Published in final edited form as:

J Cross Cult Gerontol. 2015 September ; 30(3): 269–283. doi:10.1007/s10823-015-9266-9.

Sources of Response Bias in Older Ethnic Minorities: A Case of Korean American Elderly

Miyong T. Kim¹, Ju-Young Lee¹, Jisook Ko¹, Hyunwoo Yoon², Kim B. Kim³, and Yuri Jang²

Miyong T. Kim: mkim@nursing.utexas.edu; Ju-Young Lee: juyounglee@utexas.edu

¹School of Nursing, The University of Texas at Austin, 1710 Red River, Austin, TX 78751, USA

²School of Social Work, The University of Texas at Austin, Austin, TX, USA

³Korean Resource Center, Ellicott City, MD, USA

Abstract

The present study was undertaken to investigate potential sources of response bias in empirical research involving older ethnic minorities and to identify prudent strategies to reduce those biases, using Korean American elderly (KAE) as an example. Data were obtained from three independent studies of KAE ($N=1,297$; age = 60) in three states (Florida, New York, and Maryland) from 2000 to 2008. Two common measures, Pearlin's Mastery Scale and the CES-D scale, were selected for a series of psychometric tests based on classical measurement theory. Survey items were analyzed in depth, using psychometric properties generated from both exploratory factor analysis and confirmatory factor analysis as well as correlational analysis. Two types of potential sources of bias were identified as the most significant contributors to increases in error variances for these psychological instruments. Error variances were most prominent when (1) items were not presented in a manner that was culturally or contextually congruent with respect to the target population and/or (2) the response anchors for items were mixed (e.g., positive vs. negative). The systemic patterns and magnitudes of the biases were also cross-validated for the three studies. The results demonstrate sources and impacts of measurement biases in studies of older ethnic minorities. The identified response biases highlight the need for re-evaluation of current measurement practices, which are based on traditional recommendations that response anchors should be mixed or that the original wording of instruments should be rigidly followed. Specifically, systematic guidelines for accommodating cultural and contextual backgrounds into instrument design are warranted.

Keywords

Cultural bias; Ethnic minority elderly; Response bias

Correspondence to: Ju-Young Lee, juyounglee@utexas.edu.

Ethical Statement The manuscript has not been and will not be submitted elsewhere for publication and the contents of this manuscript are not now under consideration for publication elsewhere. There are no directly related manuscripts or abstracts, published or unpublished, by any authors of this paper. All authors have reviewed and approved the complete manuscript (including tables and figures).

Introduction

Despite increases in research on older ethnic minorities, there has been relatively little methodological investigation of sources of measurement error or discussion of ways to deal with measurement error in research on such population groups (Okazaki and Sue 1995). One source of measurement error in survey research is response bias, which can result in inaccurate assessments and false conclusions. This is especially important for researchers who use translated instruments in studies of ethnic minority populations or in cross-cultural research. Over the years, generations of specialists have made significant contributions to psychometrics and survey methodology, providing contemporary researchers with well-validated traditional survey instruments that follow classical instrumentation guidelines (Nunnally 1970). One such guideline recommends the intermittent switching of response anchors (e.g., from positive to negative and vice versa) to avoid measurement errors due to potentially mechanical responses (Dillman 2000). Yet although the influence of question wording and format on the responses to questionnaire items for the general population has been discussed in survey research (Benson and Hovevar 1985; Locker, Jokovic, and Allison 2007; Schwarz 1999), the pros and cons of switching anchors midstream within survey instruments are not well discussed in the literature, and we do not know how this practice may influence the results obtained with data from ethnic minority populations or in cross-cultural groups.

Another gap in the science of measurement as it is used in cross-cultural research is related to a lack of discussion about practical guidelines concerning the assessment and handling of “cultural bias” in survey items. There are many recommended and well-accepted methods for the translation of instruments and the achievement of conceptual equivalence between original instruments and their translated versions, such as back-translation (Brislin 1970; Berkanovic 1980) or committee review (Guillemin, Bombardier, and Beaton 1993). Back-translation is the most commonly recommended method for obtaining the cultural equivalency of instruments in cross-cultural studies. Committee review is also recommended in order to obtain semantic, idiomatic, and conceptual equivalence in translation. But these methods are applied before surveys are actually administered, and researchers do not commonly assess potential cultural bias, response bias, or their magnitudes after the survey data have been collected. Given the critical role of psychometrically sound instruments in psychological and behavioral research, further methodological investigation and discussion are needed; both may entail challenges to some of the traditional practices used in survey research on ethnic minority populations.

In this study, we investigate whether sources of response bias might be linked to cultural or contextual response styles or response formats, and we examine strategies to reduce such errors. To do so, we use multiple datasets of Korean American Elderly (KAE) as our example. Korean Americans are one of the fastest growing ethnic minorities in the United States, constituting the fifth largest Asian American subgroup (U.S. Census Bureau 2012). To date, research focusing on measurement error has been limited, especially for studies of ethnic minorities. The present study represents an attempt to further such research.

