HHS Public Access

Author manuscript

Fam Community Health. Author manuscript; available in PMC 2022 January 01.

Published in final edited form as:

Fam Community Health. 2021 ; 44(1): 1–9. doi:10.1097/FCH.0000000000000279.

Heart Attacks, Bloody Noses, and Other "Emotional Problems":

Cultural and Conceptual Issues With the Spanish Translation of Self-Report Emotional Health Items

Michael A. Flynn, MA, **Donald E. Eggerth, PhD**, **C. Jeffery Jacobson Jr, PhD**, **Sarah M. Lyon, PhD**

Division of Science Integration, National Institute for Occupational Safety and Health, Cincinnati, Ohio (Mr Flynn and Dr Eggerth); Consortium for Multicultural Psychology Research, Michigan State University, East Lansing (Mr Flynn and Dr Eggerth); Departments of Anthropology (Dr Jacobson) and Family and Community Medicine (Dr Jacobson), University of Cincinnati, Cincinnati, Ohio; and Department of Anthropology, University of Kentucky, Lexington (Dr Lyon).

Abstract

This article examines how respondents understood items in the Spanish versions of the Short-Form 36 (SF-36v2). Cognitive interviews of the SF-36 were conducted in 2 phases with 46 Spanish speakers living in the United States. Roughly one-third (17/46) of respondents had difficulty understanding the Role Emotional items upon their initial reading, and almost half (21/46) provided examples that were inconsistent with the intended meaning of the items. The findings of this study underscore the importance of conducting cognitive testing to ensure conceptual equivalence of any instrument regardless of how well validated it appears to be.

Keywords

cognitive testing; conceptual equivalence; Latinx immigrants; mental health; SF-36

INTERNATIONAL MIGRATION has long been understood as a significant source of stress even under the best circumstances.[1,2] Some studies have shown increased psychological distress for immigrants relative to their counterparts back home.[3,4] Other studies have shown that immigrants report better mental health than United States (US)-born minorities and that the mental health of immigrants declines the longer they live in the US.[5–8] These findings have increased interest in mental health research with immigrants. One key area of interest has been finding accurate ways of measuring and tracking the mental health of these communities.

Self-report health status questionnaires have become core tools in both clinical and community-based health research, and they play an important role in documenting health

Correspondence: Michael A. Flynn, MA, Division of Science Integration, National Institute for Occupational Safety and Health, 1090 Tusculum Ave, Cincinnati, OH 45226 (mflynn@cdc.gov).

The findings and conclusions in this article are those of the authors and do not necessarily represent the views of the NIOSH.

The authors declare they have no real or potential conflict(s) of interest that may inappropriately impact or influence the research and interpretation of findings.

disparities in the US and internationally. These instruments are very attractive to researchers because they are short, use general terminology, and have been correlated with objective clinical measures and outcomes. However, there is growing recognition that the meanings of the construct items (conceptual equivalence) and the psychometric properties of the items (eg, measurement invariance) can change when used with individuals from different racial, ethnic, and socioeconomic groups.[9,10] Yet, conceptual equivalence is often overlooked or not described in survey development.[11] Threats to conceptual equivalence, such as understandings of health and illness that differ from the biomedical Western model, educational differences, local dialects, etc, may cause the respondent to understand the item differently from that by the researcher, limiting the applicability of such tools across populations.

Establishing the conceptual equivalence of a new instrument is a lengthy, resource-intensive process involving translation, back translation, and many iterative rounds of cognitive testing of items with members of the intended target audience.[12–14] Many research groups lack the internal expertise and/or capacity to conduct this process adequately. In addition, when the intended audience is perceived as a small or challenging-to-recruit population, researchers may be reluctant to “waste” potential research subjects in an activity that they do not consider central to the research question they intend to answer. To avoid these problems, many researchers turn to already existing instruments with the expectation that the creators of these instruments have exercised due diligence in their development. One of the most widely used health status questionnaires is Quality Metric’s Short-Form 36 (SF-36). The measure consists of 36 items that assess 8 subscales of health (Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional, and Mental Health).[15] However, some research suggests that respondent gender, age, race, ethnicity, socioeconomic status, language, and nationality may impact the validity of SF-36 items.[16–20] Although this article presents findings from cognitive testing of the SF-36, it is important to recognize that threats to cross-cultural equivalence are not a problem unique to that instrument. It is hoped that highlighting the shortcomings of an instrument as ubiquitous, if not reified, in the biomedical outcome literature as the SF-36 will serve as a cautionary tale for researchers conducting cross-cultural research. This article is an exploration of how important it is to conduct cognitive testing to ensure conceptual equivalence of any instrument that is being used with any population for the first time—particularly, populations differing markedly from the investigators in terms of language, culture, and/or literacy.

