correlations were .87 for Perceived Burdensomeness and .46 for Thwarted Belongingness. Spanish average interitem correlations were .85 for Perceived Burdensomeness and .35 for Thwarted Belongingness.

Following test–retest reliability guidelines by Koo and Li (2016), 1-week test–retest reliability was calculated using the intraclass correlation coefficient (ICC[3,1]; absolute agreement). In English, the ICCs for Perceived Burdensomeness (ICC = .99, 95% confidence interval (CI) [.96, 1.00], $p < .001$) and Thwarted Belongingness (ICC = .91, 95% CI [.75, .97], $p < .001$) were excellent. In Spanish, the ICC was excellent for Perceived Burdensomeness (ICC = .98, 95% CI [.92, .99], $p < .001$) and good for Thwarted Belongingness (ICC = .89, 95% CI [.67, .96], $p < .001$). Among those who took both the English and Spanish versions, the cross-language ICCs were moderate: .52 (95% CI [.21, .74], $p < .01$) for Perceived Burdensomeness and .61 (95% CI [.32, .79], $p < .001$) for Thwarted Belongingness.

Language effect.: Supplementary Table 2 of the online supplemental materials provides descriptive and test statistics. To test the effect of language, a 2×2 repeated measures analysis of variance was used, with language of administration as a within-subjects variable and order of administration as a between-subjects variable (Spanish-English vs. English-Spanish). This strategy allows for the analysis of a potential language effect separate from any order or time effect. For perceived burdensomeness, there was no significant effect for language ($p = .31$), order of administration ($p = .50$), or their interaction ($p = .18$). For

thwarted belongingness, there was no significant effect for language ($p = .07$), although Spanish scores were slightly higher than English scores. There was no significant effect for order of administration ($p = .83$) or its interaction with language ($p = .39$).

Discussion.—The Spanish translation of the INQ did not display language effects in a pilot sample of bilingual college students, suggesting little difference in overall scores within participants that can be attributed to language. All INQ items were significantly associated across time when administered in Spanish both times. A few items were not associated across time when measures were completed in a different language from the first to second time points, potentially indicating translation differences (e.g., the word *care* in English could not be translated directly). Even when the measure was given in English both times, some of the items were not significantly associated across time. This could indicate that some items are more sensitive to change over time, especially for thwarted belongingness. Also, a large number of thwarted belongingness items are reverse-worded, which may negatively impact item reliability (Barnette, 2000).

Although this pilot administration was limited by a smaller sample size, it served as a first step toward validating the Spanish translation of the INQ among bilingual English and Spanish speakers. Overall, given evidence that item sensitivity to change may fluctuate over time regardless of language, no further revisions were made to the translation and the final draft was used in the proceeding validation analyses.

Aim 2: Confirmatory Factor Analysis

Confirmatory factor analyses (CFAs) examining the previously established two-factor measurement model for the INQ-15 in English (Van Orden et al., 2012) were conducted in three Spanish-speaking samples: American college students (Sample 1), college students from Spain (Sample 2), and clinical inpatients from Mexico (Sample 3). Studies were approved by the FSU IRB and participating sites.

Method

Sample 1.—Sample 1 consisted of 281 heritage Spanish-speaking adults from FSU ($n = 84$) and Florida International University (FIU; $n = 197$). Participants were recruited via the psychology departments' psychology research participation system as well as flyers. Participants participated in person or online and, following informed consent in language of preference, completed a demographics form and the INQ in Spanish. Participants were debriefed and compensated (course credit or entered into a raffle). All participants reported fluent or native competency in Spanish and learning Spanish before Age 6. Participants from FSU and FIU did not differ significantly from each other on factors of sex, age, race, ethnicity, marital status, religion, or Spanish proficiency. Given no differences in demographic factors, the two groups were combined to improve statistical power.

The majority of the sample was female (77.6%), single or never married (85.9%), had at least some college education (86.8%), and identified as Catholic (52%). Mean age was 22.20 years ($SD = 6.73$ years; range = 18–59 years). Self-identified race was as follows: 71.2% White/Caucasian, 1.4% Black, 0.4% Native American/Alaskan Native, and 30% Other. Most

participants who identified as Other specified their race as Hispanic, Latino, Mestizo, or multiracial/mixed, with the rest claiming their nationality as their race ($n = 3$) or not specifying ($n = 6$). Most participants identified as Hispanic (97.9%; 2.1% non-Hispanic). Most participants (87.2%) indicated learning Spanish since birth, and most frequently reported native competency in reading (50.5%), writing (39.1%), speaking (57.7%), and listening to (75.4%) Spanish (see Supplementary Table 1 of the online supplemental materials for self-reported language fluency).

Sample 2.—Sample 2 consisted of 1,016 adults who were firstor second-year psychology majors at the Universidad Nacional de Educación a Distancia (UNED) in Spain. Although college students, the majority of students were employed, as UNED is a distance-learning institution. Thus, participants are more representative of the general population than traditional college students. The majority of the sample identified as White (85.0%) or Black (15.0%), female (75.1%), and single (65.7%). Mean age was 33.02 years ($SD = 10.12$ years; range = 19–64 years). Participants completed a demographics form and the INQ-S online, among other measures. All participants were native Spanish speakers.

Sample 3.—Sample 3 consisted of 181 individuals receiving treatment for substance use at rehabilitation treatment facilities in Mexico. The participants did not differ significantly from each other in regard to their rates of depression and suicidal behaviors. Participants were invited to participate in the study, signed an informed consent form, and completed the INQ-S as part of a larger research project already being conducted at substance use treatment facilities in Mexico. Most of the sample was male (91.2%) and identified as Catholic (71.3%). Mean age was 27.06 years ($SD = 10.24$ years; range = 14–58 years). All participants were Mexican nationals living in Mexico and native Spanish speakers.

Data analytic plan.—CFA was conducted with each sample data using the robust maximum likelihood (MLR) estimation method. Other less restrictive estimation methods (e.g., exploratory structural equation modeling) were not examined, as the INQ has a strong theoretical foundation and a well-established measurement model (Hill et al., 2015; Van Orden et al., 2012). All the analyses were conducted in Mplus (Version 7.4; Muthén & Muthén, 1998 –2015).

A two-factor measurement model, consistent with the interpersonal theory of suicide and the English INQ (Van Orden et al., 2012), was examined. Even though previous psychometric analyses of the INQ have rejected an alternative one-factor model (Van Orden et al., 2012), we also tested a one-factor model, as perceived burdensomeness and thwarted belongingness items could conceptually represent a broader Social Disconnection factor. Three and four-factor models, however, were not tested, as previous research has indicated that a three-factor solution is uninterpretable and a four-factor solution has too few indicators per factor to ensure stability and viability of the factor (see Van Orden et al., 2012).

The two-factor model included 15 observed variables and two latent variables (i.e., perceived burdensomeness and thwarted belongingness). Each observed variable was an indicator for only a single latent variable (i.e., six for perceived burdensomeness and nine for

thwarted belongingness). The first factor loading was fixed to 1 to manage scale dependency for each latent variable. The covariance between the two latent variables was estimated.[7]

It is important to note that although responses to the INQ items are continuous, responses to some items may not be normally distributed (see Van Orden et al., 2012). In the current study, based on skew and kurtosis values for Samples 1 to 3 (see Supplementary Table 4 of the online supplemental materials), most of the INQ items were also not normally distributed in at least one of the samples. Accordingly, data did not meet multivariate normality based on Mardia's multivariate test of skewness (g1) and kurtosis (g2) in Samples 1 to 3 (g1, $p < .05$; g2, $p < .05$). Following procedures by Van Orden et al. (2012), transformations were not used as these items measure constructs that are not expected to be normally distributed in the population. Instead, we used MLR because it is an estimation procedure robust to violations of non-normality (Brown, 2015).

Model fit for all samples was examined using several fit indices, as each index of model fit has unique properties characterized by different strengths and weaknesses. Robust indices to be evaluated included: two indices of absolute fit, chi-square (χ^2) and standardized root mean square residual (SRMR); two indices of comparative fit, the comparative fit index (CFI) and the Tucker-Lewis index (TLI); and a parsimony corrected fit index, the root mean square error of approximation (RMSEA). The following recommendations were used when cutoff criteria was available: CFI and TLI values greater than .95, RMSEA values close to or below .06, and SRMR values close to or below .08 suggest good fit (Hu & Bentler, 1999). We report the Yuan-Bentler scaled chi-square (YB χ^2), which is a chi-square test of overall model fit for continuous nonnormal outcomes (nonsignificant values indicate good fit).

