



Published in final edited form as:

*Health Psychol.* 2014 May ; 33(5): 490–499. doi:10.1037/hea0000055.

## New English and Spanish Social Health Measures Will Facilitate Evaluating Health Determinants

Elizabeth A. Hahn<sup>a</sup>, Darren A. DeWalt<sup>b</sup>, Rita K. Bode<sup>c</sup>, Sofia F. Garcia<sup>a</sup>, Robert F. DeVellis<sup>d</sup>, Helena Correia<sup>a</sup>, David Cella<sup>a</sup>, and for the PROMIS Cooperative Group

<sup>a</sup>Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL

<sup>b</sup>Division of General Internal Medicine and Cecil G. Sheps Center for Health Services Research, University of North Carolina School of Medicine, Chapel Hill, NC

<sup>c</sup>Department of Physical Medicine and Rehabilitation, Northwestern University Feinberg School of Medicine, Chicago, IL

<sup>d</sup>Department of Health Behavior and Health Education, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC

### Abstract

**Objective**—To develop psychometrically sound, culturally relevant and linguistically equivalent English and Spanish self-report measures of social health guided by a comprehensive conceptual model and applicable across chronic illnesses.

**Methods**—The Patient-Reported Outcomes Measurement Information System (PROMIS) Social Health Workgroup implemented a mixed methods approach to evaluate earlier results (v1.0); expand and refine domain definitions and items; translate items into Spanish; and obtain qualitative feedback. Computer-based and paper/pencil questionnaire administration was conducted with a variety of U.S. respondent samples during 2009–2012. Analyses included exploratory factor analysis (EFA), confirmatory factor analysis (CFA), two-parameter logistic item response theory (IRT) modeling, evaluation of differential item functioning (DIF), and evaluation of criterion and construct validity.

**Results**—Qualitative feedback supported the conceptualization of the Social Health domain framework (Social Function and Social Relationships sub-components). Validation testing participants ( $n=2,208$  English;  $n=644$  Spanish) were diverse in terms of gender, age, education and ethnicity/race. EFA, CFA and IRT identified seven unidimensional factors with good model fit. There was no DIF by language, and good evidence of criterion and construct validity.

**Conclusions**—PROMIS English and Spanish language instruments (v2.0), including computer-adaptive tests and fixed-length short forms, are publicly available for assessment of Social Function (Ability to Participate in Social Roles and Activities, and Satisfaction with Social Roles and Activities) and Social Relationships (Companionship; Emotional, Informational and

Instrumental Support; and Social Isolation). Measures of social health will play a key role in applications that use ecologic (or determinants of health) models that emphasize how patients' social environments influence their health.

### Keywords

patient-reported outcomes; social health; social function; social relationships; Hispanic Americans; psychometrics

---

An Institute of Medicine (IOM) report recommends development of a standardized, core set of indicators focused on priority health outcomes (Institute of Medicine. Committee on Public Health Strategies to Improve Health, 2011). The IOM also recommends changes in the processes, tools and approaches for gathering information on health outcomes. Several groups have been working to identify and test concepts of health and function that are meaningful across countries and cultures (Taskforce on Health Status, 2005). The Patient-Reported Outcomes Measurement Information System (PROMIS; [www.nihpromis.org](http://www.nihpromis.org)) adopted the World Health Organization framework to define three components of health: physical, mental and social (World Health Organization, 1946) (see <http://www.nihpromis.org/measures/domainframework>). This is congruent with the biopsychosocial approach adopted by many health psychologists (Engel, 1977, 1980).

Measures of social health will play a key role in applications that use ecologic (or determinants of health) models that emphasize how patients' environments influence their health (Institute of Medicine, 2003; Institute of Medicine. Committee on Public Health Strategies to Improve Health, 2011; Whitehead, 1995). This is of particular significance given that social health has historically been a relatively neglected domain due to the lack of measures for clinical populations, as well as disagreement about how best to define and measure it (Hahn, Cella, Bode, & Hanrahan, 2010). Social health measures that can be used across chronic illness populations are essential given findings that individuals who receive appropriate social support and are integrated within their communities experience better health outcomes than those who experience social isolation (Broadhead et al., 1983; Bruhn & Philips, 1984; McDowell, 2006; Mitchell, Billings, & Moos, 1982).

At the broadest level, social health includes health outcomes as well as social processes that play an important role in influencing other health outcomes, e.g., mediating or moderating the effects of stress on physical health and illness. In some contexts (e.g., family or group therapy), processes such as social support can be outcomes of interventions. The PROMIS domain framework for Social Health (v2.0) includes two primary sub-components: Social Function and Social Relationships (McDowell, 2006) (see supplemental Figure 1). As described in detail elsewhere, PROMIS initially developed two sets of Social Function items (*Ability to Participate* and *Satisfaction with Participation*, v1.0) and tested them in a general population English-speaking sample (Hahn et al., 2010). Results showed that some item refinement was necessary, and that clinical samples should be included in future testing to evaluate our conceptual models and produce measures for use across chronic illnesses. Additionally, development of Social Relationships items was needed as well as Spanish language versions in order to address the rising need for assessment tools appropriate for

Latinos/Hispanics, who constitute the fastest growing group among minority populations that are often underserved in health care settings and excluded from research studies (Garber & Arnold, 2006; Yancey, Ortega, & Kumanyika, 2006). Establishing cultural equivalence across populations is particularly important for social health, which inherently includes cultural norms against which social roles and support are assessed (McDowell, 2006).

The aim of the work we present here was to develop psychometrically sound, culturally relevant and linguistically equivalent English and Spanish language self-report measures of adult social health guided by a comprehensive conceptual model and applicable across chronic illnesses. This manuscript describes the work of the PROMIS Social Health Workgroup to refine and expand the PROMIS Social Health measures (v2.0), test the items with large samples of English- and Spanish-speaking adults, and derive item response theory (IRT)-based item banks (sets of calibrated items). These calibrated item banks are comprised of numerous items that allow for computerized adaptive testing (CAT) and development of multiple short forms of varying length that provide accurate measurement with low response burden (Cella, Gershon, Lai, & Choi, 2007).