THE SF-36

The SF-36 was originally developed for use in the US and has consistently shown to be a valid measure of health and a strong predictor of future morbidity.[21–23] Research conducted by McHorney et al,[24] using a large clinical sample, demonstrated that the Physical Health and the Role Emotional scales were robust enough to reliably discriminate between diagnostic groups and to predict diagnostic group membership. In 1996, an ambitious effort was undertaken to develop validated translations of the instrument for use in international and multinational health studies.[25] The initial translation effort involved 14 countries (13 in Western Europe and Japan) and 10 languages and used translation/back translation and pilot

testing with focus group feedback to ensure content validity.[26] Problems adapting the instrument were identified and corrected, resulting in a second version of the tool (SF-36v2), which was shown to have good psychometric properties for both the emotional and physical health items within each of the populations studied.[27,28] As a result, the SF-36v2 was dubbed the *international version* and has subsequently been translated into more than 170 languages/dialects and is used across the globe.[29] However, these subsequent translation efforts have not been as well investigated as the original, and questions about cross-cultural validity have surfaced when the instrument is used in non-Western populations such as Japanese[30] and older ethnic Maori and Samoans in New Zealand.[31] Scott and colleagues[31] have argued that these complications suggest that the 2-dimensional structure (physical/mental) of the instrument does not hold up well when used with cultures lacking a straightforward mind-body dualism that is central to the Western concept of health.

Spain was the only Spanish-speaking country to participate in the original translation effort.[32] That study found that the expression *problemas emocionales* (emotional problems), which appears in all 3 items of the Role Emotional scale, was not readily understood. To clarify the term, a parenthetical qualifier "such as feeling depressed or anxious" was added to the end of the first of the 3 items (Table).[28] While this improved comprehension for Spanish respondents, in a separate study in Southeast Mexico, Zuniga and colleagues[33] reported complications with the same 3 items in the Role Emotional scale despite the inclusion of the parenthetical qualifier. They found that the Role Emotional scale was the only one for which respondents, recruited from a clinical setting, reported better average health than the control group of government employees. In addition, it was the only scale whose internal consistency (Cronbach $\alpha = 0.56$) was below desirable levels. Translation problems and cultural factors impacting respondent understanding of the items were among the reasons suggested for the inconsistencies. They recommend further investigation of how the items were being understood by the target audience.

There are currently 15 Spanish language versions of the SF-36 for use in specific countries throughout Latin American as well as in the US.[29] Many of these versions differ from one another only in terms of the behaviors used to illustrate item intent. For example, the English language version contains an item asking whether the respondent has difficulty performing effortful physical activities such as "moving furniture." The Mexican-Spanish version substitutes "playing baseball" for "moving furniture" and the Guatemalan version references "dancing" in this item, but these versions do not differ substantially from one another.

In general, studies evaluating different country-specific Spanish versions of the SF-36v2 conclude that they are appropriate for use with the target population.[32,34–37] However, despite the extensive body of literature on establishing measurement equivalence in cross-cultural and cross-ethnic research,[9,38–41] such conclusions are almost inevitably based upon a single measure of the psychometric properties of scale items—the coefficient α. It appears implicit to these arguments that as long as an acceptable level of α is achieved, similar to that achieved in the development of the original version, then one is safe assuming similar levels of internal consistency for the newer version. Although such "bootstrapping" is widely reported in the literature, it stretches the meaning of α far beyond acceptable bounds. Coefficient α represents the mean of all possible split-half correlations using the Rulon

method and, as used in the manner discussed here, is an index of internal consistency.[42] As such, it is a measure of reliability rather than validity. Consider the analogy of a team of horses harnessed to a wagon. Coefficient α indicates whether the horses are pulling in the same direction (reliability). However, it says nothing about whether they are on the right track (validity).

In an effort to better understand the cultural and conceptual issues with the Spanish translation of self-report emotional health items, the current study presents findings from cognitive interviews with Spanish-speaking individuals in the US. It focuses on the 3 items in the Role Emotional scale (Table) and, in particular, the respondents' understanding of the expression "emotional problems" that appears in each of these items.

METHODS

This study collected data in 2 phases. The first phase of data collection was part of a broader, qualitative, cross-cultural investigation of cultural and ethnic variations of health cognitions related to the self-reporting of health. The data collection indicated significant problems with comprehension of items contained in the Role Emotional scale among the Spanish-speaking immigrants in the study. To follow up on this initial finding, it was decided to conduct a second phase of data collection with a new sample of Spanish-speaking immigrants, focusing on comprehension of the Role Emotional items.