For construct validity of the proposed measurement theory, we examined convergent validity by calculating the average variance extracted (AVE) for each latent construct (i.e., degree to which indicators of each construct shared a high proportion of variance). An AVE   .5 suggests adequate convergence (Hair, Black, Babin, & Anderson, 2010). The AVE was then compared with the squared correlation estimate of the constructs to determine discriminant validity. An AVE higher than the squared correlation estimate indicates good discriminant validity (i.e., latent construct explains more of the variance in its indicators than it shares with another construct). We report McDonald's (1999) omega coefficient (m), a measure of reliability based on factor loadings. McDonald's omega is less likely to over or underestimate reliability than Cronbach's alpha even under violations of tau-equivalence (Dunn, Baguley, & Brunsden, 2014).

Results

Description of data.—Descriptive statistics for all three samples (i.e., mean, standard deviation, skew, kurtosis) are provided for each INQ item in Supplementary Table 4 of the online supplemental materials. There were no missing data. To examine the viability of the

---

[7]Although the final measurement model specified by Van Orden et al. (2012) included correlated residual terms (i.e., between two belongingness items), we did not correlate these error terms, as it is not recommended when establishing a unidimensional construct (Anderson & Gerbing, 1988; Landis, Edwards, & Cortina, 2009).

CFA model in each of the samples, fit statistics and standardized factor loadings appear in Tables 1 and 2, respectively.

Model fit.—Indices of overall model fit indicated poor fit on all fit indices for the 15-item two-factor model in all three samples. Some items (Bel, Bel3, Bel6 in Sample 1; Bel1–3 in Sample 2; Bur1, Bel3, Bel5–6 in Sample 3) did not appear to be precise indicators of their construct as their factor loadings or R^2 values were small.[8] Importantly, most of these items were directly worded. It is possible that some of these items performed worse because of the unequal number of reverse- and directly worded items within the Thwarted Belongingness subscale.

Notably, Bel1 ("Other people care about me") and Bel3 ("I rarely interact with people who care about me") are both items that do not have literal translations into Spanish, with the closest word to *care* being represented by *worry* (*preocupar*). For the Mexican clinical sample (Sample 3), the low R^2 value for Bel3 is important to note in light of its negative (although statistically significant) factor loading, which is inconsistent with the theory, indicating that this item may not be measuring the construct as intended. Again, this may suggest that translation differences (i.e., *care* vs. *worry* for Bel3; *outsider* vs. *stranger* for Bel6) may make these items imprecise indicators of thwarted belongingness. Furthermore, because of translation differences (*care* vs. *worry*), clinical inpatients that otherwise endorse feelings of thwarted belongingness may also endorse that others are "worried" about them.

The alternative one-factor model was examined in all three samples to determine whether it would provide better fit than the two-factor model. Robust values for χ^2(all samples, $p < .001$), CFI (.47–.56), TLI (.38–.49), RMSEA (.13–.14), and SRMR (.12–.16) values suggested poor fit across all three samples (see Supplementary Table 5 of the online supplemental materials).

As not all translated items may conceptually capture the original constructs (e.g., *care* vs. *worry*), we considered items that had low R^2 values (i.e., <.20) for elimination (Hooper, Coughlan, & Mullen, 2008), including Bur1 ("People in my life would be better off if I were gone"), Bel1 ("Other people care about me"), Bel2 ("I feel like I belong"), Bel3 ("I rarely interact with people who care about me"), Bel5 ("I feel disconnected from other people"), and Bel6 ("I often feel like an outside in social gatherings"). Indeed, removing these six items (resulting in a nine-item INQ-S) significantly improved model fit in all three samples, as indicated by YB χ^2 difference tests (Sample 1, $\chi^2[24] = 210.58$, $p < .001$; Sample 2, $\chi^2[24] = 660.29$, $p < .001$; Sample 3, $\chi^2[24] = 204.98$, $p < .001$). Results indicated good fit in all samples. While the χ^2 test of model fit in Sample 2 (Spain) was inconsistent with adequate fit, this index may underestimate model fit given large sample sizes (Brown, 2015).

Factor determinacies ranged from .91 to .95 for Perceived Burdensomeness and .86 to .93 for Thwarted Belongingness, indicating evidence for the reliability of the current factor solution. Standardized factor loadings for the latent variable were significant and ranged

---

[8]Standardized factor loadings .50 were considered acceptable and .70 were considered strong (Hair et al., 2010). R^2 values less than .20 were considered low (Hooper et al., 2008).

from .62 to .89 for Perceived Burdensomeness and .62 to .84 for Thwarted Belongingness. R^2 values ranged from .38 to .79, with most values indicating large magnitudes (i.e., >.25). Covariance between the latent variables was significant in Samples 1 and 2 (see Table 2). The alternative one-factor model also displayed poor fit on robust indices across all three samples with the nine-item solution (all samples: χ^2, $p < .001$, CFI = .52–.65, TLI = .36 –.53, RMSEA = .14 –.16, SRMR = .13–.17; see Supplementary Table 5 of the online supplemental materials).

Construct validity.—The AVE for Thwarted Belongingness was .55 in Sample 1, .61 in Sample 2, and .42 in Sample 3. The AVE for Perceived Burdensomeness was .58 in Samples 1 and 2, and .49 in Sample 3. These values are consistent with convergent validity in Samples 1 and 2 but not Sample 3. AVE for Thwarted Belongingness and Perceived Burdensomeness was higher than the squared correlation estimate between the two constructs in each sample, indicating good discriminant validity (see Supplementary Table 6 of the online supplemental materials). Perceived Burdensomeness displayed good reliability in Sample 1 (ω= .85, 95% CI [.77, 92]), Sample 2 (ω = .88, 95% CI [.85, .91]), and Sample 3 (ω = .82, 95% CI [.78, .87]). Thwarted Belongingness displayed good reliability in Sample 1 (ω = .83, 95% CI [.79, .87]) and Sample 2 (ω = .86, 95% CI [.84, .87]), and acceptable reliability in Sample 3 (ω = .74, 95% CI [.68, .80]).

Discussion

It is important to interpret results in light of study limitations. First, although all participants in Samples 2 and 3 were native Spanish speakers, language proficiency was not assessed. Although language proficiency was assessed in Sample 1, it was not included in the model; future research could use multiple indicators multiple causes (MIMIC) models to enter language proficiency as a possible covariate. Second, samples varied greatly in size; larger sample sizes can improve precision and statistical power of the model's parameter estimates as well as indices of overall model fit (Brown, 2015). Finally, functional equivalence of meaning was difficult to achieve for some items of the INQ (especially for thwarted belongingness), suggesting that future studies should examine whether removing or adding novel items to the INQ-Spanish could further improve cross-cultural assessment of the constructs. In such cases, use of less restrictive estimation methods, such as exploratory structural equation modeling, which allow for examination of possible cross-loadings of items, may be recommended.

Overall, both of the latent variables seem to be best measured by the indicators included in the nine-item INQ-S (INQ-S-9), and most of these indicators were reasonably good indicators of the latent constructs of perceived burdensomeness and thwarted belongingness. Thus, the results of the model estimation suggest that a nine-item model, with two distinct but related constructs, represents a viable measurement model for the INQ-S.

Of note, all items retained for the Thwarted Belongingness subscale in the nine-item model are reverse-worded and should be reverse scored. As such, to derive observed total scores for Perceived Burdensomeness, Items 1 to 5 should be summed. However, to derive observed total scores for Thwarted Belongingness, Items 6 to 9 should be reverse coded and summed

(so that higher scores represent greater *thwarted* belongingness). Importantly, as mentioned previously, reverse-worded items can negatively affect item reliability and may not actually conceptually reflect the opposite of its content (Barnette, 2000; Segura & González-Romá, 2003). Future studies should evaluate the psychometric impact of rewording these items to retain a range of risk items (e.g., "I am *not* close to other people") for the 15-item INQ in Spanish, as it may improve item performance on this subscale. The Appendix lists all original 15 items in Spanish as well as which items on the original Thwarted Belongingness subscale require recoding.[9]

Aim 3: Measurement Invariance

Measurement invariance for the INQ-S-9 was examined via three multiple-group CFAs using the MLR estimator in Mplus (Version 7.4; Muthén & Muthén, 1998 –2015).These analyses examine whether students from the U.S. and Spain, and inpatients from Mexico, respond to items on the INQ-S in a manner that produces equivalent factor structures (equal form), factor loadings, and indicator intercepts to responses from an *independent* sample of Mexican college students (Sample 4). Sample 4 was selected as the comparison group, as it consisted of native Spanish speakers that resembled the population in which the original INQ was developed (i.e., college students). Equal forms (i.e., configural invariance) indicate that the latent constructs have similar patterns of loadings across groups. Equal factor loadings (i.e., metric invariance) indicate that each item contributes to its respective latent construct similarly across groups. Equal indicator intercepts (i.e., scalar invariance) indicate that latent (mean) scores are related to observed scores regardless of group, allowing for the comparison of (latent) means across groups (Brown, 2015).