Our investigation focuses on two survey instruments: Pearlin's Mastery Scale (Pearlin and Schooler 1978) and the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff 1977). These instruments, designed to measure social competence and mental health, have been used in many studies of ethnic minorities, and they have been translated for use with Korean Americans (Jang et al. 2005a; Jang, Roh, and Chiriboga 2014; Kim et al. 2005). However, researchers have reported lower internal consistency for these scales when examining Korean American populations as opposed to other populations whose primary language is English (Jang et al. 2005b). It should be noted that such lower internal consistencies were not found for the CES-D in one study of younger Korean Americans, presumably because they were more acculturated (Kim, Landis, and Cain 2013). In the present study, therefore, we examine response bias in three KAE cohorts on two commonly used scales.

Methods

Data Sources

This study is a retrospective analysis of data from three descriptive studies using cross-sectional designs with KAE subjects ($N=1,297$) in three states (Florida, New York, and Maryland) from 2000 to 2008. The subjects in the three studies were self-identified Korean Americans 60 years old or older and living in the community. Basic characteristics of the samples from the three studies are summarized in Table 1.

The first data set was collected from 675 KAE living in Florida in 2008; the study was funded by the National Institute of Mental Health (Jang, Chiriboga, Allen, Kwak, and Haley 2010; Study 1). This study was designed to help reduce mental health disparities among underserved ethnic minority including KAE living in Florida. The second data set was obtained from the Advanced Research Institute in Geriatric Mental Health in order to examine the physical and mental health status of 420 KAE in the New York City metropolitan area (Jang et al. 2014; Study 2). Data sets used as the third data set were collected with support from the Agency for Healthcare Research and Quality for a study conducted to develop and test an innovative self-help program aimed at improving the control of high blood pressure among KAE living in Maryland (Kim, Han, Kim, and Duong 2002; Study 3).

Instruments

Pearlin's Mastery Scale (Pearlin and Schooler 1978) comprises seven items that measure the extent to which one feels confident in one's ability to control or master one's life—*sense of mastery*. Items are scored on a 4-point scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). The scale includes five negatively anchored items ("There is really no way I can solve some of the problems I have"; "Sometimes I feel that I'm being pushed around in life"; "I have little control over things that happen to me"; "I often feel helpless in dealing with the problems in life"; "There is little I can do to change many important things in my life") and two positively anchored items ("I can do just about anything I really set my mind to do"; "What happens to me in the future depends on me"). Responses to the five negatively anchored items are reverse coded, and total scores range from 7 (low sense of

mastery) to 28 (high sense of mastery). The mastery scale has shown evidence of validity and reliability in English (Bengtsson-Tops 2004) as well as in Korean (Jang, Poon, Kim, and Shin 2004; Shin, Han, and Kim 2007). Internal consistency reliabilities for the Korean version of the Mastery Scale have varied from .51 to .77 (Jang et al. 2004; Noh and Avison 1996; Shin et al. 2007).

The CES-D is a self-reported, 20-item Likert-type instrument (Andresen, Malmgren, Carter, and Patrick 1994; Radloff 1977). The 10-item CES-D short form assesses the frequency of depressive symptoms during the past week on a 4-point scale where 0=*rarely or none of the time*, 1=*some of the time*, 2=*much of the time*, and 3=*most or all of the time*. The scale includes two positively worded items (“I felt hopeful” and “I was happy”) and eight negatively worded items (e.g., “I felt depressed” and “I felt lonely”). The two positively worded items are reverse coded. The total scores have a potential range from 0 (no depressive symptoms) to 30 (severe depressive symptoms). The CES-D was first translated into Korean by Kuo in 1984; its psychometric properties have been validated in subsequent studies (e.g., Cho, Nam, and Suh 1998; Noh, Avison, and Kaspar 1992). Alpha coefficients have ranged from 0.87 to 0.89 (Cho and Kim 1998; Hurh and Kim 1990), and the test–retest reliability coefficient was 0.68 in Cho et al.’s study of the Korean population.

Analysis

Data were analyzed using SPSS 22 software for exploratory factor analysis (EFA; IBM, 2013) and the AMOS 18 (Arbuckle 2006) program for confirmatory factor analysis. Mastery scores were obtained from Studies 1, 2, and 3. CES-D scores were not available from Study 3, so we used CES-D scores from Studies 1 and 2.

A series of psychometric tests was performed, including various correlation analyses for both item-level and total-scale-level properties. In addition, EFA was used to further investigate and estimate the magnitude of each bias and to identify the source of each bias.

First, we assessed the internal consistency reliability index of each scale in order to estimate the magnitude of potential measurement error within each scale; the reliability index can be a useful measure to assess measurement error according to classical measurement theory (reliability index, $r_{tt} = \frac{V_{\infty}}{V_t}$; measurement error = $1 - r_{tt}$). Specifically, Cronbach’s alpha for each scale was calculated for the three data samples. Second, item analysis was performed to identify potential items that did not contribute to explaining the substantive variance and that exhibited possible signs of bias. Items that demonstrated low inter-item correlations and item–total correlations ($r < .30$) were subject to further qualitative examination to determine the nature of bias.