## Method

The main goal of PROMIS was to develop a set of publicly available, efficient and flexible measurements of patient-reported outcomes (Cella et al., 2010). State-of-the-science qualitative and quantitative methods were implemented to develop unidimensional item sets that fit a two-parameter IRT model and do not exhibit measurement bias (differential item functioning; DIF) across gender, age, education, language and administration method (computer vs. paper).

## Domain Definitions and Item Development

The initial work of the PROMIS Social Health Workgroup is described in detail elsewhere (Castel et al., 2008; DeWalt, Rothrock, Yount, Stone, & PROMIS Cooperative Group, 2007; Hahn, Devellis, et al., 2010). Briefly, this involved a qualitative item review process that included identification of existing items, development of new items, item revision, readability levels, focus group exploration of domain coverage and cognitive interviews. Building on a previous study of adults with cancer (Hahn, Cella, et al., 2010), 56 items were developed in each of two broad domains of Social Function (*Ability to Participate* and *Satisfaction with Participation*). In initial PROMIS testing with a general population sample, 956 respondents answered all 112 items (Hahn, Devellis, et al., 2010). Analyses included exploratory and confirmatory factor analysis, and IRT modeling. Two preliminary item banks (v1.0) were calibrated (*Satisfaction with Participation in Social Roles, and Satisfaction with Participation in Discretionary Activities*).

Based on these results, the workgroup conducted another systematic review of social health definitions, content and item wording. Some Social Function items were revised, some were removed, and some new items were written to fill important gaps in content. Domain names for Social Function (v2.0) were slightly revised: *Ability to Participate in Social Roles and Activities*, and *Satisfaction with Social Roles and Activities* (see supplemental Figure 1). Work began on Social Relationships, focusing on positive aspects of social support and one

negative component of social isolation. The workgroup implemented PROMIS procedures for item development, including reviews by content experts, focus groups and cognitive interviews with patients, and thematic analysis of qualitative data (DeWalt, et al., 2007).

### Qualitative Patient Feedback

Focus groups and cognitive interviews were conducted with adults receiving care for cancer in the Chicago metropolitan area (at a private hospital and a safety net county hospital). Individuals were purposefully sampled so that approximately half were racial or ethnic minorities, and all received a \$20 incentive. The focus group participants discussed physical and psychosocial aspects of their cancer experience. Cognitive interviews were then conducted to review definitions of Social Health and to gather patient input on overall item content. Another round of cognitive interviews took place with another sample of adults with cancer, to determine how clearly and accurately respondents understood each item. Participants first completed a subset of the Social Health items (v2.0). This ensured that no participant was burdened with more than 40 of the approximately 150 items and that each item was reviewed by at least five participants. Interviewers inquired about what participants considered in choosing each response and what additional comments they had about specific items. Lastly, interviewers asked participants general questions about the Social Health items, including face validity and timeframe, and any perceived content gaps or difficulties with comprehension. Two additional cognitive interviews were conducted to ensure the comprehensibility of 25 modified items. All interview responses were independently reviewed by three raters.

### Spanish Translations

The items were translated into Spanish with a multi-step forward-backward process (Eremenco, Cella, & Arnold, 2005; Wild et al., 2005). Each translated item was administered to five native Spanish speakers, following an interview protocol similar to that in English. Respondents were also asked to explain the meaning of the items in their own words and to consider alternative wording. All comments were analyzed to determine if and where wording changes were needed.

### Participant Recruitment and Assessment Procedures

**English item testing**—PROMIS Social Function and Social Relationships items (v2.0) were administered to several large, diverse, convenience samples manifesting varied clinical problems associated with functional limitation. Testing took place at three PROMIS sites with existing study cohorts. The three sites were chosen because of their clinical and demographic diversity (UNC sample: arthritis cohort, age 45 years; Stanford sample: arthritis cohort, age 18 years; Duke sample: cardiovascular disease cohort, age 18 years). Participants provided signed informed consent in accordance with IRB regulations, and received an incentive in accordance with the existing study cohort initiatives (UNC: \$20 gift card; Stanford: booklet of 20 postage stamps; Duke: \$80). Items were also administered to an online survey panel of general population respondents (age 18 years; [www.toluna-group.com](http://www.toluna-group.com)). Purposeful sampling required at least half of the online panel to be younger than age 60 and at least 20% to be age 60 or older; a small incentive (\$2 to \$5) was

provided. Computer-based testing was used for all respondents except for the majority of those from Stanford who completed paper-and-pencil questionnaires. To minimize respondent burden and to ensure adequate sample sizes for each domain, items were divided into three sets, with some item overlap. Respondents also completed a small number of other instruments: sociodemographics, clinical history, and 15 items from “legacy” (i.e., widely used and accepted) instruments to evaluate criterion validity (SF-36, version 2, acute timeframe (Ware, Kosinski, & Dewey, 2000) and FACT-GP, version 4 (Brucker, Yost, Cashy, Webster, & Cella, 2005)). Some respondents also completed the 10-item PROMIS Global instrument (Hays, Bjorner, Revicki, Spritzer, & Cella, 2009).

**Spanish item testing**—A randomly selected group of adult (age 18 years) Spanish-speaking members of an online survey panel were notified of a new survey opportunity ([www.toluna-group.com](http://www.toluna-group.com)). Those who indicated that they were age 18 or older and of Hispanic ethnicity were then presented with the four-item Short Acculturation Scale for Hispanics (SASH) (Marin, 1987). Only those with a SASH score < 3.0 (speak only or mostly Spanish) were considered eligible for this study. Respondents completed the full set of PROMIS Social Function items (v2.0) and a small number of other instruments (sociodemographics, clinical history, PROMIS Global (Hays, et al., 2009), PROMIS Sleep Disturbance and PROMIS Sleep-Related Impairment). Due to respondent burden, it was not possible to add the Social Relationships items. A small incentive (\$2 to \$5) was provided.

### Psychometric and Statistical Analyses

Analyses followed the PROMIS guidelines (Reeve et al., 2007), separately by language. Preliminary analyses were conducted to identify unused or sparsely used categories, and to examine whether the average measures in response categories increased monotonically. Data from the English online panel and clinical samples were pooled for analysis.