Method

Sample 4.—Sample 4 consisted of 239 college students from a private university in Central Mexico. Students were recruited from the psychology department and completed the INQ-S as part of a larger research project. Most of the sample was female (73.2%), with a mean age of 21.64 years ($SD = 2.42$ years; range = 18–29 years). All participants were Mexican nationals living in Mexico and native Spanish speakers. The nine-item two-factor solution for the INQ-S provided good fit per robust indices, $\chi^2(26) = 29.55$, $p = .29$, CFI = .99, TLI = .99, RMSEA = .02, 90% CI [.00, .06], SRMR = .03, and displayed good reliability (Perceived Burdensomeness, $\omega = .88$, 95% CI [.85, .92]; Thwarted Belongingness, $\omega = .85$, 95% CI [.81, .88]).

Data analytic plan.—The first multiple-group CFA compared American (Sample 1) and Mexican (Sample 4) college students. The second multiple-group CFA compared native Spanish-speaking college students from Spain (a random subsample of Sample 2) and Mexico (Sample 4). The third multiple-group CFA compared clinical (Sample 3) and nonclinical (Sample 4) native Spanish speakers from the same country (Mexico).

---

[9]Using the original item numbering of the INQ-15, the following items were eliminated: 1 (Bur1), 7 (Bel1), 8 (Bel2), 9 (Bel3), 11 (Bel5), and 12 (Bel6).

Beginning with tests of equal forms, models were tested sequentially with increased restrictiveness—equivalence of factor loadings followed by equivalence of intercepts—until invariance no longer held. Equivalence of forms, factor loadings, and intercepts was tested by constraining the factor structure, all factor loadings, and indicator intercepts, respectively, to be the same across groups. Invariance of form is supported by acceptable goodness-of-fit indices. Invariance of factor loadings is supported by a nonsignificant YB χ^2 difference test of the equal forms solution with the nested more restricted solution (equal factor loadings). Similarly, invariance of intercepts is supported by a nonsignificant YB χ^2 difference test of the equal factor loadings solution with the nested more restricted solution (equal intercepts). Importantly, as chi-square difference testing can be differentially influenced by large and unequal sample sizes, balanced group sizes are recommended so that one group does not contribute considerably more to the χ^2 (Brown, 2015). For this reason, we used a random subsample ($n = 250$) of Sample 2 in the multiple-group CFA with students from Spain instead of the full sample size ($N = 1,016$). Differences between CFI values of the nested and comparison model were also examined, such that a decrease of .01 or less supports invariance (Cheung & Rensvold, 2002).

Invariance of intercepts is necessary for valid comparisons of group mean differences on the latent constructs, as, otherwise, differences may be related to construct measurement differences (Brown, 2015). Latent means analysis involves conducting multiple-group CFAs when factor variances are constrained and the reference group's mean is fixed to zero (the mean in the other group is freely estimated and represents the deviation from the reference group's mean).

Results

Results appear in Table 3. Models that imposed the same factor structure (i.e., test of equal form) for both groups (across all three comparisons) were consistent with good fit, indicating equivalent factor structure for the INQ-S-9 across populations of college students from America and Spain and clinical inpatients from Mexico (when compared with Mexican college students). Tests of equivalent factor loadings using the YB χ^2 difference test indicated invariance across all three comparisons. The CFI difference supported equality of factor loadings for comparisons with American college students and Mexican inpatients as well but was equivocal for college students from Spain (yet supported partial invariance when one indicator was freed). Using the YB χ^2 difference test, invariance also held for tests of equivalence of intercepts for comparisons of college students from America and Spain with those from Mexico; this was consistent with the CFI difference. However, invariance of intercepts did not hold for clinical inpatients when compared with Mexican college students based on YB χ^2 or CFI difference values. Modification indices suggested that the intercepts of Bur3 ("I think I am a burden on society") and Bel8 ("I am close to other people") were not invariant across Mexican inpatients and students, and should be freed. Partial invariance was established after allowing these intercepts to vary, as the YB χ^2 difference test was nonsignificant and the CFI difference was —.001.

As invariance for intercepts was established for comparisons with college students, we proceeded with latent means analysis. On average, college students from America scored .33

units lower on Perceived Burdensomeness ($p < .001$) and .41 units lower on Thwarted Belongingness ($p < .001$) than those from Mexico. College students from Spain scored .31 units lower than those from Mexico on Perceived Burdensomeness ($p < .001$) but did not differ on Thwarted Belongingness. Clinical inpatients from Mexico scored .80 units higher on Perceived Burdensomeness ($p < .001$) and .41 units higher on Thwarted Belongingness ($p < .01$) than students from Mexico.

Discussion

These multiple-group analyses help clarify the generalizability of the psychometric properties of the INQ-S-9 across populations varying in nationality and clinical severity, parameters that are known to influence suicide rates. As suggested by results from the previous single-group CFAs, the equivalent latent structures of the INQ-S-9 hold across college students and clinical inpatients. This suggests that the two-factor measurement model of the INQ-S-9 is viable for Spanish-speaking populations of different nationalities and clinical severity. Invariance of factor loadings indicates that items on INQ-S-9 are also appropriate indicators of these constructs across levels of nationality and clinical severity. Furthermore, among varying nationalities of college students, equivalence of indicator intercepts suggested that the levels of the underlying items are equal across those groups. However, lack of equivalence of intercepts for clinical inpatients (when compared with college students of the same nationality) suggests that this may not be the case for samples with greater clinical severity. Item weighting may improve item precision in measuring perceived burdensomeness and thwarted belongingness in clinical samples (Brown, 2015). Of note, intercepts invariance was not established for the English INQ-15 in comparisons of outpatients and older adults with undergraduates (Van Orden et al., 2012).

Finally, latent means analysis indicated that, on average, Mexican clinical inpatients had higher Perceived Burdensomeness and Thwarted Belongingness scores than Mexican college students, whereas Mexican college students had higher Perceived Burdensomeness scores than students from America or Spain as well as higher Thwarted Belongingness scores than American students.

Aim 4: Reliability, Construct Validity, and Criterion Validity

In this section, we present findings on test–retest reliability, construct validity (i.e., convergent and divergent validity), and criterion validity for the INQ-S-9 in an independent sample. According to Van Orden et al. (2012), thwarted belongingness, loneliness, social support, self-liking (self-esteem), and relatedness are related but distinct constructs. Similarly, perceived burdensomeness, competence, autonomy, and responsibility to family should be related but distinct constructs. All of these constructs may also be associated with each other, as they reflect the construct of social connectedness (Berkman, Glass, Brissette, & Seeman, 2000). Thwarted belongingness and perceived burdensomeness are expected to be associated with suicide ideation, as they are proposed proximal causes of suicidal desire (Joiner, 2005; Van Orden et al., 2010).

Method

Sample 5.—Participants were 104 Spanish-speaking U.S. adults recruited online via Mechanical Turk (MTurk). Mean age was 34.63 years (SD = 10.91 years; range = 18 – 69 years). Most participants were male (61%), single or never married (51.9%), Catholic (47.1%), and had a 4-year college degree (33.7%). Approximately half (51.9%) were born in the United States. Race was 51.9% White/Caucasian, 26.9% Mestizo, 5.8% Indigenous, 10.6% Other, 2.9% Mulatto, 1.0% Black, and 1.0% Asian. All participants who identified their race as "Other" specified (in a write-in field) Hispanic/Latino or their nationality as their race instead. A majority of participants (87.5%) identified as Hispanic/Latino; those who did not reported their nationality as their ethnicity. All participants reported learning Spanish before Age 6 and met MTurk qualifications for fluency in Spanish.

Participants completed a demographics form, the INQ-S-9 (Perceived Burdensomeness, ω = .95, 95% CI [.91, .98]; Thwarted Belongingness, ω = .88, 95% CI [.84, .92]), as well as measures assessing competence, autonomy, responsibility to family, loneliness, social support, self-liking, relatedness, and suicide ideation online in Spanish.[10] Participants were invited to complete the INQ-S-9 again 1 week later (M = 7.85 days, SD = 1.00 days), with 41.3% (n = 43) completing the INQ-S-9 at both time points. The FSU IRB approved all study procedures.

Measures.

Basic Psychological Need Satisfaction and Frustration Scale (BPNSFS; Chen et al., 2015).—The BPNSFS was designed to measure the satisfaction and frustration of three basic needs: autonomy, competence, and relatedness. The scale was validated in four culturally diverse samples, including Peru (Chen et al., 2015). The scale consists of 24 items, rated from 1 (*completely false*) to 5 (*completely true*). Items correspond to one of six subscales: Competence Satisfaction (α = .83), Competence Frustration (α = .79), Autonomy Satisfaction (α = .85), Autonomy Frustration (α = .84), Relatedness Satisfaction (α = .90), and Relatedness Frustration (α = .76). Higher scores represent either greater needs satisfaction or frustration. We used the satisfaction subscales in the subsequent analyses to be consistent with Van Orden et al. (2012).