Third, upon identification of items with significant measurement bias, we recalculated the internal consistency reliability index and item analyses for each scale after removing the problematic items one by one. If the Cronbach’s alphas and the results of correlational analysis improved after deletion of an item, that item was considered to show additional evidence of systematic measurement bias. Fourth, EFA was conducted to examine whether similar patterns existed between the different psychometric properties after item analysis. We used principal axis analysis with oblique rotation as our factor extraction method.

Unlike principal components analysis, principal axis analysis specifically models error variance (Steger 2006).

Prior to the EFA, the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity (Hair, Anderson, Tatham, and Black 1995; Tabachnick and Fidell 2007) were used to check assumptions for factor analysis of the Mastery Scale and the CES-D, respectively. The appropriateness of EFA for the mastery data was demonstrated by both the Kaiser-Meyer-Olkin measure (Study 1, 0.718; Study 2, 0.734; Study 3, 0.632) and Bartlett's test (Study 1, $\chi^2=1,006.47$, $p<.001$; Study 2, $\chi^2=850.86$, $p<.001$; Study 3, $\chi^2=182.22$, $p<.001$). The appropriateness of EFA for the CES-D data was also demonstrated by the Kaiser-Meyer-Olkin measure (Study 1, 0.861; Study 2, 0.889) and Bartlett's test (Study 1, $\chi^2=1,811.25$; $p<.001$; Study 2, $\chi^2=1,405.39$, $p<.001$). Scree plots of each factor were used to examine the magnitude of each factor as well as the relative distances between that factor and any dominant or secondary factors that might or might not contribute in explaining the latent variable (the conceptual domain) of interest. As in the item analysis, after examining each factor load and the total explained variance (λ and R^2) of the full scale, a series of EFAs were conducted by deleting each item that demonstrated a low factor load on the latent factor. Throughout these empirical analytic processes, a team of bilingual and bicultural researchers held regular discussions to determine the existence and the sources of each bias.

Confirmatory factor analysis was performed to further evaluate the model fit before and after removing the problematic items, which were identified from the EFA. The goodness of fit was estimated by using the root mean square error of approximation, standardized root mean square residual, goodness of fit index, adjusted goodness of fit index, normed fit index, and comparative fit index (Bollen and Long 1993).

Results

Demographic Characteristics

The first data set was collected from 675 KAE living in Florida in 2008 (Jang et al. 2010; Study 1). Their mean age was 70.2 years ($SD=6.9$), the majority were women (58.6 %), most were married (76.6 %), and the majority lived with others (86.5 %). About two thirds (70.7 %) had less than a high school education. All were foreign born, and the average years of residence in the U.S. was 28.3 ($SD=11.8$). Around 60 % considered their health status to be good. Detailed information can be found in Jang et al. (2010).

The second data set came from a study of KAE in the New York metropolitan area (Jang et al. 2014; Study 2). These data were collected from 420 Korean immigrants, 60 years old or older, with a mean age of 71.6 years ($SD=7.6$). Half of the sample were women (51.7 %), 60.9 % were married, and 73.7 % lived with others. The majority had completed at least a high school education (69.0 %), and most practiced some form of religion (90.8 %). Their mean number of years living in the U.S. was 23.4 ($SD=7.6$), and about 45 % rated their current health status as either excellent or good. Additional details can be found in Jang et al. (2014).

The third data set was collected from 205 KAE living in Maryland, with a mean age of 67.6 years ($SD=17.8$; M. T. Kim et al. 2002; Study 3). The majority were women (64.4 %), and most were married (68.5 %). About two thirds (61.3 %) had less than a high school education and a mean length of residence in the U.S. of 15.8 years ($SD=6.7$). A little over half of these respondents considered their health to be as good as or better than that of their peers (55.1 %). Further details are presented in Kim et al. (2002).

Descriptive Psychometric Statistics

Table 2 presents means, standard deviations, Cronbach's alpha reliability coefficients, and score ranges for the three samples on the Mastery Scale and the CES-D. The mean scores for the Mastery Scale were 20.1 ($SD=2.4$) in Study 1, 19.3 ($SD=2.2$) in Study 2, and 15.4 ($SD=4.3$) in Study 3. The mean CES-D scores were 7.2 ($SD=4.6$) in Study 1 and 9.3 ($SD=4.9$) in Study 2. The Cronbach's alphas for the Mastery Scale ranged from 0.518 to 0.691 for the three studies; for the CES-D scale, alpha was 0.771 in Study 1 and 0.779 in Study 2.