**Assessment of dimensionality: Social Function**—Separate sets of analyses were conducted for each Social Function domain (*Ability to Participate in Social Roles and Activities*, and *Satisfaction with Social Roles and Activities*). The English sample was randomly split in half for use in either exploratory factor analysis (EFA) or categorical confirmatory factor analysis (CFA); only CFA was performed for Spanish. For EFA, polychoric correlations were used in Mplus and analyzed using an unweighted least squares estimation procedure with quartimin rotation (Muthen, B. O., du Toit, & Spisic, 1987; Muthen, L. K. & Muthen, 2006; Reeve, et al., 2007). Factors were identified by eigenvalues greater than 1.0. Items loading 0.40 or above on a factor were examined to describe the factor. For CFA, polychoric correlations were entered into Mplus and analyzed using a weighted least squares estimation procedure that is robust to non-normality. A value greater than 0.95 on the comparative-fit index (CFI) was considered evidence of good model fit; a value greater than 0.90 was considered acceptable (Hu & Bentler, 1999). The results of one-factor models were examined; acceptable model fit provided some support for unidimensionality. Local dependence was also examined, defined as a residual correlation between item pairs of 0.20 or greater.

**Assessment of dimensionality: Social Relationships (English only)**—EFA was first performed to determine the factor structure of the Social Support items. Items loading 0.40 or above on a factor were examined to describe the factor. After deleting items that loaded on more than one factor, the sample was randomly split in half for use in either EFA or CFA, using the methods described above. Separate sets of analyses were conducted for each identified Social Support sub-domain. Another set of analyses was conducted for the Social Isolation items, randomly splitting the sample in half for EFA or CFA.

**Estimation of IRT parameters**—MULTILOG software (Thissen, 1991) was used to implement the graded response model (Samejima, 1969) to estimate item parameters and evaluate model fit for each identified domain or sub-domain. In this two-parameter logistic IRT model, item responses are used to estimate the “measure” (the person’s score on the latent trait). The two parameters are item location on the latent trait and item slope, which indicates how well the item discriminates (distinguishes) between person differences across the latent trait (van der Linden & Hambleton, 1997). Item characteristic curves were examined for the distribution of responses across categories, item thresholds were examined for the range being measured by the items, and slopes  $> 2.0$  were used as an indicator that the item has acceptable discrimination between individuals (see example in Rose et al., 2008 (Rose, Bjorner, Becker, Fries, & Ware, 2008)). Model fit was assessed with likelihood-based chi-square statistics ( $S-X^2, p > 0.01$ ) (Orlando & Thissen, 2000). The standard error of measurement (SEM) - an indicator of precision - was estimated for each person. An advantage of IRT is that precision is estimated at different points across the latent trait. An SEM of 3.0 or less is generally considered a good indicator of precision because it is roughly equivalent to a reliability of 0.9 and is often used as a cut point for stopping rules in CAT (Babcock & Weiss, 2012).

**Differential item functioning (DIF)**—Three group comparisons were evaluated: by gender, by age (<65 versus  $\geq 65$  in English; <45 versus  $\geq 45$  in Spanish) and by education (high school/GED or less versus higher education). DIF by administration method (computer versus paper) was also examined in the English samples. Uniform and non-uniform DIF were evaluated using the LORDIF R package (Choi, Gibbons, & Crane, 2011). Uniform DIF is in the same direction across the entire latent trait, e.g., if women were more likely to report having someone to take them to the doctor compared with men at the same level of instrumental support. Non-uniform DIF means that an item favors one group at certain trait levels, and other groups at other levels, e.g., if women were more likely to report having someone to take them to the doctor compared with men at low levels of instrumental support, but less likely to report this at high levels of instrumental support. Non-uniform DIF can be viewed as significant group-by-trait interaction. Hierarchically nested IRT models were compared; specifically, one model that fully constrained parameters to be equal between two groups was compared to other models that allowed parameters to be freely estimated. Differences between model 1 (scores) and model 2 (scores, group) were used to identify uniform DIF, and between model 2 and model 3 (scores, group, interaction between scores and group) to identify non-uniform DIF, using the criterion of a change of 0.20 or greater in the R-square as an indicator of a meaningful DIF effect size. Similar analyses were conducted to evaluate DIF by language.

**Scoring and evaluation of validity**—PROMIS T-scores were used (mean=50; standard deviation=10), and descriptive statistics were calculated for gender, age and education subgroups. Scores for the legacy instruments (SF-36 (Ware, et al., 2000) and FACT-GP (Brucker, et al., 2005)) were calculated using the developers' guidelines.

Criterion validity (Scientific Advisory Committee of the Medical Outcomes Trust, 2002) for the English data was evaluated by calculating Pearson correlation coefficients between PROMIS T-scores, three SF-36 subscales (Social Functioning, Role-Emotional and Role-Physical (Ware, et al., 2000)), and the FACT-GP Functional Well-being subscale (Brucker, et al., 2005). Correlations with the SF-36 and FACT-GP were hypothesized to be high for PROMIS Social Function domains since all of these measures focus on perceived function and abilities. Correlations for PROMIS Social Relationships were expected to be more moderate, with Social Isolation being negatively associated, and Social Support being positively associated, with the criterion measures. Construct validity (Scientific Advisory Committee of the Medical Outcomes Trust, 2002) was evaluated by comparing mean scores between respondents with and without comorbidities, and between English online panel respondents and those from the clinical samples, using independent group *t*-tests. Effect sizes were calculated as the mean difference between groups divided by the overall standard deviation, and interpreted in terms of a minimally important difference in patient-reported outcome measures (0.33) (Yost & Eton, 2005). Respondents without comorbidities and those from the online panel were hypothesized to have higher PROMIS Social Function scores than the comparison groups.

**Development of short forms**—Fixed-item short forms were created to provide an alternative to CAT when internet-based data collection may not be feasible. The criteria for item inclusion were content representativeness (inclusion of items from each context), maximized range of difficulty (inclusion of items across the calibration range), and acceptable discrimination levels (inclusion of items that distinguish between people across the latent trait).