Beck Scale for Suicide Ideation (BSS; Beck & Steer, 1991).—The BSS contains 19 groups of statements, rated from 0 to 2, that assess suicide ideation. The BSS has adequate internal consistency; high face, convergent, and construct validity; and adequate test–retest reliability (Beck, Brown, & Steer, 1997). The official Pearson Spanish-language version of the BSS was used (Pearson Education, 2018). Sample internal consistency was good (α = .87).

Responsibility to Family (Linehan, Goodstein, Nielsen, & Chiles, 1983).—Responsibility to Family is a seven-item subscale of the Reasons for Living Inventory (RLI; Linehan et al., 1983). The RLI measures the degree of importance of reasons not to die by

---

[10]McDonald's omega (ω) is reported as a measure of reliability only for measures included in the SEM as latent variables (i.e., the INQ-S-9), as it assumes a congeneric model. Coefficient alpha (α) is reported for all other scales, which were used as direct observed variables in the SEM.

suicide, of which responsibility to family is one. The Spanish version of the RLI Responsibility to Family subscale (α = .89) was used (Garza & Cramer, 2011). The RLI has displayed strong construct validity, internal consistency, and convergent/divergent validity in English and Spanish (Garza & Cramer, 2011; Osman et al., 1999). Items are scored from 1 (*not at all important*) to 6 (*very important*). Higher scores represent greater responsibility to family.

Revised UCLA Loneliness Scale (Russell, Peplau, & Cutrona, 1980).—The UCLA Loneliness Scale is a 20-item measure designed to assess perceptions of loneliness. Items are rated from 1 (*never*) to 4 (*often*), with higher scores indicating greater loneliness. The Spanish version was used (Morejón & García-Bóveda, 1994). The scale has displayed good internal consistency and construct validity in English and Spanish (Morejón & García-Bóveda, 1994; Russell et al., 1980). Internal consistency was excellent (α = .96) in the current sample.

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965).—The RSES is a 10-item measure of global self-esteem. Scores range from 0 (*strongly disagree*) to 3 (*strongly agree*), with higher total scores representing greater self-esteem. The Spanish version (Martín-Albo, Núñez, Navarro, & Grijalvo, 2007) was used as a measure of self-liking in analyses (α = .91). The RSES has displayed good internal consistency, temporal stability, and convergent validity in English and Spanish (Martín-Albo et al., 2007; Robins, Hendin, & Trzesniewski, 2001).

Multidimensional Scale of Perceived Social Support (MSPSS; Zimet, Dahlem, Zimet, & Farley, 1988).—The MSPSS is a 12-item scale designed to measure perceptions of support from family, friends, and significant others. Items are score on a scale from 1 (*very strongly disagree*) to 7 (*very strongly agree*), which produce a total score or three subscales representing family, friends, and significant others. The Spanish version of the MSPSS was used (Mantuliz & Castillo, 2002). The MSPSS has displayed good internal and test–retest reliability, as well as construct validity in English and Spanish (Mantuliz & Castillo, 2002; Zimet et al., 1988). For the purpose of our analyses, we used only the total scale score (α = .97).

Data analytic plan.: For test–retest reliability, we calculated the stability of subscale total scores on the INQ-S-9 across time (i.e., 1 week) using the ICC (ICC [3,1]; absolute agreement; Koo & Li, 2016). For convergent and divergent validity, we used a structural equation model (SEM) to examine the associations between the latent constructs and related constructs (similar to those examined in Van Orden et al., 2012). Following procedures by Van Orden et al. (2012), we regressed seven observed variables (total scale scores for each related construct) onto the INQ-S nine-item measurement model using MLR. We hypothesized that the first four observed variables (loneliness, social support, self-liking, relatedness) were associated with thwarted belongingness, whereas the last three observed variables (competence, autonomy, responsibility to family) were associated with perceived burdensomeness. All observed variables in the structural model were allowed to covary.

Following procedures by Van Orden et al. (2012) for criterion validity, we tested two models to examine the association of the INQ-S-9 with concurrent suicide ideation (observed variable). First, we conducted an SEM regressing suicide ideation (present vs. absent) onto the latent variables of thwarted belongingness and perceived burdensomeness. Second, as dichotomizing a continuous measure can result in (a) reduced statistical power to detect an effect, (b) underestimation of variation in outcome, and (c) concealment of nonlinear relationships (Altman & Royston, 2006), we constructed a negative binomial regression model with suicide ideation severity regressed on perceived burdensomeness and thwarted belongingness (observed sum scores). A negative binomial regression model was selected because of overdispersion (variance greater than the mean) on the BSS; similar results were obtained with using Poisson regression.

Univariate outliers (standard scores ±3.0) were identified for observed sum scores of perceived burdensomeness, competence satisfaction, autonomy satisfaction, responsibility to family, social support, self-esteem, and suicide ideation. Outliers were addressed by bringing the score in question to the next highest value within three standard deviations. The interaction of thwarted belongingness and perceived burdensomeness was not examined because of the low prevalence of suicide ideation. Supplementary Table 7 of the online supplemental materials provides the intercorrelations.

Results and Discussion

Test–retest reliability.—The ICC was excellent for Perceived Burdensomeness (ICC = . 97, 95% CI [.95, .99], $p < .001$) and acceptable for Thwarted Belongingness (ICC = .75, 95% CI [.58, .85], $p < .001$).

Convergent and divergent validity.—Supportive of convergent validity, the associations between thwarted belongingness and loneliness, social support, self-liking, and relatedness were significant and in the expected directions. Perceived burdensomeness was significantly associated in the expected directions with competence and autonomy but not responsibility to family. Figure 2 provides standardized regression coefficients for the simultaneous regression.

However, similar to findings reported by Van Orden et al. (2012), divergent validity was not strongly supported. Thwarted belongingness was significantly associated with competence and autonomy, although these associations were less strong than those with conceptually linked measures. Burdensomeness was also significantly associated with loneliness, social support, self-liking, and relatedness, although thwarted belongingness displayed much stronger associations with those measures. Van Orden et al. suggested that low evidence for divergent validity might not indicate poor scale construction or psychometric properties; rather, it may reflect the broad relation of social disconnection across a variety of psychological experiences. Future studies may better examine divergent validity using measures completely unrelated with social disconnection or more closely tied with burdensomeness (e.g., meaning in life).

Of note, responsibility to family was not associated with thwarted belongingness or perceived burdensomeness. Responsibility to family has been suggested to be relevant to

familism and found to be associated with an attitudinal measure of familism (Garza & Pettit, 2010). Thus, it may be that familism is independent of feelings of belonging or burden among Hispanics. Garza and Pettit (2010) found no associations between familism and responsibility to family (also using the Responsibility to Family subscale) with perceived burdensomeness or suicide ideation among Mexican women.

Criterion validity.—Greater odds of reporting suicide ideation were associated with higher levels of both the Thwarted Belongingness (odds ratio [OR] = 2.52, $p < .001$) and Perceived Burdensomeness ($OR = 2.92$, $p < .001$) latent variables. Examining severity of suicide ideation, higher BSS scores were concurrently associated with greater scores on the Perceived Burdensomeness subscale (95% CI of incidence rate ratio [IRR] [1.08, 1.30], $p < .001$) and Thwarted Belongingness subscale (95% CI of IRR [1.05, 1.26], $p < .01$).

Limitations.—It is important to keep in mind limitations associated with MTurk samples. First, MTurk participants tend to have greater survey experience and may be strongly financially motivated (Chandler & Shapiro, 2016). Second, MTurk participants report a greater degree of clinical symptoms than traditional nonclinical samples (Arditte, Çek, Shaw, & Timpano, 2016). Although it is unclear whether these differential characteristics hold for Spanish-speaking MTurk samples, results may not be generalizable to or representative of the general population.

General Discussion

The first aim of the current study was to translate the 15-item INQ into Spanish and examine potential language effects in a pilot sample of bilingual college students. Results indicated no language effects, strong item reliability, and 1-week test–retest reliability for the translation. Certain items were not significantly associated across time with the same item (mostly for thwarted belongingness) when administered in English at both time points or cross-language (i.e., English vs. Spanish). This is not unexpected, as perceived burdensomeness and thwarted belongingness are posited to be fluctuating and dynamic states. Language and order effects were not observed. Overall, findings suggested that a functional translation was achieved.