Potential Items with Bias

Using item analysis and EFA to examine psychometric properties, we identified several items that contained significant measurement errors. In subsequent panel discussions among bicultural and bilingual researchers, we found two distinct types of measurement bias: (1) cultural or contextual bias and (2) response anchor or format-related bias.

Example of Cultural or Contextual Bias—Cultural or contextual bias was best shown in two items on the Sense of Mastery Scale: Item 4 ("I can do just about anything I really set my mind to do") and Item 6 ("What happens to me in the future depends on me"; see Table 3). Item analysis demonstrated that these two items consistently performed poorly (the corrected item-total correlation was low: $r=.043$ to $.266$ for Item 4 and $r=.086$ to $.162$ for Item 6 across the three studies), and when we removed them from the seven-item scale, the Cronbach's alpha for the five-item scale increased in all three studies (0.691 to 0.756 in Study 1, 0.663 to 0.805 in Study 2, and 0.518 to 0.650 in Study 3). Given that correlation coefficients in psychometric analysis are generally sensitive to the number of items in the scale (a higher number of items in a scale will likely yield higher coefficients), the increases in Cronbach's alpha with fewer items highlight the magnitude of bias contained in those two items.

The results of EFA on the sense of mastery scores from the three studies produced two factors with eigenvalues greater than one. The magnitude of the first factor extracted from all three studies was substantial; it contributed significantly to explaining a latent factor of "sense of mastery" (Study 1, 32.1 %/58.6 % total explained variance; Study 2, 40.8 %/64.1 % for two factors; Study 3, 21.1 %/31.9 % for two factors). Items 4 and 6 were loaded as the second factor on all three studies. Although the second factor's size was not negligible at first glance, we concluded that the factor was not a contributor to explaining the latent construct of "sense of mastery," but rather an indication of common bias in these two items. When we removed Items 4 and 6, the EFA result revealed that all items loaded on a single factor accounted for 51.0 %, 56.3 %, and 28.8 % of the total variance in Studies 1, 2, and 3,

respectively. Inspections of the scree plots also indicated a change in the shape after removal of the biased items (see Fig. 1). The sense of mastery plot revealed a clear elbow at the one-dimensional solution for all three studies. Together, these results indicated that the data from this scale were best portrayed when we used a one-factor solution with those two biased items removed.

The panel discussion concluded that the two items were stated in such a way that even KAE with a strong sense of mastery would find them very difficult to endorse positively, resulting in cultural bias. The attitude of placing (or stating) extreme optimism or supreme confidence in the future is discouraged in Korean culture, because such an attitude can be considered a sign of “immaturity.”

Example of Response Bias Related to Format—To determine any response bias that might be related to response anchor or format, we examined the data from the two studies that used the CES-D. Two items from the scale produced significant method bias: Item 5 (“I felt hopeful about the future”) and Item 8 (“I was happy”; see Table 4).

Item analysis demonstrated that these two items consistently performed poorly (the corrected item–total correlation was low: $r=.024$ and $.258$ for Item 5, and $r=-.139$ and $.285$ for Item 8). The scale’s internal reliability increased with the removal of these two items; Cronbach’s alpha increased significantly from 0.771 to 0.834 in Study 1, and from 0.779 to 0.872 in Study 2.

EFA indicated that the majority of items in the CES-D scale loaded on Factor 1, whereas negatively keyed items loaded on Factor 2 (see Table 4). In Study 1, Factor 1 explained 38.7 % of the variance and Factor 2 accounted for 13.1 % of the variance. In Study 2, Factor 1 accounted for 43.2 % of the variance and Factor 2 accounted for 14.1 %.

After removal of the biased items (4 and 6), all items loaded on a single factor, accounting for 47.1 % of the total variance in Study 1 and 53.6 % of the total variance in Study 2. The scree plot showed a change in shape after removal of the biased items (see Fig. 1). The CES-D plot changed to a clear elbow at the one-dimensional solution for both studies.

Similarly, data from each study were fitted to the a priori one-domain structure using confirmatory factor analysis. In particular, various fit indexes generated by data with and without the problematic items identified by EFA were compared. All fit indices (goodness of fit, adjusted goodness of fit, normed fit, comparative fit) were significantly improved after removal of the items that contained potential error variances (see Table 5). For example, comparative fits generated from the Sense of Mastery scale data across all three studies demonstrated dramatic increases when those items were removed (range before, 0.71–0.73; range after, 0.90–0.95).

Discussion

The findings of our secondary data analysis demonstrate a need for cultural sensitivity in measurement when researchers are conducting cross-cultural research or research involving cultural minority groups. Even well-established “cross-validated” scales leave ample room

for improvement in order to prevent measurement errors. Indeed our findings reinforce a discussion of measurement that is already emerging among researchers who have conducted psychological descriptive studies of Korean immigrants in North America (Han, Choi, Kim, Lee, and Kim 2008; Noh et al. 1992; Noh, Kaspar, and Chen 1998).