## Results

### Domain Definitions, Item Development and Qualitative Patient Feedback

Four focus groups ( $n=21$  total participants) and domain-level cognitive interviews ( $n=25$ ) were conducted to inform item modifications and the writing of new items to fill content gaps. The focus groups highlighted the need to include a negative aspect of Social Relationships (Social Isolation) while corroborating the need to include different kinds of Social Support (e.g., instrumental and emotional). Cognitive interview responses supported a framework in which Social Function is comprised of *Ability to Participate in Social Roles and Activities* and *Satisfaction with Social Roles and Activities* domains, with both including the contexts of family, friends, work and leisure.

Semi-structured cognitive interviews were conducted on the revised pool of items with 21 English-speaking adults with a variety of cancer diagnoses (mean age, 56 years; 57% Female; 52% White and 33% African-American). Patient feedback, along with expert review by the Social Health Workgroup, resulted in the deletion of a few items, modification

of a few items, and the decision to drop the use of a timeframe (“past seven days” or “lately”) for the items. When asked what relevant questions were not included, none of the participants mentioned any missing content in Social Health. Instead, they mentioned content covered in other PROMIS domains (e.g., Emotional Health) or currently outside of the PROMIS framework (e.g., patient-provider communication).

Spanish translations were modified for a small number of items as a result of the cognitive interviews with native Spanish speakers. Several items were considered too long and confusing, and a few participants reported having to read some items more than once to understand them. No items were considered irrelevant or offensive and none of the participants suggested the need to add new items. The translation of “leisure activities” was revised from “actividades de ocio” to “actividades de tiempo libre” since several participants did not understand the word “ocio.” The Spanish word used for “holidays” (“Do you have someone with whom you can celebrate holidays?”), “festividades”, proved to be ambiguous, so it was revised to “dias festivos.” The register (formality) of the word used for “close” (“I feel that I am no longer close to anyone”), had to be lowered by adding an alternative word, while still maintaining a universal understanding of the item.

The final definition of Social Health is that it encompasses participation in activities with others, carrying out one’s usual roles and responsibilities, and relationships and connections with important others. These include the ability to relate to individuals, groups, communities and society as a whole. Supplemental Table A summarizes definitions, and example wording and rating scales. Computer-based testing was subsequently conducted on 79 Social Function and 69 Social Relationships items.

### Respondent Characteristics

Participants in the item testing studies were fairly diverse in terms of gender, age and education (Table 1). The English language samples were primarily non-Hispanic Whites; by design, the Spanish language sample was Hispanic/Latino.

**Psychometric Analyses: English**—For each domain or sub-domain, all of the response categories were used; thus, no category collapsing was needed. Missing data ranged from zero to 5% across items. The average measures increased monotonically across categories.

**Dimensionality analyses**—The EFA for the two Social Function domains (abbreviated here as *Ability* and *Satisfaction*) identified a large first factor (see Table 2). Each CFA had good model fit and no local dependence (all residual correlations <0.20).

The initial EFA for *Social Support* included a total of 54 items (not shown). EFA identified five factors with eigenvalues greater than 1.0. After removing nine items that cross-loaded on multiple factors, the EFA was repeated. Four factors were identified (Companionship, Emotional Support, Informational Support, Instrumental Support) and all items loaded on their hypothesized factors (item loadings, 0.53 to 0.91; see supplemental Table B). For each factor, the sample was randomly split in half for use in either EFA or CFA.



The CFA for *Companionship* had good model fit (CFI, 0.99) and no local dependence. Results of the initial IRT analysis suggested that two items should be removed due to model misfit or gender DIF (these items asked about celebrating holidays and special occasions). EFA and CFA were then repeated with the reduced item set. EFA identified one factor; CFA had good model fit and no local dependence (see Table 2). For *Emotional Support*, *Informational Support*, *Instrumental Support* and *Social Isolation*, one factor was identified by EFA. Each CFA had good model fit and no local dependence.

**IRT calibration analyses**—All *Ability* and *Satisfaction* items fit the IRT models, with slopes greater than 2.0 for all but one item and a small mean SEM (see Table 2). Over 90% of respondents had an SEM  $\leq 3.0$ , indicating that the measure provides excellent precision for a large part of the sample. All *Companionship*, *Emotional Support*, *Informational Support*, *Instrumental Support* and *Social Isolation* items fit the IRT models, with slopes greater than 2.0 for most items. The proportion of respondents with an SEM  $\leq 3.0$  ranged from 59.0% to 84.3%.

**Differential item functioning (DIF)**—There was no evidence of DIF by gender, age or education for *Ability*, *Companionship*, *Emotional Support*, *Informational Support* and *Instrumental Support*. For *Satisfaction* and *Social Isolation*, five and two items, respectively, exhibited some DIF by administration method only (data not shown).

### Psychometric Analyses: Spanish

For each Social Function domain, there were no missing data and all of the response categories were used; thus, no category collapsing was needed. The average measures increased monotonically across categories for *Ability*. Two items had category inversions in *Satisfaction*; specifically, the average measure was lower for respondents selecting the highest category (*Muchísimo*) compared to the average measure for the adjacent category (*Mucho*).

**Dimensionality analyses**—The CFA for *Ability* had good model fit and no local dependence (see Table 2). The CFA for *Satisfaction* had acceptable model fit; one item pair exhibited local dependence.

**IRT calibration analyses**—All *Ability* and *Satisfaction* items fit the IRT models with slopes greater than 2.0 for all but one item and a small mean SEM. Over 90% of respondents had an SEM  $\leq 3.0$ .

**Differential item functioning (DIF)**—There was no evidence of DIF by gender, age or education.

### Psychometric Analyses: English versus Spanish

There was no evidence of DIF by language. The English IRT calibrations were therefore used to create a measure (score) for each person who completed the English or Spanish language items; descriptive statistics are summarized in Table 3.

## Criterion and Construct Validity

The English data demonstrated good evidence of criterion validity. Pearson correlation coefficients between PROMIS Ability or Satisfaction T-scores and the SF-36 and FACT-GP legacy subscales ranged from 0.52 to 0.76 (see supplemental Table C). As expected, correlations for the PROMIS Social Support sub-domains were smaller with the SF-36 (0.11 to 0.30) and the FACT-GP (0.34 to 0.48). Correlations were negative and moderate between PROMIS Social Isolation and the SF-36 or FACT-GP (−0.30 to −0.57).