Confirmatory Factor Analysis

The second aim of this study was to examine the measurement model for the INQ-S using CFA. Results indicated that the 15-item two factor measurement model developed using English-speaking undergraduates (see Van Orden et al., 2012) did not provide good fit for Spanish-speaking American and Spanish college students or Mexican inpatients using the INQ-S. Rather, a nine-item two-factor measurement model (i.e., after eliminating poor fitting items) provided good fit across all three samples. Poor-fitting items contained, at face value, different concepts when translated (e.g., translating *care* in English as *worry* in Spanish) or were directly worded. One-factor models also consistently displayed poor fit across all three samples for both the 15-item and nine-item solutions, lending support for the appropriateness of the two-factor measurement model for the INQ-S-9. Consistent with the

interpersonal theory of suicide, this suggests that perceived burdensomeness and thwarted belongingness are related but distinct.

Measurement Invariance

The third aim of this study was to examine measurement invariance for the INQ-S-9. Results indicated equivalence in factor structure and factor loadings across the three samples (when compared with an independent Mexican college student sample). This supports the generalizability of the INQ-S measurement model and suggests that the items display comparable associations with the latent constructs across nationality and clinical severity in Spanish-speaking samples. Invariance was found for intercepts among the college student samples, with partial invariance established for the clinical sample (indicating potential indicator bias for clinical samples). Latent means were greater for both constructs for Mexican inpatients compared with college students of the same nationality, whereas Mexican college students had greater latent means than students of other nationalities.

Reliability, Construct Validity, and Criterion Validity

The fourth aim of this study was to examine the reliability, construct validity, and criterion validity of the INQ-S-9 in an independent sample of U.S. Spanish-speaking adults (in the community). Internal consistency for the latent variables of perceived burdensomeness and thwarted belongingness on the INQ-S-9 were excellent and good, respectively. One-week test–retest reliability was excellent for perceived burdensomeness and acceptable for thwarted belongingness. Although participants may still remember their responses after 1 week, this time frame was used because perceived burdensomeness and thwarted belongingness are proximal and dynamic (fluctuating) states. Future research may examine the temporal stability of the INQ-S-9 over longer periods of time.

In regard to construct validity, evidence for convergent, but not divergent, validity was found for thwarted belongingness and perceived burdensomeness. The low divergent validity is similar to the English INQ-15 (Van Orden et al., 2012) and not necessarily unexpected, as social disconnection broadly is related to a variety of psychological outcomes (Van Orden et al., 2012). Despite this, the Thwarted Belongingness subscale did demonstrate much stronger associations with more closely related constructs (such as loneliness). Future research should continue to examine divergent validity with measures that may be less strongly associated with belonging. Finally, in regard to criterion validity, thwarted belongingness and perceived burdensomeness were also positively associated with concurrent suicide ideation (presence–absence and severity).

Implications

The current study is the first to develop and evaluate a Spanish-language version of the INQ, which can be used for future adaptation and examination by researchers and practitioners. We recommend that the INQ-S be further examined in independent samples (e.g., testing the nine- and 15-item solutions). Suggestions for further development include (a) examining whether Bel2 ("I feel like I belong") performs better if specifying belonging *to* someone or something; (b) rewording Bel7 ("I feel that there are people I can turn to in times of need") to ". . . if I need it" (i.e., ". . . acuidir si lo necesito"); and (c) using the phrase "the people

that surround me" ("la gente que me rodea") instead of "the people in my life" ("la gente en mi vida") for Bur1, Bur2, Bur4, Bur5, and Bur6. The term *cercano* instead of *cerca* for Bel8 may also be considered (note: the use of cerca vs. cercano to mean "close" varies across countries). We also recommend adapting the instructions to the form of application (e.g., changing "circle that number" to "select" or "mark") as well as the removal of the phrase "these days" ("Estos dias") from the beginning of each item to reduce instruction redundancy and aid with clarity for readers.

Given our goal of obtaining a Spanish version of the INQ as similar in item content to the English version, we did not create new items (or revise existing item core concepts). Developing items that may capture the latent constructs more accurately in Spanish (as opposed to removing items that do not culturally reflect the constructs) should be considered. However, when revising an existing scale to include new items (to improve cultural relevancy), comparisons across cultures may be limited by measurement differences of the constructs. Research is also needed to examine the predictive validity of the INQ-Spanish for suicide ideation.

Results should be considered in the context of limitations of the current project. Larger sample sizes or simulation studies could be useful in improving model fit and increasing the precision with which the latent structure of the INQ-Spanish can be modeled. Notably, items retained for the Thwarted Belongingness subscale were all reverse-worded. Researchers may consider rewording items on the Thwarted Belongingness subscale to address psychometric issues associated with reverse-wording and examine whether doing so improves item performance and retention. Furthermore, although we developed a standard Spanish version of the INQ at an accessible reading level, measurement invariance should be examined in non-college-educated and more diverse samples (e.g., outpatient). Item response theory methods can also be used to examine whether the endorsement of certain INQ-S items varies in difficulty and could indicate a more or less severe level of the constructs. This could suggest the utility of differential item weighting when calculating total scores, especially considering the partial invariance of intercepts between clinical and nonclinical (student) samples of the same nationality (Mexico).

These data have implications for the interpersonal theory as well as for clinical work with Spanish-speaking suicidal patients. According to the U.S. Census in 2010, 16% of the population (approximately 50.5 million) self-identifies as Hispanic or Latino (Ennis et al., 2011), reflecting a Hispanic population growth of 43% since 2000. The U.S. Hispanic population is expected to double to 119 million by 2060 (Colby & Ortman, 2015), constituting over 25% of the nation's population. The 2011 American Community Survey conducted by the U.S. Census Bureau indicated that almost 13% of U.S. adults report speaking Spanish in the home, and, of these, approximately 44% were classified as needing English assistance (Ryan, 2013). The changing cultural demographics in the United States have important implications for research, especially in the assessment of cross-culturally sensitive topics related to mental health. A Spanish version of the INQ will further suicide research by facilitating the testing of theoretical and etiological models of suicidal behaviors among millions of people in the United States and around the world.

The INQ has the potential to be a useful instrument for use in suicide risk assessment because it can supplement assessments that ask specifically and exclusively about desire and plans/intent for suicide through the assessment of proximal causes for suicidal desire. Additionally, the results of the present study suggest that this measure can be used across Spanish-speaking populations, in both community and clinical settings. Through the translation process, the translation team, which was comprised of native Spanish speakers of different nationalities, had extensive discussions to ensure that the meaning of the original items was retained and the measure could be generalized across a variety of Spanish speakers.

The interpersonal constructs assessed by the INQ-S can also indicate targeted areas for intervention (e.g., assisting patients in developing social support). For suicidal patients, perceived burdensomeness and thwarted belongingness can be targeted areas of crisis management and emotion regulation (e.g., contributing to others). In addition, patients who are initially reluctant to endorse suicide ideation because of cultural factors or fear of hospitalization may be more likely to report their thoughts and feelings on the INQ. This could then alert clinicians to the presence of proximal causes of suicidal desire that could indicate the presence of suicide risk. Toward this end, future research should seek to develop norms or cut-off scores at which elevated scores on the INQ-Spanish would indicate an elevated risk for suicide. Overall, given the increasing need for linguistically and culturally sensitive mental health services, continued research examining the reliability and validity of common clinical and research instruments and their translations with diverse populations remains vitally important.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

Altman DG, & Royston P (2006). The cost of dichotomising continuous variables. BMJ, 332, 1080 10.1136/bmj.332.7549.1080 [PubMed: 16675816]

Anderson JC, & Gerbing DW (1988). Structural equation modeling in practice: A review and recommended 2-step approach. Psychological Bulletin, 103, 411–423. 10.1037/0033-2909.103.3.411

Arditte KA, Çek D, Shaw AM, & Timpano KR (2016). The importance of assessing clinical phenomena in Mechanical Turk research. Psychological Assessment, 28, 684–691. 10.1037/pas0000217 [PubMed: 26302105]

Barnette JJ (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using negatively worded stems. Educational and Psychological Measurement, 60, 361–370. 10.1177/00131640021970592

Beck AT, Brown GK, & Steer RA (1997). Psychometric characteristics of the Scale for Suicide Ideation with psychiatric outpatients. Behaviour Research and Therapy, 35, 1039–1046. 10.1016/S0005-7967(97)00073-9 [PubMed: 9431735]

Beck AT, & Steer RA (1991). Manual for the Beck Scale for Suicide Ideation San Antonio, TX: Psychological Corporation.

Berkman LF, Glass T, Brissette I, & Seeman TE (2000). From social integration to health: Durkheim in the new millennium. Social Science & Medicine, 51, 843–857. 10.1016/S0277-9536(00)00065-4 [PubMed: 10972429]

Breslau J, Kendler KS, Su M, Gaxiola-Aguilar S, & Kessler RC (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. Psychological Medicine, 35, 317–327. 10.1017/S0033291704003514 [PubMed: 15841868]

Brislin RW, Lonner WJ, & Thorndike RM (1973). Cross-cultural research methods New York, NY: Wiley.