Phase 1

Participants—Participants for this phase were a convenience sample (n = 11; 6 men, 5 women) recruited through local community centers in Cincinnati, Ohio. All participants were born outside the US (Guatemala: n = 8; Mexico: n = 3). No identifying information was collected from participants.

The ages of the participants ranged from 18 to 58 years, with an average of 31 years. Participants had completed between 2 and 9 years of formal education, with an average education level of 6 years. Five (4 Guatemalans, 1 Mexican) respondents identified as Indigenous, with the remainder identifying as mixed race (Ladino/Mestizo). All reported being fluent in Spanish. Three of the Guatemalans reported speaking an Indigenous language (Mam) in addition to Spanish. All other respondents identified as native Spanish speakers.

Data collection—The 30- to 60-min cognitive interviews incorporated methods recommended by the National Center for Health Statistics and were conducted by bilingual members of the research team.[43,44] The interviews were highly structured and focused on eliciting participant understandings of the concept and their decision-making processes in the selection of a response. The interviews were audio-recorded and transcribed verbatim in Spanish, with any inadvertent personal identifiers removed. The data collection in phase 1 was reviewed and approved by the University of Cincinnati Institutional Review Board.

Analysis—The Spanish language transcripts were independently reviewed by the bilingual members of the research team. Analysts were tasked with assessing and noting respondent

comprehension difficulties and/or alternative conceptual understandings for each of the items under consideration. The Spanish-speaking members of the team then met to discuss those items for which problems in comprehension or interpretation were most discernible. Among other items, responses to the Role Emotional items were identified inductively as worthy of further analysis and characterization. To broaden the analytic team and facilitate English language dissemination of our findings, we then translated responses to the Role Emotional items into English and met on several occasions to discuss and further characterize the various types of comprehension problems apparent to us. Specific responses were discussed by the team, sorted into the categories and subcategories discussed later, and recorded and tracked using NVivo software and an Excel spreadsheet.

Phase 2

Participants—Participants (n = 35) for phase 2 of this study were recruited as a convenience sample at several immigrant-serving, nongovernmental organizations in Louisville, Kentucky. For phase 2 of this study, participants had to meet the following criteria: have no more than a high school education; be native Spanish speakers or proficient Spanish speakers (in the case of the 4 indigenous participants who speak Spanish as a second language); be of Latin American nativity; and currently reside in the US.

A total of 35 individuals (25 women, 10 men) were interviewed during phase 2 of the study. All participants were born outside the US (Mexico: n = 17; Guatemala: n = 10; Honduras: n = 4; Cuba: n = 2; El Salvador: n = 1; Colombia: n = 1). The ages of the participants ranged from 24 to 65 years, with an average of 39 years. Participants completed between 0 and 12 years of formal education, with an average education level of 6.8 years. The majority of participants (80%; n = 28) reported having migrated from a rural area of their home country, with the remainder (20%; n = 7) reported being from an urban area. All participants reported being fluent in Spanish; however, 2 respondents reported being non-native Spanish speakers. Both identified Mam as their first language. An additional 5 participants identified as being able to speak a second language (English: n = 3; Mam: n = 1; Cakchiquel: n = 1).

Data collection—The 20-minute cognitive interviews again incorporated methods recommended by the National Center for Health Statistics.[43,44] The interviews were conducted by 2 bilingual contractors, each having many years of experience working with members of the Latin American immigrant community in the region. Both were fluent Spanish speakers. The interviews were highly structured and focused on eliciting participant understandings of the central concepts contained in the Role Emotional scale (Table) from the SF-36. Specifically, respondents were asked a series of probes focused on eliciting their understanding of the terms *salud fisica* [physical health] and *problemas emocionales* [emotional problems]. Participants were also asked to describe the distinction between the 2 concepts. All interviews were digitally recorded, and the researchers also took handwritten notes on the printed surveys. No identifying information was collected from participants. Participants were compensated with a $15 gift card for their time. The data collection in phase 2 was reviewed and approved by the University of Kentucky Institutional Review Board.

Analysis—The Spanish language audio recordings and interviewer notes were transcribed and reviewed using the same criteria used in phase 1 of this study. Our initial review of these transcripts revealed comprehension problems similar to those seen in phase 1, and so the team applied (deductively) phase 1 categories of analysis and sorting to phase 2 data. Responses were catalogued and tracked using NVivo software and an Excel spreadsheet. Once this initial sorting was complete, responses that were determined by the team to represent a nuanced or unique understanding of the item were analyzed by the group. Our further analyses and subcategorizations of data collected in both phases can be found in the "Results" and "Discussion" sections.