There was also good evidence of construct validity in English and Spanish. As hypothesized, respondents without comorbidities had higher Ability and Satisfaction scores compared to those with comorbidities (English effect sizes, 0.94 and 0.85, respectively; Spanish effect sizes, 0.79 and 0.50, respectively; Table 4). Effect sizes were small for Social Relationships (English only). Interestingly, Social Support scores for the online panel were lower than the clinical samples (effect sizes, −0.31 to −0.47). Social Isolation scores were higher (poorer) for the online panel (effect size, 0.45).

## Item Banks and Short Forms

Seven Social Health item banks are available in English and Spanish through the PROMIS internet-based data collection and management system ([www.assessmentcenter.net](http://www.assessmentcenter.net)). The item banks can be administered as a CAT, or users can select a fixed-item short form (4, 6, or 8 items) if internet administration is not feasible. Custom short forms can also be created. CAT will provide the most precise estimate (score); short forms are based on the same IRT calibrations as a CAT, but will be slightly less precise because they are not targeted to an individual respondent.

## Discussion

A primary objective of the PROMIS physical, mental and social health workgroups was to develop a comprehensive approach to assessing self-reported health that would produce English and Spanish language instruments that are psychometrically sound and culturally appropriate. The Social Health Workgroup efforts were reported here, resulting in 136 items measuring seven aspects of social health, each producing a unidimensional item bank that enables users to select subsets of items or CATs from each bank to create efficient measures. The study presented in this article significantly expanded our previous work (Hahn, Devellis, et al., 2010) by developing item banks of Social Relationships in addition to Social Function, comprehensively translating all items, and testing them in English- and Spanish-speaking clinical populations.

Our initial testing of English language Social Function items (v1.0) was conducted with a general population sample. Results revealed unexpected and somewhat disappointing features of the *Ability to Participate* item pool (e.g., poor model fit, difficulties arising from positively and negatively worded items within a single subdomain), and two preliminary item banks were created for *Satisfaction with Participation* (Hahn, Devellis, et al., 2010). After a process of careful item review, the revised Social Function items (v2.0) demonstrated substantial improvements when tested on multiple, diverse samples that varied

with respect to health status. Both the English and Spanish versions revealed highly acceptable psychometric properties providing evidence of reliability and validity. Likewise, the newly developed English language versions for Social Relationships showed good model fit and discrimination parameters, and good precision for estimating respondent measures. In addition, there was no evidence of measurement bias (DIF) by gender, age, education or language.

Creating assessment tools that can be used readily in a variety of populations is central to the PROMIS mission. The decision to create Spanish language versions for self-reported Social Health was driven by the growing number of Hispanics/Latinos (50.5 million) in the U.S., the majority (76%) of whom are Spanish-dominant or bilingual (Pew Hispanic Center, 2011), as well as by the need to have translated measures for use in global research. By using a comprehensive translation methodology and in finding no DIF by language, the Spanish language versions are culturally responsive. Having culturally relevant, linguistically equivalent and psychometrically sound patient-reported measures in languages other than English helps to overcome some common barriers to including underrepresented groups in research and to conducting cross-cultural research. In particular, accurate Spanish language measures will be useful in helping to unravel whether the Hispanic/Latino epidemiologic paradox is substantiated, and whether social support contributes to the health outcomes among Hispanics/Latinos that are often found to be comparable to or better than that of other racial/ethnic groups in the U.S despite their low levels of average income and education (Eschbach, Ostir, Patel, Markides, & Goodwin, 2004; Franzini, Ribble, & Keddie, 2001). In general, appropriately translated PROMIS measures allow for inclusion of substantial demographic subgroups in research studies, and better examination of cultural differences in patient-reported outcomes and health disparities among vulnerable populations.

Similarly, the PROMIS method allows for comparability across populations with different illnesses or levels of health because it measures common, generic experiences that apply to people in a variety of contexts or with a variety of diseases (Cella, et al., 2010). Social Health includes item content with universal applicability along with comprehensive coverage of social relationships and function. The World Health Organization's International Classification of Functioning, Disability and Health (ICF) uses the term "participation" to describe social health and functioning, and merges activities and participation into a single taxonomy with multiple domains, including learning, communication, mobility, self-care, interpersonal relationships and social life (World Health Organization, 2001). A recent review identified eight contemporary participation measures with linkages to the ICF, yet these instruments are not equivalent because, even when linked to a single ICF code, items represent different aspects of a dimension (Magasi & Post, 2010). Instruments are also diverse in their theoretical approach, format, psychometric properties and targeted use (Anatchkova & Bjorner, 2010).

Other researchers applied the PROMIS IRT-based item banking approach to develop a generic role functioning item bank in English (Anatchkova, Ware, & Bjorner, 2011). A key distinction between PROMIS and this role functioning item bank is that the latter includes health attribution in each item, e.g., *because of health [I] have to stop work*. These

investigators acknowledge that this decision narrowed the definition of the construct and also made the items longer and more cumbersome.

This study is not without limitations. Although convenience sampling was used, purposeful enrollment of diverse samples helped to better understand the item bank properties. There is a possibility of differential response to items based on site or diagnosis. Future research could be designed to evaluate measurement equivalence across these and other characteristics. Only cross-sectional testing has been performed to date, so there are no data on change or responsiveness. Items for Social Relationships have been translated into Spanish but not yet tested or evaluated for criterion and construct validity. The psychometric analysis methods implemented by PROMIS are working, practical tools, and so might be limited in representing the multidimensional space of social health optimally. This should be a topic for further research.