Brown TA (2015). Confirmatory factor analysis for applied research (2nd ed.). New York, NY: Guilford Press.

Bryan CJ (2011). The clinical utility of a brief measure of perceived burdensomeness and thwarted belongingness for the detection of suicidal military personnel. Journal of Clinical Psychology, 67, 981–992. 10.1002/jclp.20726 [PubMed: 20715022]

Bryan CJ, Morrow CE, Anestis MD, & Joiner TE (2010). A preliminary test of the interpersonal-psychological theory of suicidal behavior in a military sample. Personality and Individual Differences, 48, 347–350. 10.1016/j.paid.2009.10.023

Centers for Disease Control and Prevention. (2017). Fatal injury reports, national, regional and state, 1981–2016. Web-Based Injury Statistics Query and Reporting System (WISQARS) Atlanta, GA: National Center for Injury Prevention and Control Retrieved from https://webappa.cdc.gov/sasweb/ncipc/mortrate.html

Chandler J, & Shapiro D (2016). Conducting clinical research using Crowdsourced convenience samples. Annual Review of Clinical Psychology, 12, 53–81. 10.1146/annurev-clinpsy-021815-093623

Chapman DW, & Carter JF (1979). Procedures for cross cultural use of measurement instruments. Educational Evaluation and Policy Analysis, 1, 71–76. 10.3102/01623737001003071

Chen BW, Vansteenkiste M, Beyers W, Boone L, Deci E, Van der Kaap-Deeder J, . . . Verstuyf J (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. Motivation and Emotion, 39, 216–236. 10.1007/s11031-014-9450-1

Cheung GW, & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9, 233–255. 10.1207/S15328007SEM0902_5

Colby SL, & Ortman JM (2015). Projections of the size and composition of the U.S. population: 2014 to 2060 (Report No. P25–1143). Current Population Reports Washington, DC: U.S. Census Bureau Retrieved from https://www.census.gov/content/dam/Census/library/publications/2015/demo/p25-1143.pdf

Cukrowicz KC, Cheavens JS, Van Orden KA, Ragain RM, & Cook RL (2011). Perceived burdensomeness and suicide ideation in older adults. Psychology and Aging, 26, 331–338. 10.1037/a0021836 [PubMed: 21401264]

Dunn TJ, Baguley T, & Brunsden V (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. British Journal of Psychology, 105, 399–412. 10.1111/bjop.12046 [PubMed: 24844115]

Ennis SR, Rio-Vargas M, & Albert NG (2011). The Hispanic population: 2010. 2010 Census Briefs Washington, DC: U.S. Census Bureau Retrieved from https://www.census.gov/prod/cen2010/briefs/c2010br-04.pdf

Fernández Huerta J (1959). Medidas sencillas de lecturabilidad [Simple measures of readability]. Consgina, 214, 29–32.

Fortuna LR, Perez DJ, Canino G, Sribney W, & Alegria 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, 572–581. 10.4088/JCP.v68n0413 [PubMed: 17474813]

Garcia C, Skay C, Sieving R, Naughton S, & Bearinger LH (2008). Family and racial factors associated with suicide and emotional distress among Latino students. The Journal of School Health, 78, 487–495. 10.1111/j.1746-1561.2008.00334.x [PubMed: 18786041]

Garza MJ, & Cramer RJ (2011). The Spanish Reasons for Living Inventory (SRFL-I): Factor structure and association with suicide risk among Spanish speaking Hispanics. Archives of Suicide Research, 15, 354–371. 10.1080/13811118.2011.615704 [PubMed: 22023643]

Garza MJ, & Pettit JW (2010). Perceived burdensomeness, familism, and suicidal ideation among Mexican women: Enhancing understanding of risk and protective factors. Suicide and Life-Threatening Behavior, 40, 561–573. 10.1521/suli.2010.40.6.561 [PubMed: 21198325]

Hair JF, Black WC, Babin BJ, & Anderson RE (2010). Multivariate data analysis (7th ed). Upper Saddle River, NJ: Pearson Prentice Hall.

Hill RM, Rey Y, Marin CE, Sharp C, Green KL, & Pettit JW (2015). Evaluating the Interpersonal Needs Questionnaire: Comparison of the reliability, factor structure, and predictive validity across five versions. Suicide and Life-Threatening Behavior, 45, 302–314. 10.1111/sltb.12129 [PubMed: 25308815]

Hooper D, Coughlan J, & Mullen M (2008). Structural equation modelling: Guidelines for determining model fit. Electronic Journal of Business Research Methods, 6, 53–60.

Hoppe SK, & Martin HW (1986). Patterns of suicide among Mexican Americans and Anglos, 1960 – 1980. Social Psychiatry, 21, 83–88. 10.1007/BF00578747 [PubMed: 3726604]

Hu LT, & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55. 10.1080/10705519909540118

Joiner TE (2005). Why people die by suicide Cambridge, MA: Harvard University Press.

Joiner TE, Ribeiro JD, & Silva C (2012). Nonsuicidal self-injury, suicidal behavior, and their co-occurrence as viewed through the lens of the interpersonal theory of suicide. Current Directions in Psychological Science, 21, 342–347. 10.1177/0963721412454873

Joiner TE, Jr., Van Orden KA, Witte TK, Selby EA, Ribeiro JD, Lewis R, & Rudd MD (2009). Main predictions of the interpersonal-psychological theory of suicidal behavior: Empirical tests in two samples of young adults. Journal of Abnormal Psychology, 118, 634–646. 10.1037/a0016500 [PubMed: 19685959]

Koo TK, & Li MY (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15, 155–163. 10.1016/j.jcm.2016.02.012 [PubMed: 27330520]

Landis RS, Edwards BD, & Cortina JM (2009). On the practice of allowing correlated residuals among indicators in structural equation models. In Lance CE & Vandenberg RJ (Eds.), Statistical and methodological myths and urban legends: Doctrine, verity, and fable in the organizational and social sciences (pp. 193–215). New York, NY: Routledge.

Limon FJ, Lamson AL, Hodgson J, Bowler M, & Saeed S (2016). Screening for depression in Latino immigrants: A systematic review of depression screening instruments translated into Spanish. Journal of Immigrant and Minority Health, 18, 787–798. 10.1007/s10903-015-0321-y [PubMed: 26660481]

Linehan MM, Goodstein JL, Nielsen SL, & Chiles JA (1983). Reasons for staying alive when you are thinking of killing yourself: The reasons for living inventory. Journal of Consulting and Clinical Psychology, 51, 276–286. 10.1037/0022-006X.51.2.276 [PubMed: 6841772]

Mantuliz MCA, & Castillo CM (2002). Validación de una escala de apoyo social percibido en un grupo de adultos mayores adscritos a un programa de hipertensión de la región metropolitana [Validation of a scale of perceived social support in a group of elders under control in a hypertension program in the metropolitan region]. Ciencia y Enfermería, 8, 49–55.

Martín-Albo J, Núñez JL, Navarro JG, & Grijalvo F (2007). The Rosenberg Self-Esteem Scale: Translation and validation in university students. The Spanish Journal of Psychology, 10, 458–467. 10.1017/S1138741600006727 [PubMed: 17992972]

McDonald RP (1999). Test theory: A unified treatment Mahwah, NJ: Erlbaum.

Morejón A, & García-Bóveda RJ (1994). RULS: Escala de soledad UCLA revisada: Fiabilidad y validez de una versión Española [RULS: The revised UCLA Loneliness Scale: Reliability and validity of a Spanish version]. Revista Psicologia e Saúde, 6, 45–54.

Muñoz Fernández A (2016). Legibilidad 2 [Readability 2]. Retrieved from https://legible.es

Muthén LK, & Muthén BO (1998 –2015). Mplus user's guide (7th ed). Los Angeles, CA: Author.