RESULTS

Analysis of the responses found evidence of problems with conceptual equivalence for the 3 items on the Role Emotional scale. This scale consists of 3 items, all of which contain the term and idea *problema emocional* (emotional problem) as a central construct (Table). While there are unique versions of the SF-36 for Guatemala, Mexico, and US-Spanish, the translations vary little from one country to another. For example, the Mexican and Guatemalan versions differ by just 3 words. The wording of the Role Emotional items is identical in all 3 versions. The term *problemas emocionales* elicited a wide range of responses that could be categorized into 3 themes: (1) the term was understood as intended; (2) the term was not understood; and (3) the term was understood differently from its intended meaning.

Understood the items as intended

Upon initial reading of the items, roughly half of the respondents (n = 21) provided explanations of the phrase *problemas emocionales* that were consistent with its intended meaning. The most common examples of emotional problems given related to mood (ie, feeling sad, stressed, depressed) and relationships (ie, being separated from spouse and children, fighting with family members). Follow-up probes were generally unsuccessful in eliciting more information from respondents to determine whether these were ongoing problems or simply temporary, everyday experience of negative emotions. For example, one respondent who initially reported not understanding the term tentatively offered the example of not eating when she "is angry" (*enojada*). It was determined that while some of the responses were not as clear as others, they all represented familiarity with the term "emotional problem" and generally demonstrated an understanding of the item as intended.

Difficulty understanding the item meaning

Slightly more than a third (n = 17) of the respondents had difficulty understanding the Role Emotional items as intended upon their initial reading. These respondents required further explanation, beyond simple repetition of the item, before they could provide a response consistent with the intended meaning. As a result, they are considered as having difficulty understanding the items. However, once additional explanations of the items and the key concept of *problemas emocionals* were provided, 11 respondents were able to answer the original items and provide examples of emotional problems that were consistent with the meaning of the concept as intended. Follow-up explanations often included examples of

specific emotional problems and symptoms, such as feeling sad, lonely, or without energy. Six respondents were unable to provide an example or explanation of an emotional problem even after the concept was described by the researcher. The excessive explanation of the items may have limited the number of "I don't understand" responses that were received, and some respondents may have ultimately answered "never" to the original items simply because they were not sure they fully understood the question. In addition, the extensive probing resulted in multiple answers from a single participant, some of which could be classified as one category of response with others being classified in another category. As a result, some respondents appear in more than 1 category.

Probes of the parenthetical explanations for emotional problems (depressed and anxious) contained in item 5 (see the Table) of the SF-36 were generally understood. The majority of respondents were able to generate examples of both depression (n = 34) and anxiety (n = 30) with limited probing or explanation from the researchers. *Ansioso* was generally understood as something someone wants and cannot get, anxiety, and feelings of desperation (or a feeling of panic). Most participants associated *deprimido* with feelings of sadness, loneliness, and a lack of interest and/or motivation. Only 6 respondents reported not understanding the term *ansioso* (anxious), and only one reported not understanding the term *deprimido* (depressed) after being provided additional explanation.

Different understanding of the item meaning

Almost half of respondents (n = 21) reported understanding the items but gave examples and explanations that were inconsistent with the intended meaning of the items. When asked to discuss their understanding of the question, the words *problemas emocionales* were consistently described as a positively inflected emotional outburst involving, what in English equates to exhilaration, excitement, or surprise. Some of the respondents gave examples of physical problems that resulted from getting overexcited. For example, one respondent described a *problema emocional* as the bloody nose he received as a child when he fell off his horse after he got excited and tried to make the horse jump. Other examples included physical problems, such as a heart attack, resulting from the shock of a surprise party or being overcome with joy.

Some respondents (n = 11) described negatively inflected emotions such as disappointment or anger that followed from unmet expectations. Some of the examples given referred to situations that resulted in temporary negative emotional responses rather than ongoing conditions. For example, one respondent gave the example of being turned down when asking someone out to dinner. While these responses appropriately identify emotional rather than physical problems, most gave examples where an anticipated event did not occur. These examples generally referred to a situation that evoked a temporary negative emotional response such as disappointment (*desilusionado*) as opposed to an ongoing emotional state such as depression or anxiety. They were classified as being understood differently, given the temporary nature of the negative emotion that was exclusively the consequence of an anticipated positive event not coming to fruition. No examples were given that did not include a positively inflected antecedent to the negatively inflected emotions being described. Others (n = 5), however, described more chronic feelings of disappointment such

as the ongoing inability to "get ahead" (*salir adelante*) in the US despite taking the opportunity seriously. These were classified as consistent with the intended meaning of the item, given their chronic nature, and are not included this section.