For example, Han et al. (2008) reported that Korean Americans are reluctant to express feelings of positive affect on the CES-D. Han et al. argued that Korean elders have difficulty in switching their cognitive anchor as they respond to questions in which the anchors are reversed. We find their argument persuasive, because we identified similar patterns and were able to document the magnitude of response bias as a consequence of having positive anchors in a predominantly negatively anchored depression scale. Noh et al. (1998), who studied Korean immigrants in Canada, also reported such a pattern and argued that Korean immigrants' responses to positively keyed items on the CES-D jeopardized the (intended) authenticity of their responses. In a prior study, Noh et al. (1992) found that Korean immigrants in Canada tended to score 3.4–4.1 times lower on positive affect items and 1.1–1.4 times higher on depressive symptom items than did European Americans. In addition, there are reports that, for various cultural reasons, Koreans and people from societies influenced by Chinese culture are not usually skilled in using affective terms to communicate emotional states (Kleinman 1982; Kleinman and Good 1985). The Korean culture discourages open or public discussion of emotional problems, because Koreans believe that one's maturity is measured by an ability to control one's feelings (Pang 1984). Some authors have speculated that these cultural influences are significant factors in the distinctive manifestation patterns of emotional problems, namely somatization, in Far Eastern Asian people, including Koreans (Kleinman 1982).

Our finding of cultural or contextual bias demonstrates that empirical evidence can show how certain cultural influences can create systematic measurement error. The notion that cultural norms and expectations can influence individuals' responses to psychological survey questions and thus create distinct response biases can be found in the literature (Kim 1995; Matsumoto 1994). As we have observed in our analysis of the Mastery Scale, KAE's reluctance to endorse items exhibiting extremely optimistic forecasts of the future can be traced to their cultural roots. Cultures influenced by Confucian philosophy (Korean, Chinese, and Japanese) generally encourage caution about the future (Kim 1995), for they believe that expressing extreme optimism about the future is a sign of immaturity that can invite "bad luck" (Pang 1984). This sort of cultural bias has also been mentioned in studies of Chinese Americans (Li and Hicks 2010) and Japanese in Japan (Iwata and Roberts 1996).

A more fundamental question related to the variable of sense of mastery might be related to this variable's ultimate utility as a buffering factor between certain antecedents (e.g., stress) and health outcomes (e.g., depression). The role of sense of mastery in a classic model of stress, coping, and health outcomes is based on the following premises: first, individuals can be evaluated in terms of the extent to which they view changes in their environment as the result of their own initiatives (control). Second, the concept of self-mastery assumes that perceived competence in interpersonal skills is an important aspect of personal resources (Fry 1989). But it is not clear whether these premises are applicable to groups that differ culturally from the dominant Western population. Because the concepts of "control" and

“manipulating the environment” represent Western perspectives, it is possible that people from different cultures have different notions about those concepts. The traditional Korean worldview places a certain emphasis on “fatalism,” which encourages people to accept negative life events without blaming themselves (or others). Yet these fatalistic attitudes are qualitatively different from attribution to an “external locus of control” or a “sense of resignation.” Rather, fatalistic beliefs within Korean culture often contribute to individuals’ spirituality, enabling them to view life from a more global perspective. Of course, for studies involving any immigrant group, it should be recognized that individual differences among members of the group may affect each individual’s value system and the manifestations of his or her worldview. Future studies of Korean Americans that use sense of mastery as a variable should test the appropriateness of the construct for the population by examining relevant contextual factors, such as the acculturation process of the individual.

It has been reported that Koreans and Japanese are prone to making response mistakes when survey items contain what would be considered “negative sensitive” expressions in Korean and Japanese (Park 2014). Although there are many guidelines for achieving a linguistic cultural equivalence in the translation of instruments, it would seem that even with the most methodical and faithful translation, the cultural adequacy of a survey or diagnostic tool in a given culture does not guarantee its reliability or validity in another (Canino et al. 1997).

Interpretation of the present results must be cautioned because the data came from three studies of a single ethnic group and because we resorted to classical measurement theory to inform our analysis. This would have been a much richer study had we included other cultural groups or age groups for comparison or contrast. In addition, a study of multiple ethnic groups using a more modern measurement theoretical framework such as item response theory would have been helpful for calculating item differential scores for items identified to be problematic. Despite these limitations, however, the results of this study make several important contributions to the literature.

First, the findings of this investigation provide evidence to propel methodological discussion in contemporary survey research further, particularly research involving ethnic minority elders. In clearly illustrating the magnitude of measurement errors created by biases related to culture or to response format, we are urging researchers to pay attention to more than the meticulous translation of survey instruments. Moreover, we are challenging current instrumentation practices such as the switching of an instrument’s anchors midstream. While such a practice makes sense in theory because it ought to prevent mechanical response bias, there is no guideline regarding how to handle potentially significant measurement error variance such as that found in studies like ours. As the geriatric population continues to grow rapidly in many parts of the world, researchers’ discussion of appropriate measurement protocols for this population must continue. This will include focused discussion about the continued application of the legacy instrumentation guideline of switching anchors for cognitively fragile elders or ethnic minority elders.