Oquendo MA, Ellis SP, Greenwald S, Malone KM, Weissman MM, & Mann JJ (2001). Ethnic and sex differences in suicide rates relative to major depression in the United States. The American Journal of Psychiatry, 158, 1652–1658. 10.1176/appi.ajp.158.10.1652 [PubMed: 11578998]

Oquendo MA, Lizardi D, Greenwald S, Weissman MM, & Mann JJ (2004). Rates of lifetime suicide attempt and rates of lifetime major depression in different ethnic groups in the United States. Acta

Psychiatrica Scandinavica, 110, 446–451. 10.1111/j.1600-0447.2004.00404.x [PubMed: 15521829]

Osman A, Kopper B, Linehan MM, Barrios FX, Gutierrez PM, & Bagge CL (1999). Validation of the adult suicidal ideation questionnaire and the reasons for living inventory in an adult psychiatric inpatient sample. Psychological Assessment, 11, 115–123. 10.1037/1040-3590.11.2.115

Pearson Education. (2018). Beck Scale for Suicide Ideation (BSS) – Spanish record forms Retrieved from https://www.pearsonclinical.com/psychology/products/100000157/beck-scale-for-suicide-ideation-bss.html

Rasmussen KM, Negy C, Carlson R, & Burns JM (1997). Suicide ideation and acculturation among low socioeconomic status Mexican American adolescents. The Journal of Early Adolescence, 17, 390–407. 10.1177/0272431697017004003

Rivera FI, Guarnaccia PJ, Mulvaney-Day N, Lin JY, Torres M, & Alegría M (2008). Family cohesion and its relationship to psychological distress among Latino groups. Hispanic Journal of Behavioral Sciences, 30, 357–378. 10.1177/0739986308318713 [PubMed: 19444326]

Roberts RE, & Chen YW (1995). Depressive symptoms and suicidal ideation among Mexican-origin and Anglo adolescents. Journal of the American Academy of Child & Adolescent Psychiatry, 34, 81–90. 10.1097/00004583-199501000-00018 [PubMed: 7860463]

Robins RW, Hendin HM, & Trzesniewski KH (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. Personality and Social Psychology Bulletin, 27, 151–161. 10.1177/0146167201272002

Rosenberg M (1965). Society and the adolescent self-image (Vol. 11). Princeton, NJ: Princeton University Press 10.1515/9781400876136

Russell D, Peplau LA, & Cutrona CE (1980). The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. Journal of Personality and Social Psychology, 39, 472–480. 10.1037/0022-3514.39.3.472 [PubMed: 7431205]

Ryan C (2013). Language use in the United States: 2011 Washington, DC: U.S. Census Bureau Retrieved from https://www.census.gov/prod/2013pubs/acs-22.pdf

Ryan C, Huebner D, Diaz RM, & Sanchez J (2009). Family rejection as a predictor of negative health outcomes in white and Latino lesbian, gay, and bisexual young adults. Pediatrics, 123, 346–352. 10.1542/peds.2007-3524 [PubMed: 19117902]

Segura SL, & González-Romá V (2003). How do respondents construe ambiguous response formats of affect items? Journal of Personality and Social Psychology, 85, 956–968. 10.1037/0022-3514.85.5.956 [PubMed: 14599257]

Silva C, & Van Orden KA (2018). Suicide among Hispanics in the United States. Current Opinion in Psychology, 22, 44–49. 10.1016/j.copsyc.2017.07.013 [PubMed: 30122277]

Sorenson SB, & Golding JM (1988). Suicide ideation and attempts in Hispanics and non-Hispanic whites: Demographic and psychiatric disorder issues. Suicide and Life-Threatening Behavior, 18, 205–218. 10.1111/j.1943-278X.1988.tb00157.x [PubMed: 3188137]

Tortolero SR, & Roberts RE (2001). Differences in nonfatal suicide behaviors among Mexican and European American middle school children. Suicide and Life-Threatening Behavior, 31, 214–223. 10.1521/suli.31.2.214.21519 [PubMed: 11459254]

Ungemack JA, & Guarnaccia PJ (1998). Suicidal ideation and suicide attempts among Mexican Americans, Puerto Ricans, and Cuban Americans. Transcultural Psychiatry, 35, 307–327. 10.1177/136346159803500208

Van Orden KA, Cukrowicz KC, Witte TK, & Joiner TE (2012). Thwarted belongingness and perceived burdensomeness: Construct validity and psychometric properties of the Interpersonal Needs Questionnaire. Psychological Assessment, 24, 197–215. 10.1037/a0025358 [PubMed: 21928908]

Van Orden KA, Witte TK, Cukrowicz KC, Braithwaite SR, Selby EA, & Joiner TE, Jr. (2010). The interpersonal theory of suicide. Psychological Review, 117, 575–600. 10.1037/a0018697 [PubMed: 20438238]

Van Orden KA, Witte TK, Gordon KH, Bender TW, & Joiner TE, Jr. (2008). Suicidal desire and the capability for suicide: Tests of the interpersonal-psychological theory of suicidal behavior among adults. Journal of Consulting and Clinical Psychology, 76, 72–83. 10.1037/0022-006X.76.1.72 [PubMed: 18229985]

Wadsworth T, & Kubrin CE (2007). Hispanic suicide in U.S. metropolitan areas: Examining the effects of immigration, assimilation, affluence, and disadvantage. American Journal of Sociology, 112, 1848–1885. 10.1086/512711

Wiebe JS, & Penley JA (2005). A psychometric comparison of the Beck Depression Inventory-II in English and Spanish. Psychological Assessment, 17, 481–485. 10.1037/1040-3590.17.4.481 [PubMed: 16393015]

Zayas LH, Lester RJ, Cabassa LJ, & Fortuna LR (2005). Why do so many Latina teens attempt suicide? A conceptual model for research. American Journal of Orthopsychiatry, 75, 275–287. 10.1037/0002-9432.75.2.275 [PubMed: 15839764]

Zimet GD, Dahlem NW, Zimet SG, & Farley GK (1988). The multidimensional scale of perceived social support. Journal of Personality Assessment, 52, 30–41. 10.1207/s15327752jpa5201_2

Public Significance Statement

The present study provides a 9-item Spanish-language version of the Interpersonal Needs Questionnaire to promote research on perceived burdensomeness and thwarted belongingness—two robust suicide risk factors—among Hispanic populations. The study details the translation process as well as the factor structure, measurement invariance, reliability, construct validity, and criterion validity of the measure.
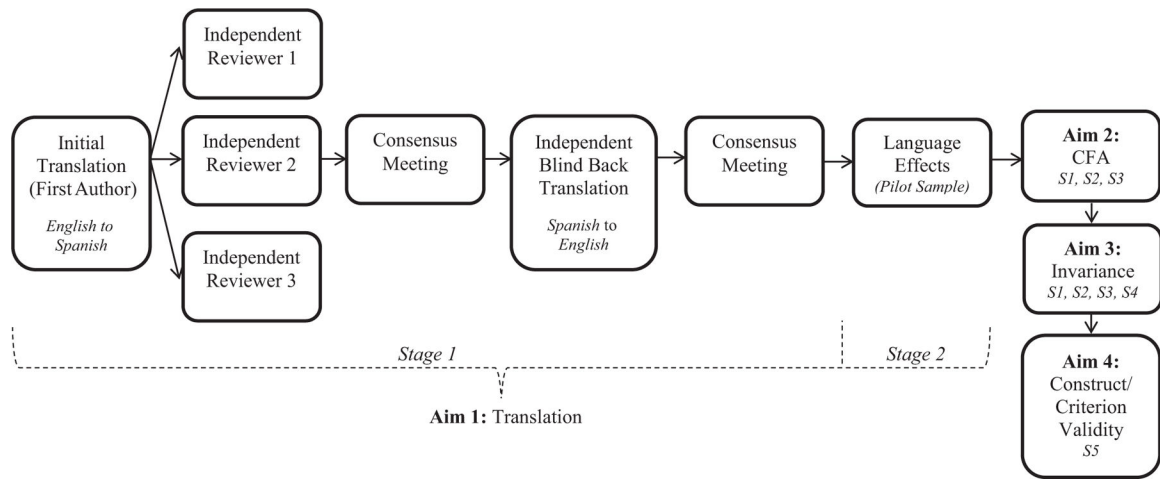
Figure 1.
Flowchart of translation and validation procedures. Pilot Sample = U.S. bilingual college students; S1 = Sample 1 (U.S. college students); S2 = Sample 2 (Spain college students); S3 = Sample 3 (Mexico psychiatric inpatients); S4 = Sample 4 (Mexico college students); S5 = Sample 5 (U.S. Spanish-speaking adults).
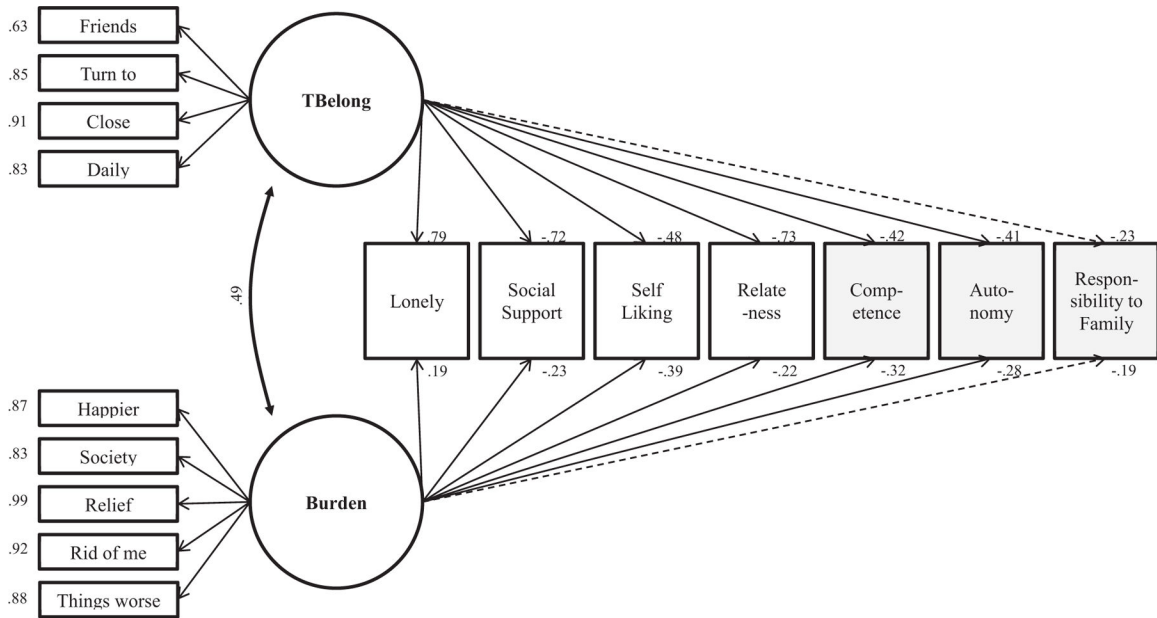
Figure 2.
Structural equation model results examining construct validity in Spanish-speaking U.S. adults (Sample 5). Standardized loadings appear on the lines; dotted lines indicate statistically insignificant loadings; all other loadings are significant at $p < .05$. Gray-shaded boxes (observed variables) were posited to be more strongly related to perceived burdensomeness, whereas the unshaded boxes (excluding the measurement indicators) were posited to be more strongly related to Thwarted Belongingness. Lonely = Revised UCLA Loneliness Scale; Social Support = Multidimensional Scale of Perceived Social Support; Self-Liking = Rosenberg Self-Esteem Scale; Relatedness, Competence, Autonomy = same-named satisfaction subscales of the Basic Psychological Needs Satisfaction and Frustration Scale; Responsibility to Family = subscale of the Reasons for Living Scale.