These descriptions represent an understanding of the question that is distinct from the item intent, although the term "emotional problems" was correctly translated in the instrument as *problemas emocionales*. These data suggest that the misunderstanding may come from the confusion of the adjective *emocional*, which means "emotional" with the verb *emocionarse*, which means "to get excited." Many respondents used the verb *emocionarse* in their explanation of *problemas emocionales*. When probed, they describe physical problems resulting from overexuberance or name negatively inflected emotions that resulted from unfulfilled expectations. However, the initial association was exclusively with positive events. Even after probing, some respondents still did not consider negatively inflected events or emotions to be related to the phrase *problema emocional* despite the fact that it contains the word *problema*, which has clear negative connotations in Spanish.

Another area of misunderstanding (n = 3) dealt with an individual's ability to perceive the problem. These respondents described emotional problems as those that occurred inside the body and were often not felt or noticed as opposed to a physical problems that occurred on the outside of the body such as getting hit by something. For example, one respondent identified cancer as an emotional problem because, she explained, "it is inside you and you don't even know it's there." In contrast, she explained that a cough or a fever was a physical problem because they had outward manifestations and could be perceived.

A final type of alternate understanding had to do with who was experiencing the problem. When asked to describe the difference between physical and emotional problems, 4 respondents classified physical problems as those having to do with oneself and emotional problems as those resulting from relationships with others. One respondent described physical problems as having to do with how one feels, self-esteem, and physical illness whereas emotional problems are caused by good or bad news about family members or loved ones.

DISCUSSION

The wide variation of responses to the term *problemas emocionales* found in this study is consistent with the original effort to translate the SF-36 into Spanish, which also identified the term as problematic.[28] The results also provide a potential explanation for the irregular results that Zuniga and colleagues[33] found with the Role Emotional scale in their study with Mexicans. The results suggest that the term *problemas emocionales* may be problematic, especially for use with individuals from Latin America, if they have limited formal education or exposure to Western concepts of health. The data show that one problem for many of the respondents was the potential confusion of the adjective *emocional* (emotional) with the reflexive verb *emocionarse* (to get or become excited). One remedy might be to simply identify a different term that is commonly used by the target audience to describe this cluster of aliments. Identifying an alternative term for *problemas emocionales* was beyond the scope of this study.

The results may also suggest that the respondents do not distinguish between physical and emotional health in the same way as the creators of the instrument. In short, conditions such as depression and anxiety may not be conceived of as a related group of ailments or experienced as distinct from the physical body. Scott and colleagues[31] argue that the mind-body duality central to the SF-36v2 complicates its use with non-Western populations and may be a contributing factor to the misunderstanding of the term *problemas emocionales* reported in this study.[31]

Examination of the transcripts indicated that even when respondents understood the items as intended, it is not clear that their responses accurately reflected their health. Some participants reported their health was good but, later, would talk about physical or mental health issues they were experiencing that they felt negatively impacted their health. This was not a universal finding; however, many of the participants seemingly experienced significant stress and many reported some sort of emotional or psychological issues that are frequently minimized, denied, or repressed as part of their coping mechanisms. Often, participants would reveal these issues and mental health concerns over the course of the interview in response to the various probing questions. Overall, many immigrants seem to experience physical pains or illness (due to the type of work they do and/or the mental stress they experience) yet will minimize these until they are specifically asked about them. Denial of health problems in response to general questions about health only to divulge serious health problems upon further probing of specific health concerns is a recurring theme from previous research with foreign-born Latinx workers.[45]

Participant comprehension appeared to be negatively impacted by the cognitive burden resulting from the length and complexity of some of the Role Emotional items. Respondents were required to remember a lot of information in order to answer these items. The researchers often had to repeat parts of the items before participants fully understood what was being asked. Often, participants wanted to answer the question after the initial prompt, before the researchers read the second part of the question. Therefore, the researchers had to first explain the options for the answers and then note that participants should wait to hear the whole question.

Phrasing of the items also contributed to confusion for the respondents, especially the mixing of positive and negative categories and response options. For example, items 6 and 10 (Table) ask about physical health and emotional problems interfering with social activities. The phrase *salud fisica* seemed particularly confusing to participants as respondents seemed to understand different things upon hearing it. For some participants, it seemed to indicate positive health states (some answered that it meant good health, feeling good, feeling well) as opposed to the clearly negative connotation of *problemas emocionales*. Asking about physical “problems” rather than physical “health” would establish a consistent grammatical structure with “emotional problems” later in the item, which might contribute to a better understanding.