Acknowledgments

This study used data obtained from three studies of Korean American elderly in three states (Florida, New York, and Maryland) from 2000 to 2008. Funding sources for data collection include the National Institute of Mental

Health (R21 MH081094), the Advanced Research Institute in Geriatric Mental Health (R25 MH068502), and the Agency for Healthcare Research and Quality (R01 HS13160).

References

- Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults—evaluation of a short-form of the CES-D (center for epidemiologic studies depression scale). *American Journal of Preventive Medicine*. 1994; 10(2):77–84. [PubMed: 8037935]
- Arbuckle, JL. Amos (version 7.0) [computer program]. Chicago: SPSS; 2006.
- Bengtsson-Tops A. Mastery in patients with schizophrenia living in the community: relationship to sociodemographic and clinical characteristics, needs for care and support, and social network. *Journal of Psychiatric and Mental Health Nursing*. 2004; 11(3):298–304.10.1111/j.1365 [PubMed: 15149377]
- Benson J, Hocevar D. The impact of item phrasing on the validity of attitude scales for schoolchildren. *Journal of Educational Measurement*. 1985; 22(3):231–240.
- Berkanovic E. The effect of inadequate language translation on Hispanics' responses to health surveys. *American Journal of Public Health*. 1980; 70(12):1273–1276. [PubMed: 7435745]
- Bollen, KA.; Long, JS. Testing structural equating models. Newbury Park: Sage; 1993.
- Brislin RW. Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*. 1970; 1(3):185–216.
- Canino G, Lewis-Fernandez R, Bravo M. Methodological challenges in cross-cultural mental health research. *Transcultural Psychiatry*. 1997; 34(2):163–184.10.1177/136346159703400201
- Census Bureau, U. S. The Asian population: 2010. Washington: U.S. Department of Commerce; 2012.
- Cho MJ, Kim KH. Use of the center for epidemiologic studies depression (CES-D) scale in Korea. *The Journal of Nervous and Mental Disease*. 1998; 186(5):304–310. [PubMed: 9612448]
- Cho MJ, Nam JJ, Suh GH. Prevalence of symptoms of depression in a nationwide sample of Korean adults. *Psychiatry Research*. 1998; 81(3):341–352.10.1016/S0165-1781(98)00122-X [PubMed: 9925185]
- Dillman, DA. Mail and internet surveys: The tailored design method. Vol. 2. New York: Wiley; 2000.
- Fry PS. Mediators of stress in older adults: conceptual and integrative frameworks. *Canadian Psychology/Psychologie Canadienne*. 1989; 30(4):636–649.
- Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*. 1993; 46(12):1417–1432. [PubMed: 8263569]
- Hair, JF., Jr; Anderson, RE.; Tatham, RL.; Black, WC. Multivariate data analysis. 4. Upper Saddle River: Prentice-Hall; 1995.
- Han HR, Choi YJ, Kim MT, Lee JE, Kim KB. Experiences and challenges of informal caregiving for Korean immigrants. *Journal of Advanced Nursing*. 2008; 63(5):517–526.10.1111/j.1365-2648.2008.04746.x [PubMed: 18727754]
- Huh WM, Kim KC. Correlates of Korean immigrants' mental health. *The Journal of Nervous and Mental Disease*. 1990; 178(11):703–711. [PubMed: 2230758]
- IBM Corp. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: Author; 2013.
- Iwata N, Roberts RE. Age differences among Japanese on the center for epidemiologic studies depression scale: an ethnocultural perspective on somatization. *Social Science & Medicine*. 1996; 43(6):967–974. [PubMed: 8888466]
- Jang Y, Poon LW, Kim SY, Shin BK. Self-perception of aging and health among older adults in Korea. *Journal of Aging Studies*. 2004; 18(4):485–496.
- Jang Y, Borenstein AR, Chiriboga DA, Mortimer JA. Depressive symptoms among African American and white older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2005a; 60(6):313–P319.
- Jang Y, Kim G, Chiriboga D. Acculturation and manifestation of depressive symptoms among Korean-American older adults. *Aging & Mental Health*. 2005b; 9(6):500–507. [PubMed: 16214697]