Table 1

Fit Statistics for Confirmatory Factor Analysis Two-Factor Models for Samples 1, 2, and 3

Sample	χ^2	df	p	CFI	TLI	RMSEA [90% CI]	SRMR
INQ-S-15							
Sample 1 (United States; N = 281)	228.44	89	.00	.86	.83	.08 [.06, .09]	.07
Sample 2 (Spain; N = 1,016)	669.17	89	.00	.85	.82	.08 [.07, .09]	.09
Sample 3 (Mexico; N = 181)	229.34	89	.00	.78	.74	.09 [.08, .11]	.11
INQ-S-9							
Sample 1 (United States; N = 281)	32.26	26	.18	.99	.98	.03 [.00, .06]	.05
Sample 2 (Spain; N = 1,016)	71.74	26	.00	.97	.96	.04 [.03, .05]	.04
Sample 3 (Mexico; N = 181)	32.86	26	.17	.98	.97	.04 [.00, .07]	.05

Note. INQ-S-9 removes Perceived Burdensomeness subscale Item 1 and Thwarted Belongingness subscale Items 1, 2, 3, 5, and 6. *df* = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; INQ-S-15 = 15-item Interpersonal Needs Questionnaire-Spanish; INQ-S-9 = nine-item Interpersonal Needs Questionnaire-Spanish.

Table 2

Factor Loadings for Confirmatory Factor Analysis Two-Factor Models for Samples 1, 2, and 3

Variable	Item	Item content	Sample 1		Sample 2		Sample 3	
			B	R^2	β	R^2	β	R^2
INQ-S-15								
Perceived Burdensomeness	Bur1	Better off	.81**	.66**	.94**	.88**	.47**	.22**
	Bur2	Happier w/o me	.93**	.87**	.95**	.91**	.79**	.62**
	Bur3	Burden to society	.69**	.47**	.69**	.48**	.70**	.50**
	Bur4	Death as relief	.72**	.52**	.72**	.52**	.67**	.45**
	Bur5	Rid of me	.73**	.54**	.61**	.37**	.67**	.45**
	Bur6	Makes worse	.66**	.44**	.73**	.54**	.64**	.41**
Thwarted Belongingness	Bel1	Others care	.21**	.04	.30**	.09**	.72**	.52**
	Bel2	I belong	.61**	.38**	.23**	.05**	.72**	.52**
	Bel3	Rarely interact	.28**	.08**	.36**	.13**	−.27**	.07
	Bel4	Friends	.73**	.54**	.72**	.53**	.68**	.46**
	Bel5	Disconnected	.55**	.30**	.54**	.29**	.06	.003
	Bel6	Outsider	.42**	.17**	.57**	.32**	.11	.012
	Bel7	Turn to	.64**	.41**	.77**	.59**	.52**	.27**
	Bel8	Close to others	.81**	.66**	.85**	.72**	.67**	.46**
	Bel9	Daily interact	.75**	.57**	.74**	.54**	.56**	.31**
BEL w/ BUR			.50**		.38**		.25**	
INQ-S-9								
Perceived Burdensomeness	Bur2	Happier w/o me	.86**	.75**	.89**	.79**	.78**	.61**
	Bur3	Burden to society	.74**	.54**	.74**	.55**	.71**	.51**
	Bur4	Death as relief	.79**	.62**	.76**	.58**	.70**	.49**
	Bur5	Rid of me	.75**	.56**	.65**	.43**	.67**	.45**
	Bur6	Makes worse	.67**	.45**	.76**	.57**	.62**	.38**

Variable	Item	Item content	Sample 1		Sample 2		Sample 3	
			B	R^2	β	R^2	β	R^2
Thwarted Belongingness	Bel4	Friends	.70 **	.50 **	.73 **	.53 **	.63 **	.40 **
	Bel7	Turn to	.66 **	.43 **	.80 **	.64 **	.65 **	.42 **
	Bel8	Close to others	.81 **	.65 **	.84 **	.71 **	.68 **	.46 **
	Bel9	Daily interact	.79 **	.62 **	.74 **	.55 **	.62 **	.38 **
BEL w/ BUR			.48 **		.33 **		.16	

Note. INQ-S-9 removes Perceived Burdensomeness subscale Item 1 and Thwarted Belongingness subscale Items 1, 2, 3, 5, and 6. The first item for each latent variable was set to 1.00 to handle scale dependency. β = standardized (STDYX) estimate; INQ-S-15 = 15-item Interpersonal Needs Questionnaire-Spanish; INQ-S-9 = nine-item Interpersonal Needs Questionnaire-Spanish; Bel = thwarted belongingness; Bur = perceived burdensomeness item; BEL w/ BUR = thwarted belongingness and perceived burdensomeness latent covariance; w/o = without.

*
$p < .05$

**
$p < .001$.

Table 3

Tests of Measurement Invariance for INQ-S-9 for Samples 1, 2, and 3 Compared With Sample 4 (Mexico)

Group/variable	χ^2	df	MLR scaling	$\text{YB}\chi^2_{\text{diff}}$	df	CFI	CFI	TLI	RMSEA [90% CI]	SRMR
Sample 1 (United States)										
Equal forms	61.93	52	1.66			.99		.99	.03 [.00, .05]	.04
Equal factor loadings	75.66	59	1.90	11.10	7	.98	.007	.98	.03 [.00, .05]	.06
Equal indicator intercepts	81.81	66	1.81	4.20	7	.98	−.001	.98	.03 [.00, .05]	.06
Sample 2 (Spain)										
Equal forms	146.53**	52	1.62			.91		.87	.09 [.07, .10]	.05
Equal factor loadings	135.13**	59	1.95	6.10	7	.93	−.017	.91	.07 [.06, .09]	.06
Equal indicator intercepts	145.96**	66	1.87	7.67	7	.92	.003	.92	.07 [.06, .09]	.06
Sample 3 (Mexico)										
Equal forms	62.18	52	1.48			.99		.98	.03 [.00, .06]	.04
Equal factor loadings	75.99	59	1.50	13.38	7	.98	.008	.98	.04 [.00, .06]	.06
Equal indicator intercepts	98.02*	66	1.45	26.94**	7	.96	.017	.96	.05 [.03, .07]	.06

Note. $\text{YB}\chi^2_{\text{diff}}$ reflects the nested difference tests between the previous and more restrictive solution (i.e., equal forms vs. equal factor loadings; equal factor loadings vs. equal indicator intercepts). INQ-S-9 = nine-item Interpersonal Needs Questionnaire-Spanish; df = degrees of freedom; MLR = robust maximum likelihood; $\text{YB}\chi^2_{\text{diff}}$ = Yuan-Bentler scaled χ^2 difference; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual.

*
$p < .05$

**
$p < .001$.