Study limitations

This study is limited by the small sample size, the nature of self-report data and possible influence of social desirability of provided responses, and reliance on self-ratings of fluency

in Spanish by the respondents whose first language was Indigenous. However, all these individuals completed the interview in Spanish and had responses consistent with participants who were monolingual Spanish speakers. Despite these limitations, the results highlight potential problems with the Role Emotional scale.

The findings have practical applications and serve as a cautionary tale. Few instruments are as established as the SF-36. Researchers using “off-the-shelf” tools with such an impressive record in the literature may not see a need to “retest” the items with their target population. However, even using the best and most rigorous efforts, it is almost impossible to fully account for the diversity (class, gender, dialect, race, ethnicity, etc) contained in the population of an entire country. Therefore, it is simply best practice to conduct cognitive interviews with a sample of the target population before using any data collection instrument whether the items have existed for decades or are newly written.

The ubiquity of the SF-36 has led to frequent shorthand in citation for the instrument. In the background research for this study, the authors found that many academic articles are imprecise about which version or translation was used. This imprecision also raises the possibility that the authors may have used one of the many unofficial or edited versions widely available on the Internet or, perhaps, even translated the items themselves.

Although the publisher (Optum) offers significant detail on the creation, translation, and validation of the original versions of the SF-36v2, little information can be found on the translation efforts of the country-specific versions of the SF-36v2 and the differences between them. For example, why is the word “depressed” (*deprimido/a*) used in item 5a in both the Mexico and Guatemala versions but is replaced by the word “sad” (*triste*) in item 9f only in the Mexico version? A more detailed description of the evolution of these second-generation versions seems warranted.

CONCLUSION

A significant number of Latinx immigrants to the US come directly from the rural areas of Mexico and Central America. They tend to have less formal education and exposure to Western cultural concepts than their counterparts in the urban areas. This population represents many unique challenges for public health professionals, not least of which are assessing and treating mental health conditions. The SF-36v2 and other self-report health status questionnaires are very attractive tools for researchers interested in providing a rapid assessment of the health of this community. However, the results of this study suggest that, while some improvements have been made, more are needed to ensure the instrument accurately assesses the health of Latinx immigrant communities in the US. Ultimately, the safest course, leading to the best science, requires cognitive testing of any instrument, new or already existing, when using it with a given population for the first time.

Acknowledgments

Funding for phase 1 of the study was provided by the UC College of Medicine’s “Deans Discovery Fund.” Phase 2 was a National Institute for Occupational Safety and Health (NIOSH)-led project, and no external financial resources were sought to support the research. Existing internal funds were used to support a purchase order (214-2016-M-90262) for data collection in phase 2.