- Jang Y, Chiriboga DA, Allen JY, Kwak J, Haley WE. Willingness of older Korean-American adults to use hospice. *Journal of the American Geriatrics Society*. 2010; 58(2):352–356.10.1111/j.1532-5415.2009.02684.x [PubMed: 20374409]
- Jang Y, Roh S, Chiriboga DA. The impact of acculturation on depressive symptoms: a comparison of older Korean Americans in two areas. *Asian American Journal of Psychology*. 2014; 5(3):200–205.10.1037/a0032591
- Kim MT. Cultural influences of depression experiences among Korean Americans. *Journal of Psychosocial Nursing*. 1995; 33(2):13–18.
- Kim MT, Han HR, Kim KB, Duong DN. The use of traditional and Western medicine among Korean American elderly. *Journal of Community Health*. 2002; 27(2):109–120. [PubMed: 11936756]
- Kim MT, Han HR, Shin HS, Kim KB, Lee HB. Factors associated with depression experience of immigrant populations: a study of Korean immigrants. *Archives of Psychiatric Nursing*. 2005; 19(5):217–225.10.1016/j.apnu.2005.07.004 [PubMed: 16226673]
- Kim E, Landis AM, Cain KK. Responses to CES-D: European American versus Korean American adolescents. *Journal of Child and Adolescent Psychiatric Nursing*. 2013; 26(4):254–261. [PubMed: 24180605]
- Kleinman A. Neurasthenia and depression: a study of somatization and culture in china. *Culture, Medicine and Psychiatry*. 1982; 6(2):117–190.10.1007/BF00051427
- Kleinman, A.; Good, B. *Culture and depression: Studies in the anthropology and cross-cultural psychiatry of affect and disorder*. Berkeley and Los Angeles: University of California Press; 1985.
- Li Z, Hicks MHR. The CES-D in Chinese American women: construct validity, diagnostic validity for major depression, and cultural response bias. *Psychiatry Research*. 2010; 175(3):227–232.10.1016/j.psychres.2009.03.007 [PubMed: 20006386]
- Locker D, Jokovic A, Allison P. Direction of wording and responses to items in oral health-related quality of life questionnaires for children and their parents. *Community Dentistry and Oral Epidemiology*. 2007; 35(4):255–262. [PubMed: 17615012]
- Matsumoto, D. *Cultural influences on research methods and statistics*. Pacific Grove: Brooks/Cole; 1994.
- Noh S, Avison WR. Asian immigrants and the stress process: A study of Koreans in Canada. *Journal of Health and Social Behavior*. 1996; 37(2):192–206. [PubMed: 8690879]
- Noh S, Avison WR, Kaspar V. Depressive symptoms among Korean immigrants: assessment of a translation of the center for epidemiologic studies-depression scale. *Psychological Assessment*. 1992; 4(1):84–91.10.1037/1040-3590.4.1.84
- Noh S, Kaspar V, Chen X. Measuring depression in Korean immigrants: assessing validity of the translated Korean version of CES-D scale. *Cross-Cultural Research*. 1998; 32(4):358–377.10.1177/106939719803200403
- Nunnally, JC. *Introduction to psychological measurement*. New York: McGraw-Hill; 1970.
- Okazaki S, Sue S. Methodological issues in assessment research with ethnic minorities. *Psychological Assessment*. 1995; 7(3):367–375.
- Pang, KM. Unpublished doctoral dissertation. Catholic University of America; Washington, DC: 1984. *Everyday life, health, and illness of the elderly Korean immigrants: Cultural construction of illness*.
- Park K. A contrastive study of Japanese and Korean negative sensitive items: a grammaticalization approach. *Language Sciences*. 2014; 45:152–172.10.1016/j.langsci.2014.06.020
- Pearlin LI, Schooler C. The structure of coping. *Journal of Health and Social Behavior*. 1978; 19(1):2–21.10.2307/2136319 [PubMed: 649936]
- Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977; 1(3):385–401.10.1177/014662167700100306
- Schwarz N. Self-reports: how the questions shape the answers. *American Psychologist*. 1999; 54(2):93–105.
- Shin HS, Han HR, Kim MT. Predictors of psychological well-being amongst Korean immigrants to the United States: a structured interview survey. *International Journal of Nursing Studies*. 2007; 44(3):415–426.10.1016/j.ijnurstu.2006.04.007 [PubMed: 16806219]

- Steger MF. An illustration of issues in factor extraction and identification of dimensionality in psychological assessment data. *Journal of Personality Assessment*. 2006; 86(3):263–272.10.1207/s15327752jpa8603_03 [PubMed: 16740111]
- Tabachnick, BG.; Fidell, LS. Using multivariate statistics. Boston: Pearson Education; 2007.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1.
The scree plots changes of MS and CES-D after removing the biased items

Table 1

Demographic profile of the respondents (over 60 years old)

	Study 1 (N=672)	Study 2 (N=420)	Study 3 (N=205)
Sex (%)			
Male	41.4	48.3	34.6
Female	58.6	51.7	64.4
Age (year)±SD (range 60–96)	70.2±6.9	71.6±7.6	67.6±17.8
Marital status (%)			
Married	76.6	60.9	68.5
Unmarried	23.4	39.1	31.5
Education (%)			
Less than high school	29.3	31.0	61.3
High school or greater	70.7	69.0	38.7
Length of stay in US (year)	28.3±11.8	23.4±7.6	15.8±6.7
Self-rated health (%)			
Excellent	15.3	8.3	10.2
Good	47.9	36.4	44.9
Fair	28.1	40.2	22.0
Poor	8.6	15.0	22.9