Nonetheless, we believe that the fruits of the PROMIS efforts are significant, namely, seven English and Spanish language item banks (v2.0) that achieve broad representation of Social Health and are available for public use ([www.assessmentcenter.net](http://www.assessmentcenter.net)), specifically: Social Function (*Ability to Participate in Social Roles and Activities*, and *Satisfaction with Social Roles and Activities*) and Social Relationships (*Companionship; Emotional, Informational and Instrumental Support; Social Isolation*) (see supplemental Figure 1). These item banks have undergone extensive and rigorous qualitative and psychometric assessment, and demonstrate strong properties. Collectively, the availability of these item banks represents a substantial resource that is now available to investigators and clinicians who wish to accurately assess social function and social relationships using CATs and short form measures. The use of common indicators of Social Health will facilitate measurement consistency and comparison across studies and populations, and should enhance understanding of how these variables relate to other aspects of health.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The Patient-Reported Outcomes Measurement Information System (PROMIS) is a National Institutes of Health (NIH) Roadmap initiative to develop a computerized system measuring patient-reported outcomes in respondents with a wide range of chronic diseases and demographic characteristics. PROMIS was funded by cooperative agreements to a Statistical Coordinating Center (Evanston Northwestern Healthcare, PI: David Cella, PhD, U01AR52177) and six Primary Research Sites (Duke University, PI: Kevin Weinfurt, PhD, U01AR52186; University of North Carolina, PI: Darren DeWalt, MD, MPH, U01AR52181; University of Pittsburgh; PI: Paul A. Pilkonis, PhD, U01AR52155; Stanford University, PI: James Fries, MD, U01AR52158; Stony Brook University, PI: Arthur Stone, PhD, U01AR52170; and University of Washington, PI: Dagmar Amtmann, PhD, U01AR52171). NIH Science Officers on this project have included Deborah Ader, PhD, Susan Czajkowski, PhD, Lawrence Fine, MD, DrPH, Laura Lee Johnson, PhD, Louis Quatrano, PhD, Bryce Reeve, PhD, William Riley, PhD, Susana Serrate-Sztejn, MD, and James Witter, MD, PhD. Each co-author of this paper was an investigator on one of these PROMIS grants. The authors thank Karon Cook, PhD, Ron Hays, PhD, and Leo Morales, MD, PhD, for consultation on the psychometric analyses.

## Reference List

- Anatchkova M, Bjorner J. Health and role functioning: the use of focus groups in the development of an item bank. *Quality of Life Research*. 2010; 19(1):111–123. [PubMed: 20047086]
- Anatchkova M, Ware J, Bjorner J. Assessing the factor structure of a role functioning item bank. *Quality of Life Research*. 2011; 20(5):745–758. [PubMed: 21153710]
- Babcock B, Weiss DJ. Termination Criteria in Computerized Adaptive Tests: Do Variable-Length CATs Provide Efficient and Effective Measurement? *Journal of Computerized Adaptive Testing*. 2012; 1(1):1–18.
- Broadhead WE, Kaplan BH, James SA, Wagner EH, Schoenbach VJ, Grimson R, Gehlbach SH. The epidemiologic evidence for a relationship between social support and health. *American Journal of Epidemiology*. 1983; 117(5):521–537. [PubMed: 6342368]
- Brucker PS, Yost K, Cashy J, Webster K, Cella D. General population and cancer patient norms for the Functional Assessment of Cancer Therapy-General (FACT-G). *Evaluation & the Health Professions*. 2005; 28(2):192–211. [PubMed: 15851773]
- Bruhn JG, Philips BU. Measuring social support: a synthesis of current approaches. *Journal of Behavioral Medicine*. 1984; 7(2):151–169. [PubMed: 6748066]
- Castel LD, Williams KA, Bosworth HB, Eisen SV, Hahn EA, Irwin DE, DeVellis RF. Content validity in the PROMIS social-health domain: A qualitative analysis of focus-group data. *Quality of Life Research*. 2008; 17(5):737–749. [PubMed: 18478368]
- Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Quality of Life Research*. 2007; 16(Suppl 1): 133–141. [PubMed: 17401637]
- Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S. PROMIS Cooperative Group. The Patient Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *Journal of Clinical Epidemiology*. 2010; 63(11):1179–1194. doi: S0895-4356(10)00173-3 [pii]10.1016/j.jclinepi.2010.04.011. [PubMed: 20685078]
- Choi SW, Gibbons LE, Crane PK. Iordif: An R Package for Detecting Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations. *Journal of Statistical Software*. 2011; 39(8):1–30. doi: <http://www.jstatsoft.org/v39/i08/>. [PubMed: 21572908]
- DeWalt DA, Rothrock N, Yount S, Stone AA. PROMIS Cooperative Group. Evaluation of Item Candidates: The PROMIS Qualitative Item Review. *Medical Care*. 2007; 45(5 Suppl 1):S12–S21. [PubMed: 17443114]
- Engel GL. The need for a new medical model: a challenge for biomedicine. *Science*. 1977; 196(4286): 129–136. [PubMed: 847460]
- Engel GL. The clinical application of the biopsychosocial model. *American Journal of Psychiatry*. 1980; 137(5):534–544.
- Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation and the Health Professions*. 2005; 28(2):212–232. [PubMed: 15851774]
- Eschbach K, Ostir GV, Patel KV, Markides KS, Goodwin JS. Neighborhood context and mortality among older Mexican Americans: is there a barrio advantage? *American Journal of Public Health*. 2004; 94(10):1807–1812. [PubMed: 15451754]
- Franzini L, Ribble JC, Keddie AM. Understanding the Hispanic paradox. *Ethnicity and Disease*. 2001; 11(3):496–518. [PubMed: 11572416]
- Garber M, Arnold RM. Promoting the participation of minorities in research. *American Journal of Bioethics*. 2006; 6(3):W14–W20. [PubMed: 16754431]
- Hahn EA, Cella D, Bode RK, Hanrahan RT. Measuring Social Well-being in People with Chronic Illness. *Social Indicators Research*. 2010; 96(3):381–401.
- Hahn EA, DeVellis RF, Bode RK, Garcia SF, Castel LD, Eisen SV, Cella D. Measuring social health in the patient-reported outcomes measurement information system (PROMIS): item bank development and testing. *Quality of Life Research*. 2010; 19(7):1035–1044. [PubMed: 20419503]