REFERENCES

1. Ritsner M, Ponizovsky A. Psychological distress through immigration: the two-phase temporal pattern? Int J Soc Psychiatry. 1999;45(2):125–139. [PubMed: 10443255]
2. Rogler LH. International migrations: a framework for directing research. Am Psychol. 1994;49(8):701–708. [PubMed: 8092613]
3. Hovey JD, King CA. Suicidality among acculturating Mexican Americans: current knowledge and directions for research. Suicide Life Threat Behav. 1997;27(1):92–103. [PubMed: 9112727]
4. Vega WA, Sribney WM, Aguilar-Gaxiola S, Kolody B. 12-month prevalence of DSM-III-R psychiatric disorders among Mexican Americans: nativity, social assimilation, and age determinants. J Nerv Ment Dis. 2004;192(8):532–541. [PubMed: 15387155]
5. Grant BF, Stinson FS, Dawson DA, Chou SP, Ruan WJ, Pickering RP. Co-occurrence of 12-month alcohol and drug use disorders and personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2004;61(4):361–368. [PubMed: 15066894]
6. Grant B, Stinson FS, Hasin DS, Dawson DA, Chou SP, Anderson K. Immigration and lifetime prevalence of DSM-IV psychiatric disorders among Mexican Americans and non-Hispanic whites in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2004;61(12):1226–1233. [PubMed: 15583114]
7. Vega WA, Alderete E, Kolody B, Aguilar-Gaxiola S. Illicit drug use among Mexicans and Mexican Americans in California: the effects of gender and acculturation. Addiction. 1998;93(12):1839–1850. [PubMed: 9926572]
8. Burnam M, Hough R, Karno M, Escobar J, Telles CA. Acculturation and lifetime prevalence of psychiatric disorders among Mexican Americans in Los Angeles. J Health Soc Behav. 1987;28(1):89–102. [PubMed: 3571910]
9. Shigemoto Y, Thoen M, Robitschek C, Ashton M. Assessing Measurement Invariance of the Personal Growth Initiative Scale-II among Hispanics, African Americans, and European Americans. J Couns Psychol. 2015;62(3):537–544. [PubMed: 25915464]
10. Teresi JA, Stewart AL, Morales LS, Stahl SM. Measurement in a multi-ethnic society: overview to the special issue. Med Care. 2006;44(11)(suppl 3):S3. [PubMed: 17060831]
11. Bowden A, Fox-Rushby JA. A systematic and critical review of the process of translation and adaptation of generic health-related quality of life measures in Africa, Asia, Eastern Europe, the Middle East, and South America. Soc Sci Med. 2003;57(7):1289–1306. [PubMed: 12899911]
12. Fujishiro K, Gong F, Baron S, et al. Translating questionnaire items for a multi-lingual worker population: the iterative process of translation and cognitive interviews with English-, Spanish-, and Chinese-speaking workers. Am J Ind Med. 2010;53(2):194–203. [PubMed: 19650081]
13. Krause N The use of qualitative methods to improve quantitative measures of health-related constructs. Med Care. 2006;44(11)(suppl 3):S34–S38. [PubMed: 17060833]
14. Nápoles-Springer AM, Santoyo-Olsson J, O'Brien H, Stewart AL. Using cognitive interviews to develop surveys in diverse populations. Med Care. 2006; 44(11)(suppl 3):S21–S30. [PubMed: 17060830]
15. Ware JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): conceptual framework and item selection. Med Care. 1992;30:473–483. [PubMed: 1593914]
16. Bjorner JB, Kristensen TS. Multi-item scales for measuring global self-rated health: investigation of construct validity using structural equation models. Res Aging. 1999;21:417–439.
17. Lix LM, Wu X, Hopman W, et al. Differential item functioning in the SF-36 Physical Functioning and Mental Health sub-scales: a population-based investigation in the Canadian multicentre osteoporosis study. PLoS One. 2016;11(3):e0151519. [PubMed: 26998611]
18. McHorney CA, Fleishman JA. Assessing and understanding measurement equivalence in health outcome measures. Issues for further quantitative and qualitative inquiry. Med Care. 2006;44(11)(suppl 3):S205–S210. [PubMed: 17060829]
19. Perkins AJ, Stump TE, Monahan PO, McHorney CA. Assessment of differential item functioning for demographic comparisons in the MOS SF-36 health survey. Qual Life Res. 2006;15(3):331–348. [PubMed: 16547771]

20. Wolfe F, Hawley DJ, Goldenberg DL, Russell IJ, Buskila D, Neumann L. The assessment of functional impairment in fibromyalgia (FM): Rasch analyses of 5 functional scales and the development of the FM Health Assessment Questionnaire. J Rheumatol. 2000;27(8):1989–1999. [PubMed: 10955343]
21. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav. 1997;38(1):21–37. [PubMed: 9097506]
22. Knox SA, King MT. Validation and calibration of the SF-36 health transition question against an external criterion of clinical change in health status. Qual Life Res. 2009;18(5):637–645. [PubMed: 19330463]
23. Rumsfeld JS, MaWhinney S, McCarthy M, et al. Health-related quality of life as a predictor of mortality following coronary artery bypass graft surgery. JAMA. 1999; 281(14):1298–1303. [PubMed: 10208145]
24. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1994;31(3):247–263.
25. Ware JE, Kosinski M, Bjorner JB, Turner-Bowker DM. User's Manual for the SF 36v2 Health Survey. 2nd ed Lincoln, RI: Quality Metric Inc; 2007.
26. Ware JE, Gandek B. Overview of the SF-36 Health Survey and the International Quality of Life Assessment (IQOLA) project. J Clin Epidemiol. 1998;51(11):903–912. [PubMed: 9817107]
27. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA project approach. J Clin Epidemiol. 1998;51(11):913–923. [PubMed: 9817108]
28. Wagner AK, Gandek B, Aaronson NK, et al. Cross-cultural comparisons of the content of SF-36 translations across 10 countries: results from the IQOLA project. International Quality of Life Assessment. J Clin Epidemiol. 1998;51(11):925–932. [PubMed: 9817109]
29. Optum. Features and benefits. https://www.optum.com/sf-health-surveys/sf-36v2-health-survey.html. Accessed January 14, 2020.
30. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol. 1998;51(11):1037–1044. [PubMed: 9817121]
31. Scott KM, Sarfati D, Tobias MI, Haslett SJ. A challenge to the cross-cultural validity of the SF-36 Health Survey: factor structure in Maori, Pacific and New Zealand European ethnic groups. Soc Sci Med. 2000;51(11):1655–1664. [PubMed: 11072885]
32. Vilagut G, Ferrer M, Rajmil L, et al. El cuestionario de salud SF-36 español: una década de experiencia y nuevos desarrollos. Gac Sanit. 2005;19:135–150. [PubMed: 15860162]
33. Zuniga MA, Carrillo-Jimenez GT, Fos PJ, Gandek B, Medina-Moreno MR. [Health status evaluation with the SF-36 survey: preliminary results in Mexico]. Salud Publica Mex. 1999;41(2):110–118. [PubMed: 10343514]
34. Duran-Arenas L, Gallegos-Carrillo K, Salinas-Escudero G, Martinez-Salgado H. [Towards a Mexican normative standard for measurement of the Short Format 36 health-related quality of life instrument]. Salud Publica Mex. 2004;46(4):306–315. [PubMed: 15468571]
35. Peek MK, Ray L, Patel K, Stoebner-May D, Ottenbacher KJ. Reliability and validity of the SF 36 among older Mexican Americans. Gerontologist. 2004;44(3):418–425. [PubMed: 15197296]
36. Jordan-Marsh M, Cody M, Silverstein M, et al. Assessing a self-report health measure for non-English-speaking elders: issues in using the SF-36 Health Survey. Res Soc Work Pract. 2008;18:55–65.
37. Martinez-Hernandez LE, Segura-Méndez NH, Antonio-Ocampo A, et al. Validación del cuestionario SF-36 en pacientes adultos con asma y rinitis alérgica. Rev Med Inst Mex Seguro Soc. 2010;48:531–534. [PubMed: 21205503]
38. Brown SD, Hu U, Kirsten A, et al. The Multigroup Ethnic Identity Measure–Revised: measurement invariance across racial and ethnic groups. J Couns Psychol. 2014;61(1):154–161. [PubMed: 24188656]
39. Johnson TP. Methods and frameworks for crosscultural measurement. Med Care. 2006;44(11):S17–S20. [PubMed: 17060823]

40. Stewart AL, Thrasher AD, Goldberg J, Shea JA. A framework for understanding modifications to measures for diverse populations. J Aging Health. 2012;24(6):992–1017. [PubMed: 22495768]

41. Teresi JA, Fleishman JA. Differential item functioning and health assessment. Qual Life Res. 2007;16(suppl 1): 33–42. [PubMed: 17443420]

42. Crocker LM, Algina J. Introduction to Classical and Modern Test Theory. New York, NY: Holt Rinehart & Winston; 1986.

43. Miller K Conducting cognitive interviews to understand question-response limitations. Am J Health Behav. 2003;27(suppl 3):S264–S272. [PubMed: 14672387]

44. Willis GB. Cognitive Interviewing: A Tool for Improving Questionnaire Design. Thousand Oaks, CA: Sage Publications; 2004.

45. Eggerth DE, Ortiz B, Keller BM, Flynn MA. Work experiences of Latino building cleaners: An exploratory study. Am J Ind Med. 2019;62(7):600–608. [PubMed: 31104342]

TABLE.

Role Emotional Items SF-36v2 (Boldface Added for Emphasis)

Item No.	English	Spanish
5	During the last 4 wk, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any **emotional problems** (such as feeling depressed or anxious)?	*Durante las últimas cuatro semanas, cuanto tiempo ha tenido usted alguno de lossiguientes problemas con el trabajo u otras actividades diarias normales a causa de algún* ***problema emocional*** *(como sentirse deprimido/a o ansioso/a)?*
6	During the past 4 wk, to what extent has your physical health or **emotional problems** interfered with your normal social activities with family, friends, neighbors, or groups?	*Durante las últimas cuatro semanas en qué medida su salud física o sus* ***problemas emocionales*** *han dificultado sus actividades sociales normales con la familia, amigos, vecinos o grupos?*
10	During the past 4 wk, how much of the time has your physical health or **emotional problems** interfered with your social activities (like visiting with friends, relatives, etc)?	*Durante las últimas cuatro semanas cuanto tiempo su salud física o sus* ***problemas emocionales*** *han dificultado sus actividades sociales (como visitar amigos, parientes, etc)?*

Fam Community Health. Author manuscript; available in PMC 2022 January 01.