Study 1: Collected data from Florida

Study 2: Collected data from New York

Study 3: Collected data from Central Maryland

Table 2

Means, standard deviations, and Cronbach's alpha, and item total statistics of the scales

		Mean±SD	Cronbach's Alpha	Range
MS (7 items)	Study 1	20.1±3.4	0.691	7–28
	Study 2	19.3±3.3	0.663	7–28
	Study 3	15.4±4.3	0.518	7–28
CES-D (10 items)	Study 1	7.2±4.6	0.771	0–40
	Study 2	9.3±4.9	0.779	0–40

MS: The Pearlin Mastery Scale

CES-D: The Center for Epidemiological Studies-Depression Scale

Table 3

Item loadings and variance and Cronbach's alpha changes for the mastery scale

	Study 1 (N=672)		Study 2 (N=420)		Study 3 (N=205)		Corrected Item-Total Correlation (Study 1/2/3)
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	
Factors							
Sense of mastery							
1. There is really no way I can solve some of the problems I have.	0.579		0.671		0.402		.355/.466/.220
2. Sometimes I feel that I'm being pushed around in life.	0.678		0.744		0.468		.438/.508/.410
3. I have little control over things that happen to me.	0.744		0.792		0.695		.532/.622/.389
5. I often feel helpless in dealing with the problems in life.	0.754		0.729		0.555		.532/.401/.403
7. There is little I can do to change many important things in my life.			0.779		0.540		.521/.511/.265
0.760							
Items related to cultural or contextual inequivalence bias							
4. I can do just about anything I really set my mind to do.		0.828		0.880		0.589	.266/.114/.043
6. What happens to me in the future depends on me.		0.867		0.861		0.493	.162/.027/.086
Eigenvalue	2.6	1.5	2.9	1.6	2.1	1.4	
Variance (%)	32.1	21.5	40.8	23.3	21.1	10.8	
After Item 4 and 6 removed							
	Before	After	Before	After	Before	After	
Variance (%)	58.6 (2 factors)	51.0	64.1 (2 factors)	56.3	31.9 (2 factors)	28.8	
Cronbach's Alpha	0.691	0.756	0.663	0.805	0.518	0.650	

Table 4

Item loadings and variance and Cronbach's alpha changes for the CES-D

Factors	Study 1 (N=672)		Study 2 (N=420)		Corrected Item-Total Correlation (Study 1/2)
	Factor 1	Factor 2	Factor 1	Factor 2	
Depressive symptoms					
1. I was bothered by things that usually don't bother me.	0.642		0.667		.480/.521
2. I had trouble keeping my mind on what I was doing.	0.664		0.775		.483/.611
3. I felt depressed.	0.819		0.811		.666/.691
4. I felt that everything I did was an effort.	0.694		0.722		.533/.538
6. I felt fearful.	0.668		0.778		.528/.651
7. My sleep was restless.	0.679		0.624		.394/.440
9. I felt lonely.	0.679		0.723		.539/.586
10. I could not get "going."	0.699		0.750		.548/.628
Items related response format related bias					
5. I can do just about anything I really set my mind to do.		0.769		0.833	.258/.024
8. What happens to me in the future depends on me.		0.769		0.832	.285/.139
Eigenvalue	3.9	1.3	4.3	1.1	
Variance (%)	38.7	13.1	43.2	14.4	
After Item 5 & 8 removed					
	Before	After	Before	After	
Variance (%)	51.8 (2 factors)	47.1	57.6 (2 factors)	53.6	
Cronbach's Alpha	0.771	0.834	0.779	0.872	

Table 5

Confirmatory Factor Analysis for the Mastery Scale and CES-D

Fitness Index	Criteria	MS		CES-D					
		Study 1		Study 2		Study 3		Study 1	
		Before	After	Before	After	Before	After	Before	After
$\chi^2(p)$	$p > .05$	279.31 (<.001)	38.90 (<.001)	240.52 (<.001)	32.72 (<.001)	61.55 (<.001)	16.78 (.005)	222.46 (<.001)	112.00 (<.001)
df		14	5	14	5	14	5	35	20
RMSEA	<.10	0.169	0.101	0.202	0.118	0.129	0.107	0.091	0.083
SRMR	<.05	0.105	0.038	0.124	0.043	0.091	0.052	0.061	0.042
GFI	>.90	0.905	0.977	0.869	0.968	0.921	0.970	0.937	0.959
AGFI	>.90	0.811	0.931	0.737	0.905	0.841	0.909	0.901	0.927
NFI	>.90	0.725	0.949	0.703	0.945	0.667	0.879	0.875	0.932
CFI	>.90	0.733	0.955	0.713	0.953	0.710	0.908	0.892	0.943

df degree of freedom, *RMSEA* Root mean squared error of approximation, *SRMR* Standardized root mean square residual, *GFI* Goodness of fit index, *AGFI* Adjusted goodness of fit index, *NFI* Normed fit index, *CFI* Comparative fit index