- Hays RD, Bjorner J, Revicki DA, Spritzer K, Cella D. Development of physical and mental health summary scores from the Patient Reported Outcomes Measurement Information System (PROMIS) global items. *Quality of Life Research*. 2009; 18(7):873–880. [PubMed: 19543809]
- Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6(1):1–55.
- Institute of Medicine. *The future of the public's health in the 21st century*. Washington, D.C.: National Academies Press; 2003.
- Institute of Medicine. *Committee on Public Health Strategies to Improve Health. For the public's health: the role of measurement in action and accountability*. Washington, D.C.: National Academies Press; 2011.
- Magasi S, Post MW. A Comparative Review of Contemporary Participation Measures' Psychometric Properties and Content Coverage. *Archives of Physical Medicine and Rehabilitation*. 2010; 91 Supplement 1(9):S17–S28. [PubMed: 20801275]
- Marin G. Development of a Short Acculturation Scale for Hispanics. *Hispanic Journal of Behavioral Sciences*. 1987; 9(2):183–205.
- McDowell, I. *Measuring health: a guide to rating scales and questionnaires*. 3rd ed.. Oxford; New York: Oxford University Press; 2006.
- Mitchell R, Billings A, Moos R. Social support and well-being: Implications for prevention programs. *Journal of Primary Prevention*. 1982; 3(2):77–98. [PubMed: 24301450]
- Muthen BO, du Toit SHC, Spisic D. Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. 1987 from [http://www.gseis.ucla.edu/faculty/muthen/articles/Article\\_075.pdf](http://www.gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf).
- Muthen, LK.; Muthen, BO. *Mplus User's Guide*. Los Angeles, CA: Muthen & Muthen; 2006.
- Orlando M, Thissen D. Likelihood-based item-fit indices for dichotomous item response theory models. *Applied Psychological Measurement*. 2000; 24:50–64.
- Pew Hispanic Center. *National Survey of Latinos*. 2011 from <http://www.pewhispanic.org/>.
- Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Cella D. Psychometric Evaluation and Calibration of Health-Related Quality of Life Item Banks: Plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*. 2007; 45(5 Suppl 1):S22–S31. [PubMed: 17443115]
- Rose M, Bjorner JB, Becker J, Fries JF, Ware JE. Evaluation of a preliminary physical function item bank supported the expected advantages of the Patient-Reported Outcomes Measurement Information System (PROMIS). *Journal of Clinical Epidemiology*. 2008; 61(1):17–33. [PubMed: 18083459]
- Samejima F. Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, No. 17. 1969
- Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality of life instruments: attributes and review criteria. *Quality of Life Research*. 2002; (11):193–205. [PubMed: 12074258]
- Taskforce on Health Status. *Criteria for and selection of domains for the measurement of health status*; Paper presented at the Conference of European Statisticians; Budapest, Hungary. 2005.
- Thissen, D. *MULTILOG user's guide*. Multiple, categorical item analysis and test scoring using item response theory. Lincolnwood, IL: Scientific Software International, Inc.; 1991.
- van der Linden, WJ.; Hambleton, RK. *Handbook of Modern Item Response Theory*. New York: Springer-Verlag; 1997.
- Ware, JE.; Kosinski, M.; Dewey, JE. *How to score version two of the SF-36 health survey*. Lincoln, R.I.: QualityMetric; 2000.
- Whitehead, M. Tackling inequalities: A review of policy initiatives. In: Benzeval, M.; Judge, K.; Whitehead, M., editors. *Tackling inequalities in health: an agenda for action*. London: King's Fund; 1995. p. 22-52.
- Wild D, Grove A, Martin M, Eremenco S, Ford S, Verjee-Lorenz A, Erickson P. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient reported outcomes(PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health*. 2005; 8(2):94–104. [PubMed: 15804318]

- World Health Organization. Geneva: World Health Organization; 1946. Constitution of the World Health Organization.
- World Health Organization. Geneva, Switzerland: World Health Organization; 2001. International Classification of Functioning, Disability and Health (ICF).
- Yancey AK, Ortega AN, Kumanyika SK. Effective recruitment and retention of minority research participants. *Annual Review of Public Health*. 2006; 27:1–28.
- Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: The FACIT experience. *Evaluation and the Health Professions*. 2005; 28(2): 172–191. [PubMed: 15851772]

**Table 1**

## Sociodemographic and Clinical Characteristics of Study Participants

	English		Spanish
	General population ( <i>n</i> =1,008)	Clinical samples ( <i>n</i> =1,200)	General population ( <i>n</i> =644)
Site			
Online panel	1008 (100%)	---	644 (100%)
UNC	---	622 (52%)	---
Stanford	---	533 (44%)	---
Duke	---	45 (4%)	---
Method of administration			
Computer	1008 (100%)	697 (58%)	644 (100%)
Paper & Pencil	---	503 (42%)	---
Gender			
Male	336 (41%)	333 (29%)	250 (37%)
Female	487 (59%)	832 (71%)	394 (63%)
Age category			
18–29	60 (8%)	9 (1%)	172 (27%)
30–44	104 (13%)	65 (6%)	318 (49%)
45–59	294 (38%)	247 (23%)	139 (22%)
60–74	267 (34%)	576 (53%)	15 (2%)
75+	59 (8%)	192 (18%)	---
Race, ethnicity			
White, non-Hispanic	723 (88%)	956 (83%)	---
Black, non-Hispanic	64 (8%)	148 (13%)	---
Other, Non-Hispanic	2 (0%)	13 (1%)	---
Hispanic, any race	32 (4%)	42 (4%)	644 (100%)
Highest education			
< High School	15 (2%)	95 (9%)	81 (13%)
High School/GED	157 (19%)	345 (30%)	157 (24%)
Some college	336 (41%)	381 (33%)	195 (30%)
College graduate	250 (30%)	183 (16%)	151 (23%)
Advanced degree	63 (8%)	144 (12%)	60 (9%)
Comorbidities <sup>a</sup>			
None	554 (55%)	320 (27%)	247 (38%)
At least one	454 (45%)	880 (73%)	397 (62%)

Missing data were excluded from this table.

<sup>a</sup>High blood pressure, Chest pain, etc.



Table 2

Psychometric Analyses

Dimensionality Analysis <sup>a</sup>		IRT Calibration Analysis <sup>b</sup>		IRT Scoring <sup>c</sup>				
# items	EFA (eigen-value for first factor)	CFA model fit (CFI)	Range of item thresholds (theta)	Range of item slopes (theta)	SEM mean of persons (T-score)	SEM range of persons (T-score)	% of persons with SEM < 3.0 (T-score)	
<b>a. English</b>								
<b>Social Function</b>								
Ability	35	26.2	0.97	-2.6 to 1.6	1.9 to 4.5	1.4	0.7 to 6.6	92.1%
Satisfaction	44	32.5	0.96	-2.4 to 1.4	2.1 to 4.7	1.6	0.7 to 6.6	90.2%
<b>Social Relationships</b>								
Companionship	6	4.8	0.99	-3.0 to 1.0	2.3 to 5.7	3.3	1.8 to 6.6	59.0%
Emotional	16	12.2	0.99	-3.0 to 1.0	2.0 to 5.7	2.7	1.1 to 5.8	71.5%
Informational	10	7.6	0.99	-1.9 to 1.4	2.5 to 5.1	2.6	1.5 to 7.2	80.7%
Instrumental	11	8.6	0.99	-1.9 to 0.7	2.6 to 4.4	2.9	1.3 to 5.7	65.9%
Isolation	14	10.2	0.99	-1.0 to 2.6	1.8 to 4.2	2.1	1.4 to 4.9	84.3%
<b>b. Spanish</b>								
<b>Social Function</b>								
Ability	35	---	0.97	-1.9 to 1.1	2.4 to 4.9	1.2	0.7 to 3.8	92.5%
Satisfaction	44	---	0.94	-2.3 to 1.3	1.8 to 4.4	0.8	0.6 to 3.1	99.5%

EFA: Exploratory Factor Analysis  
 CFA: Confirmatory Factor Analysis  
 CFI: Comparative-fit index (> 0.95: good model fit; > 0.90: acceptable model fit)  
 IRT: Item Response Theory (see text)  
 SEM: Standard error of measurement

<sup>a</sup>The English sample was randomly split in half for use in either EFA or CFA; only CFA was performed for Spanish. The English sample size was greater than 1,000 for each EFA or CFA, except for Social Isolation (n=418); Spanish sample size, n=644.

<sup>b</sup>IRT calibration of items

<sup>c</sup>IRT estimation of SEM for individual persons, based on transformation to T-scores (see text)

**Table 3**  
 Descriptive Statistics for Social Health Domains, by Language and Selected Characteristics

a. English	Social Function				Social Relationships			
	Ability	Satisfaction	Companionship	Emotional Support	Informational Support	Instrumental Support	Social Isolation	
Gender								
Male	49.0 (8.7)	50.1 (9.4)	51.7 (9.2)	51.3 (9.3)	51.8 (9.2)	52.8 (9.2)	48.3 (9.9)	
Female	48.4 (8.9)	49.4 (9.8)	51.5 (8.4)	52.0 (8.6)	52.9 (9.3)	52.1 (9.0)	47.6 (9.3)	
Age category								
< 65	48.8 (9.3)	49.4 (9.9)	51.2 (8.9)	51.3 (9.0)	51.9 (9.3)	51.6 (9.3)	49.0 (9.8)	
65	48.7 (8.4)	50.0 (9.3)	52.0 (8.3)	52.5 (8.4)	53.3 (9.3)	53.2 (8.7)	46.2 (8.9)	
Highest education								
High School/GED	48.8 (10.0)	48.9 (9.6)	50.7 (8.7)	51.7 (8.9)	52.2 (9.3)	52.0 (9.1)	48.4 (10.5)	
College	48.8(8.8)	49.7 (9.6)	51.6 (8.7)	51.8 (8.8)	52.5 (9.3)	52.2 (9.1)	47.8 (9.5)	
b. Spanish <sup>a</sup>								
Gender								
Male	50.2 (8.4)	48.1 (7.0)	---	---	---	---	---	
Female	49.0 (9.2)	47.8 (7.6)	---	---	---	---	---	
Age category								
< 45	49.5 (9.1)	48.2 (7.2)	---	---	---	---	---	
45	49.3 (8.2)	46.9 (7.8)	---	---	---	---	---	
Highest education								
High School/GED	48.8 (8.6)	47.1 (6.7)	---	---	---	---	---	
College	49.9 (9.1)	48.4 (7.7)	---	---	---	---	---	

Values in the table represent the mean (standard deviation)

<sup>a</sup> Social Relationships items were not tested in Spanish

Table 4

Construct Validity: Known group comparisons<sup>a</sup>

<b>a. English</b>		<b>No Comorbidities</b>			<b>Comorbidities</b>			<b>Effect Size<sup>b</sup></b>
	<i>n</i>	mean	sd	<i>n</i>	mean	sd	<i>p</i> -value	
<b>Social Function</b>								
Ability	794	54.1	8.4	1,332	45.6	7.7	<0.001	0.94
Satisfaction	776	54.8	9.2	1,333	46.6	8.6	<0.001	0.85
<b>Social Relationships</b>								
Companionship	760	52.6	8.7	1,333	50.9	8.6	<0.001	0.20
Emotional	746	52.8	9.1	1,333	51.2	8.6	<0.001	0.18
Informational	750	53.5	9.6	1,332	51.9	9.0	<0.001	0.17
Instrumental	758	52.3	9.7	1,333	52.1	8.8	0.618	0.02
Isolation	291	46.3	9.6	529	48.7	9.4	0.001	-0.25
<b>b. Spanish</b>								
<b>Social Function</b>								
Ability	247	55.7	8.5	397	48.1	9.1	<0.001	0.79
Satisfaction	247	50.8	7.5	397	47.1	7.1	<0.001	0.50
<b>c. English</b>								
		<b>Panel Sample</b>		<b>Clinical Sample</b>				
<b>Social Function</b>								
Ability	940	50.0	9.8	1,186	47.8	8.1	<0.001	0.25
Satisfaction	922	50.0	9.8	1,187	49.3	9.5	0.121	0.07
<b>Social Relationships</b>								
Companionship	910	50.0	9.4	1,183	52.7	7.9	<0.001	-0.31
Emotional	900	50.0	9.6	1,179	53.1	7.9	<0.001	-0.35
Informational	896	50.0	9.7	1,186	54.3	8.5	<0.001	-0.47
Instrumental	910	50.0	9.6	1,181	53.9	8.4	<0.001	-0.43
Isolation	412	50.0	9.8	408	45.7	8.8	<0.001	0.45

<sup>a</sup> a and b: Respondents reporting no comorbidities versus those reporting any comorbidities  
c: Respondents from the online panel sample versus those from the clinical samples

$b_2$  Effect Size: Mean difference between groups / Overall standard deviation

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript