Survey Research Methods (2020)
Vol. 14, No. 4, pp. 417-437
doi:10.18148/srm/2020.v14i4.7413

ISSN 1864-3361
http://www.surveymethods.org

Measurement Equivalence of Subjective Well-Being Scales under the Presence of Acquiescent Response Style for the Racially and Ethnically Diverse Older Population in the United States

Sunghee Lee
Institute for Social Research
University of Michigan, U.S.A

Elizabeth Vasquez
Department of Epidemiology and Biostatistics
University at Albany
State University of New York
U.S.A.

Lindsay H. Ryan
Institute for Social Research
University of Michigan. U.S.A.

Jacqui Smith
Institute for Social Research
University of Michigan, U.S.A.

Objectives. With valid assessment of subjective well-being (SWB) at the forefront of social science research, comparability of measurement scales designed to capture SWB across racial/ethnic groups has been questioned. This study examines measurement properties of well-established SWB scales and their comparability between older Hispanics and Whites in the U.S.. Methods. We analyzed the Health and Retirement Study data in order to examine measurement invariance of the satisfaction with life (SWL), positive affect (PAF) and purpose in life (PIL) scales across non-Hispanic Whites, Hispanics interviewed in English and Hispanics interviewed in Spanish through multigroup confirmatory factor analysis and examined their validity by linking their latent scores with covariates of SWB. Examinations of measurement properties further considered acquiescence response style. Results. Strict, scalar and metric invariance was observed for SWL, PAF and PIL, respectively. However, when latent scores estimated from these invariance models were regressed on the validation measures, the relationship was weaker for Hispanics than Whites, suggesting a lower level of validity for Hispanics than Whites. A lower level of invariance was observed for respondents who acquiesced than their counterpart, while the validity was not necessarily hampered by acquiescent response style. Discussion. Our analysis suggests that the traditional measurement invariance test may not be effective under the presence of acquiescent response style for SWB instruments that use the Likert-type response scales. Research into cross-cultural measurement of SWB that considers systematic difference in conceptualization of SWB as well as response styles may improve our ability to understand SWB of the increasingly diverse population.

Keywords: Subjective Well-being; Measurement Equivalence; Measurement Invariance; Minority and Diverse Populations; Cross-cultural Differences

1 Introduction

1.1 Importance of subjective well-being

Over history across cultures, there is agreement that living a "good life" benefits individuals and the societies in which they live. Although ongoing debate investigates the factors determining an individual's well-being (WB) and whether these factors are universal across populations or population subgroups, most agree that one's report of subjective well-being (SWB) is derived from a complex combination of individual characteristics and contextual factors (Diener et al., 2018; National Research Council, 2013; OECD, 2013). Considering demographic shifts around the world (United Nations, 2010), a valid assessment of SWB for an increasingly diverse population is important so that, by focusing on SWB, research can identify population subgroups likely at risk of poor SWB and develop policy to improve SWB in these groups.

In the U.S., Hispanics have emerged as the largest racial/ethnic minority group, estimated at 58.9 million as of 2017 accounting for 18.1 percent of the population, according to the American Community Survey (ACS) (Alonzo,

Contact information: Sunghee Lee, Institute for Social Research, University of Michigan, 436 Thompson St. Ann Arbor, MI 48104, (sungheel@umich.edu)

2018). While much younger than the general population currently, Hispanics are on the way to becoming an important group in the aging population. Despite the importance, a report by the National Research Council (2013) acknowledged the lack of our understandings about WB of Hispanics, particularly older Hispanics. For example, assessment studies of WB measures may collect data only in English (Kobau et al., 2010). With 2018 ACS indicating almost 40% of Hispanics having limited English proficiency, such studies apply only to a subgroup within the Hispanic population, those proficient in English enough to participate in research.

1.2 Subjective well-being and its measures

There are various measures of SWB (see Annex A and B in OECD, 2013 and Appendix 1 in National Research Council, 2013 for examples). These measures stem from three conceptually distinct yet related elements: evaluative, affective (experienced) and eudaimonic WB (Diener, Kahneman, et al., 2010; Ryan, 2016). Below, we review these elements and exemplary survey instruments.

First, evaluative WB, arguably one of the most frequently measured constructs in social science, consists of cognitive appraisals on one's life and circumstances. The satisfaction with life scale (Diener et al., 1985) and the Cantril's ladder scale (Cantril, 1966) are some of the examples of global measures of evaluative WB. As the evaluation criteria are entirely up to each individual, experiencing a major life challenge, such as unemployment or widowhood, is shown to affect evaluative WB (Diener et al., 2006; Lucas et al., 2004). Evaluative WB can help to disentangle questions about individual and group differences in factors determining a good life. For example, for older adults who reported similar levels of life satisfaction, what predicted life satisfaction differed by sex: for women, self-rated health and depressive symptoms were important, whereas widowhood mattered for men (Berg et al., 2006). The concept of evaluative WB is also considered with a nuanced focus. For instance, life satisfaction can be assessed within specific domains, such as family life or financial situation (Campbell et al., 1976). These domain-specific measures are shown to be less subject to mood or affective state than global satisfaction (Schwarz et al., 1987).

The second major element of SWB reflects affective experiences, typically assessed with positive and negative affect. The affective circumplex (Russell & Pratt, 1980) provides a useful conceptualization of affective WB, where individuals rate the extent to which they felt a variety of positively and negatively valance adjectives that cover a range of intensities, such as happy, excited, calm, bored, angry and frustrated, to name a few. A primary purpose of assessing affective WB is to understand individual's subjective experiences rather than evaluations. Examples of affective WB measures include the experience-sampling method that captures ecological momentary assessments (Csikszentmihalyi et al., 1977) and the positive and negative affect scale that assesses affect felt in the last few weeks (Watson et al., 1988). Although these measures show within-person change over time (Carstensen et al., 1999), they are largely considered measures of stable traits.

Eudaimonia adds the third element to conceptualizing SWB. Unlike the first two, eudaimonic WB focuses on functioning of a person on the assumption that people strive for more than happiness in life. This makes constructs, such as meaningfulness, self-worth, autonomy and relations with others, relevant to SWB (Deci & Ryan, 2000). Examples of eudaimonic WB measures include the purpose in life scale (Ryff, 1989), the flourishing scale (Diener, Wirtz, et al., 2010) and the meaning in life scale (Steger et al., 2006).

1.3 Cross-cultural comparative research on SWB measures: Anomaly with U.S. Hispanics

SWB scores are widely used for comparative studies. The OECD Better Life Index, for example, compares countries on the global life satisfaction score[1]. In multi-racial and multi-ethnic countries like the U.S., racial/ethnic subgroups are also widely compared. For instance, Gallup has compared Hispanics and Blacks against Whites over time[2].

From anthropological perspectives, existing comparative research on SWB, such as OECD Better Life Index, employs an etic approach where researchers impose a preconceived set of concepts or measures across groups in the comparison. As noted by Kagawa-Singer and colleagues (Kagawa-Singer et al., 2014), social science disciplines widely practice this approach. Unfortunately, under the etic framework, it becomes difficult to disentangle whether the group-level differences in SWB scores are due to substantive differences, measurement artifacts stemming from cultural norms or some combination of these two (Diener, 2009; Oishi, 2010). One of the well-known measurement artifacts of SWB arises due to response scale formats (National Research Council, 2013). In particular, response style, a tendency where respondents choose certain answers (e.g., "strongly agree" or "yes") regardless of the question content (Paulhus, 1991), has been shown to hamper measurement properties of the positive and negative affect scales (Schneider, 2016). With response style being particularly relevant for cross-cultural research (Baumgartner & Steenkamp, 2001) and with response style interacting with racial/ethnic subgroups (Hui & Triandis, 1989; Liu et al., 2018; G. Marín et al., 1992; G. Marín & Marín, 1991), SWB comparisons are likely affected, at least partially, by this artifact. Despite the increased diversity in our societies, empirical evidence on this area is limited (OECD, 2013).

Acknowledging this issue, the adult population of ages 50

[1] http://www.oecdbetterlifeindex.org/

[2] http://www.gallup.com/poll/163688/blacks-hispanics-life-satisfaction-2008.aspx?g_source=

and older in the U.S. was compared on four measures of SWB in Figure **??**. The comparison focused on Hispanics and non-Hispanic Whites, where Hispanics were further divided by interview language, following previous work (e.g., Lee & Schwarz, 2014). The first three measures were about evaluative WB: the single-item global life satisfaction (GLS), the single-item momentary life satisfaction (MLS) and the average score on the five items of the satisfaction with life scale (SWL) (Diener et al., 1985). The last was eudaimonic WB: the average score of the seven items in the purpose of life scale (PIL) (Ryff, 1989). Appendix A includes wording of these items.

Should these groups be ranked on these measures (e.g., Helliwell et al., 2013), results should be similar across measures, given that they tap into SWB-related constructs (OECD, 2013). However, Figures **??** offered a different story. On GLS and MLS, Hispanics were ranked lower than Whites. In particular, Hispanics interviewed in Spanish as a group appeared least satisfied on GLS. However, on SWL, Spanish-interviewed Hispanics ranked the highest, appearing most satisfied with life, with English-interviewed Hispanics least satisfied and Whites in the middle. On the other hand, all three groups scored similarly on PIL. One cannot draw a reasonable conclusion about which group has a better life from these comparisons. With these measures providing less than coherent results, they may be deemed lacking "statistical quality", a concept introduced by (OECD, 2008).

A closer examination on these measures provides three methodological aspects that may explain the anomaly. First, these measures differ in the number of items: GLS and MLS are measured with a single item, while SWL and PIL with multiple items. Second, they use different response scales. GLS and MLS ask respondents to rate their satisfaction on a response scale ranging from "completely satisfied" to "not satisfied at all". SWL and PIL ask respondents to rate their agreement to each item using a Likert scale from "strongly agree" to "strongly disagree". The Likert scale is deemed more prone to response style bias than other response scales (Baumgartner & Steenkamp, 2001). Although not SWB-specific, there is a large volume of literature linking U.S. Hispanics to acquiescent response styles (e.g., Davis et al., 2019; Davis et al., 2011; G. Marín et al., 1992). Third, this response style bias becomes exacerbated for multi-item measures that are directionally unbalanced (Baumgartner & Steenkamp, 2001). In our case, PIL is balanced with four out of seven items worded in the direction of low PIL (e.g., "My daily activities often seem trivial and unimportant to me.") and the rest in the direction of high PIL (e.g., "I am an active person in carrying out the plans I set for myself"). On the other hand, all five SWL items are worded in the direction of high satisfaction, making SWL directionally unbalanced.

With cultural norms being the driver of response style (Hui & Triandis, 1989), Hispanics, particularly those interviewed in Spanish, are shown more likely to acquiesce on any given statement by choosing "agree" or "strongly agree" than Whites (Davis et al., 2011; G. Marín et al., 1992). This leads to higher scores by Hispanics than Whites on multi-item measurement scales that use directionally unbalanced items and a Likert response scale. Among the four SWB measures in Figure **??**, SWL is likely to be most affected by acquiescent response style. This observation is consistent with reports about higher SWL scores for Latin American countries than other countries (Diener, 2009). This artifactual influence of cultural norms associated with race and ethnicity on the measurement of SWB is listed as one of the recommended research areas by U.S. (National Research Council, 2013) but remains largely unexplored. This study aims to examine existing SWB measures in order to improve our ability to incorporate cultural considerations in SWB measurements, which may render hypothesis-driven methodological research opportunities.

2 Data and Methods

2.1 Data source description

Data for this study came from the 2010 Health and Retirement Study (HRS). Started in 1992, HRS is an on-going population-based longitudinal study of older adults aged 50 and older in the U.S. As a study of older persons near and through the retirement, HRS includes an extensive list of SWB measures, mostly in a module administered to a random half of the total sample (see J. Smith et al. (2013) for detail about this module). Unlike most modules in HRS administered by interviewers, this module is self-administered by respondents. A total of 7,828 eligible respondents completed this self-administered module in 2010.

HRS is particularly advantageous for examining cross-racial/ethnic comparability, as interviews have been conducted in both English and Spanish since its inception, and, particularly in 2010, racial/ethnic minorities were oversampled (Ofstedal & Weir, 2011; Sonnega et al., 2014). Moreover, being a longitudinal study, responses to a question can be linked to future outcomes (e.g., subsequent mortality in Lee et al., 2016). This data set-up provides a unique opportunity to examine existing SWB measures, their comparability across racial/ethnic groups and their relationship with known covariates of SWB over the course of time.

2.2 Dependent variables

This study focuses on three multi-item SWB scales, covering all three elements of SWB: satisfaction with life (SWL) for evaluative WB; positive affects (PAF) for affective WB; and purpose in life (PIL) for eudaimonic WB, as summarized in Table **??**. See Appendix A for item wording. These scales were chosen, because they differ not only in conceptual elements of SWB but also in methodological factors associated

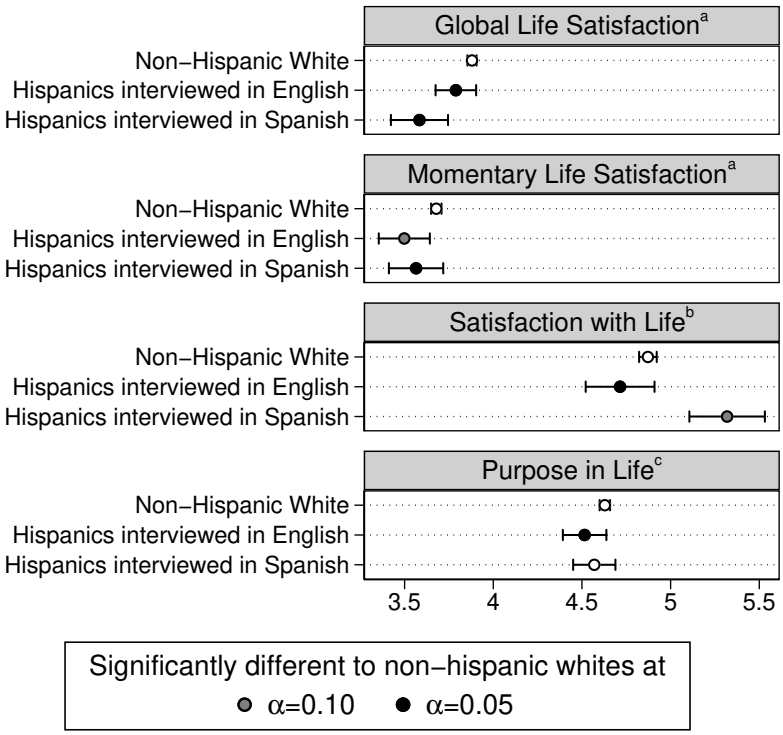

Figure 1. Means and 95% C.I. of four measures for Subjective Well-Being by ethnicity and language.
[a] 6-pt single item scale from 1, "completely satisfied" to 6, "not at all satisfied". [b] Multi-item scale with directionally unbalanced items using 7-pt agreement response scales from 1, "strongly disagree" to 7, "Strongly agree". [c] Multi-item Scale with directionally balanced items using 6-pt agreement response scales from 1, "strongly disagree" to 6 "strongly agree".

with measurement artifacts. In particular, these instruments differ with respect to 1) directional balance of items (e.g., all items written in the direction of high WB vs. a mixture of items in the direction of high and low WB), 2) response scales and 3) the number of response points. These factors allowed us to identify potential sources of measurement artifacts and examine their consequences empirically. Note that in order to make the interpretation of the results easier, responses were re-coded so that a higher score in each scale means a higher level of the measured construct.

2.3 Independent variable

The focus of this study is comparability of the selected SWB measures between Hispanics and non-Hispanic Whites, where Hispanics are further divided by interview language. While not perfect, interview language is shown to approximate acculturation levels among minorities in the U.S. (Lee et al., 2011). This, in turn, suggests that interview language itself may affect the measurement artifacts of SWB. Hence, throughout our analysis, SWL, PAF and PIL scales and their comparability were examined as a function of three groups as follows: 1) non-Hispanic Whites ($n = 5,645$), 2) Hispanics interviewed in English ($n = 407$) and 3) Hispanics interviewed in Spanish ($n = 332$). Our analysis excluded 1,444 respondents whose were neither Hispanics nor non-Hispanic Whites.

Further, we considered respondents' acquiescent response style in examining measurement properties by designating respondents as acquiescent or non-acquiescent respondents. While there are various approaches for this designation as summarized by Van Vaerenbergh and Thomas (2013)[3] given the data availability within HRS, we used illogical responses given to a balanced scale as an indicator of acquiescent respondents as described in Baumgartner and Steenkamp (2001). In particular, we used the PIL scale and classified

[3]More advanced methods require certain data set-ups (e.g., multiple scales using directionally balanced items; Billiet and McClendon, 2000)

Table 1
Description of Selected Subjective Well-Being Scales

Selected scale	Conceptual element	Number of items	Item direction	Response scale	Response points	Origin
Satisfaction with life (SWL)	Evaluative	5	Unbalanced	1. Strongly disagree to 7. Strongly agree	7-pt	Diener et al., 1985
Positive affect (PAF)	Affect	13	Unbalanced	1. Very much to 5. Not at all	5-pt	Watson et al., 1988
Purpose in life scale (PIL)	Eudaimonic	7	Balanced	1. Strongly disagree to 6. Strongly agree	6-pt	Ryff, 1989

respondents who chose "strongly agree" or "agree" on items written in the direction of high PIL as well as items written in the direction of low PIL as "acquiescers" and the remaining respondents as "nonacquiescers".

2.4 Validation measures

We chose two subjective (GLS used in Figure **??** and self-rated health) and four objective outcomes (total wealth in U.S. dollars, the number of chronic conditions, cognitive functioning scores and subsequent mortality status ascertained for 2014) as validation measures. Total wealth combined the net value of primary residence, other real estate, transportation, business, IRA, stock, checking and savings account, CDs, government bonds, other bonds, treasury bills and all other saving and subtracted total mortgage, other home loans and other debts. The chronic condition considered how many among the following eight conditions a given respondent was ever diagnosed with: high blood pressure, diabetes, cancer, lung diseases, heart problems, stroke, emotional/psychiatric problems and arthritis. The cognitive functioning reflected the ability to recall a list of 10 words immediately as well as in a delayed fashion (Ofstedal et al., 2005). Details about the objective outcome variables are provided by Bugliari et al. (2016).

These validation measures are well-known correlates of SWB (OECD, 2013). For example, income is frequently used in the analysis of SWB (e.g., Sacks et al., 2010). If a given scale captures SWB well, then the latent score of the scale should co-vary with income. This illustrates concurrent validity of that scale. If the relationship between the income and the latent score is consistent across comparison groups, then comparability in concurrent validity can be assumed.

2.5 Analysis steps

We first examined reliability of SWL, PIL and PAF through Cronbach's α for the overall sample and for three comparison groups (non-Hispanic Whites, Hispanics interviewed in English, and Hispanics interviewed in Spanish). We then conducted measurement invariance tests using factorial models where each of SWL, PIL and PAF was individually and independently modeled. Measurement invariance across three comparison groups in these models was examined through multigroup confirmatory factor analyses (MG-CFA) (Jöreskog, 1971), following the standard testing of configural, metric, scalar and strict measurement invariance models. The goodness of model fit was evaluated with comparative fit index (CFI), root mean square error of approximation (RMSEA) and standardized root mean square residual (SRMR). Following Hu and Bentler (1999) and Byrne (2009), CFI $>$ 0.95, RMSEA $<$ 0.06 and SRMR $<$ 0.09 were used as reference points of good fit. When a reasonable fit was not established for the configural invariance model (CFI $<$ 0.900 or RMSEA $>$ 0.100), we examined modifications using estimated model parameters, modification indices as well as methodological factors (e.g., item direction) before proceeding to higher invariance models. Although the use of such cut-points has been shown to lead inconsistent conclusions (Lai & Green, 2016), the large sample size of our study should eliminate inconsistency (Kenny et al., 2015). The relative fit of measurement invariance models was examined with changes in CFI primarily, supplemented by changes in RMSEA and SRMR as the latter two tend to over-reject models (Chen, 2007). The cut-off points of ΔCFI ≥ -0.010, ΔRMSEA ≥ 0.015 and ΔSRMR ≥ 0.030 were used for non-invariance of loadings; and ΔCFI ≥ -0.010, ΔRMSEA ≥ 0.015 and ΔSRMR ≥ 0.010 for non-invariance of intercepts and residuals (Chen, 2007; Cheung & Rensvold, 2002). It should be noted that we examined partial invariance models at each step (e.g., Hirschfeld & von Brachel, 2014) but excluded them from this paper as they were not effective in addressing measurement invariance issues.

With the invariance model identified from MG-CFA, the next analysis step examined the comparability of concurrent validity of each SWB scale across race/ethnicity. For this, we used structural equation models (SEM) that regressed a given latent score estimated from the invariance structure on each validation measure. Regression coefficients were estimated separately for each group and compared across groups using Whites as a reference group.

The last step introduced acquiescent response style into the examination of properties of reliability, measurement in-

variance as well as construct validity. We examined Cronbach's α for nonacquiescers and acquiescers and tested the measurement invariance model of each scale identified from earlier step separately for acquiescers and nonacquiescers in order to examine whether measurement invariance properties were independent - of the acquiescent response style. We fitted the SEM from the second analysis step separately for acquiescers and nonacquiescers and compared regression coefficients. As response style is not considered in typical measurement invariance tests, SWB scales in SEM was modelled with the invariance structure applicable for the overall sample.

R packages, psych, lavaan and SemTools, were used for the analysis. For handling missing on the measurement items in this study, we use the full information maximum likelihood (FIML) approach for fitting MG-CFA and SEM models. Under the missing mechanisms of at random, FIML is reported to be advantageous compared to other approaches (e.g., list-wise deletion, mean imputation; Arbuckle, 1996; Enders, 2001).

3 Results

3.1 Reliability, confirmatory factor models and measurement invariance

The reliability varied across SWL, PAF and PIL, with SWL and PAF associated with high Cronbach's α at 0.900 and 0.920 and PIL at 0.780. When examining the reliability across comparison groups, there was some variation on SWL and PAF, with Cronbach's α ranging from 0.850 for Hispanics interviewed in Spanish to 0.890 for Hispanics interviewed in English and to 0.900 for non-Hispanic Whites on SWL and 0.900 for Hispanics interviewed in Spanish to 0.920 for Hispanics interviewed in English and to 0.930 for non-Hispanic Whites on PAF. The variation was larger for PIL, with Cronbach's α as low as 0.660 for Hispanics interviewed in Spanish and as high as 0.790 for non-Hispanic Whites. Across SWL, PAF and PIL, the reliability was highest consistently for non-Hispanic Whites, followed by Hispanics interviewed in English and Hispanics interviewed in Spanish.

When the parameters were freely estimated for each group in the invariance model (i.e., configural invariance; see Supplemental Table B1 to B2 for standardized estimates of model parameters for each scale), none of the three scales showed a good fit. SWL as an individual scale appeared somewhat better than the other two with CFI = 0.966 and SRMR = 0.030 but RMSEA = 0.148 (see Table 2). The fit of configural invariance models of PAF and PIL was similarly poor with CFI < 0.9 and RMSEA at 0.111 and 0.125, although SRMR suggested a reasonable fit. Considering the parameter estimates from these configural invariance models, the modification indices of these three scales and item directionality of PIL, respective configural invariance models were modified as follows: for SWL, correlated residuals between Q1 and 2 and across Q3, 4 and 5 and fixed residual variance of Q1, 2 and 4; for PAF, correlated residuals of Q4, 5, 7 and 11; for PIL, Q2, 4, 5 and 6 for PIL due to their direction in low sense of eudaimonia as well as lower factor loadings. These modifications resulted in an improved fit of configural invariance models. Specifically, SWL showed an improvement on RMSEA from 0.148 to 0.089; PAF improved moderately from CFI = 0.885, RMSEA = 0.111 and SRMR = 0.046 to CFI = 0.947, RMSEA = 0.079 and SRMR = 0.35; and PIL improved substantially from CFI = 0.866, RMSEA = 0.125 and SRMR = 0.058 to CFI = 0.982, RMSEA = 0.061 and SRMR = 0.020.

With a reasonable fit established for the configural invariance models with modifications, a series of invariance models were tested as presented in Table 2. The model comparisons within each scale indicated strict invariance for SWL, scalar invariance for PAF and metric invariance for PIL.

3.2 Relationship between estimated latent scores and validation measures

Given that SWL, PAF and PIL are conceptualized as elements of SWB, their estimated latent scores were examined in relation to the well-known covariates of WB. Here, the relationship is viewed as indicative of concurrent validity. We further examined how comparable concurrent validity was across Whites and Hispanics. As shown in Table 3, the regression coefficients of the latent scores on the validation measures were largely significant. However, there was a variation across groups. For Whites, the relationship was significant across all latent scores and all validation measures; for both Hispanic groups, mortality was not related to any of the latent scores; and particularly for Spanish-interviewed Hispanics, wealth was not related to any of the latent scores.

Magnitudes of estimated coefficients varied across groups. Notably, coefficients were estimated smaller for Hispanics interviewed in Spanish than the other groups. All three latent scores were associated with single-item GLS, self-rated health and wealth consistently at a significantly higher level for non-Hispanic Whites than Spanish-interviewed Hispanics. For example, on PIL, the coefficient of wealth was estimated positive and significant at 0.181 (Std. Err.=0.015) for non-Hispanic Whites but negative and non-significant at −0.071 (Std. Err.=0.072) for Spanish-interviewed Hispanics. These estimated coefficients were significantly different ($p < 0.01$). Differences between non-Hispanic Whites and English-interviewed Hispanics observed on the coefficient estimates between latent scores and GLS. For other validation measures, non-Hispanic Whites and English-interviewed Hispanics appeared comparable. In the case of SWL, the number of chronic conditions showed a significantly higher level of relationship for English-interviewed Hispanics with an estimated coefficient at −0.283 (Std.

Table 2

Goodness of Fit for Satisfaction with Life, Positive Affects and Purposein Life. Models from Multigroup Confirmatory Factor Analysis for Non-Hispanic Whites, Hispanics Interviewed in English and Hispanics Interviewed in Spanish, 2010 Health and Retirement Study

	Overall Sample (n = 6,334)						Nonacquiescers (n=4,365)						Acquiescers (n=1,969)					
Model	df Δ	χ^2 Δ	p Δ	CFI Δ	RMSEA Δ	SRMR Δ	df Δ	χ^2 Δ	p Δ	CFI Δ	RMSEA Δ	SRMR Δ	df Δ	χ^2 Δ	p Δ	CFI Δ	RMSEA Δ	SRMR Δ
Satisfaction with life																		
Configural	15	713	< 0.01	0.97	0.15	0.03	15	563	< 0.01	0.96	0.16	0.03	15	199	< 0.01	0.97	0.14	0.03
Modified[a]																		
Configural Invariance	12	212	< 0.01	0.99	0.09	0.02	12	64	< 0.01	1.00	0.06	0.01	12	216	< 0.01	0.97	0.16	0.03
Metric Invariance	8	30	< 0.01	−0.00	−0.02	0.00	8	22	0.01	−0.00	−0.01	0.00	8	18	0.02	−0.00	−0.03	0.00
Scalar Invariance	8	73	< 0.01	−0.00	−0.00	0.00	8	28	< 0.01	−0.00	0.00	0.00	8	44	< 0.01	−0.01	−0.01	0.00
Strict Invariance	4	156	< 0.01	−0.01	0.01	0.00	4	87	< 0.01	−0.01	0.01	0.00	4	59	< 0.01	−0.01	0.00	0.00
Positive affects																		
Configural	195	5236	< 0.01	0.89	0.11	0.05	195	3638	< 0.01	0.89	0.11	0.04	195	1884	< 0.01	0.86	0.12	0.05
Modified[b]																		
Configural	177	2502	< 0.01	0.95	0.08	0.04	177	1770	< 0.01	0.95	0.08	0.03	177	989	< 0.01	0.93	0.08	0.04
Metric Invariance	24	31	0.15	0.00	−0.01	0.00	24	50	0.00	0.00	−0.00	0.00	24	24	< 0.47	0.00	−0.01	0.00
Scalar Invariance	24	346	< 0.01	−0.01	0.00	0.00	24	199	< 0.01	−0.01	−0.00	0.00	24	180	< 0.01	−0.01	0.00	0.00
Strict Invariance	26	1041	< 0.01	−0.02	0.01	0.00	26	559	< 0.01	−0.02	0.01	0.00	26	371	< 0.01	−0.03	0.01	0.01
Purpose in life																		
Configural	42	1416	< 0.01	0.87	0.12	0.06	42	478	< 0.01	0.96	0.09	0.03	42	720	< 0.01	0.67	0.16	0.08
Modified[b]																		
Configural	24	210	< 0.01	0.98	0.06	0.02	24	84	< 0.01	0.99	0.04	0.01	24	124	< 0.01	0.95	0.08	0.04
Metric Invariance	12	32	0.00	−0.00	−0.01	0.01	12	15	0.27	0.00	−0.01	0.00	12	45	< 0.01	−0.02	−0.00	0.01
Scalar Invariance	12	150	< 0.01	−0.01	0.01	0.01	12	64	< 0.01	−0.01	0.01	0.00	12	120	< 0.01	−0.05	0.01	0.01
Strict Invariance	14	408	< 0.01	−0.04	0.02	0.01	14	91	< 0.01	−0.01	0.01	0.01	14	99	< 0.01	−0.04	0.00	0.01

[a] Modified with correlated residuals and fixed residual variances [b] Modified with correlated residuals.

Table 3

Relationship between Latent Scores and Validation Measures of Subjective Well-Being by Race/Ethnicity/Interview Language and by Acquiescent Response Style, 2010 and 2014 Health and Retirement Study

			Hispanics							
	Non-Hispanic Whites[a]		English Interview		Spanish Interview		Nonacquiescers[b]		Acquiescers	
Latent Score and Validation Measure	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.	Est.	Std. Err.
Satisfaction with life[c]										
Single-item global life satisfaction	0.585[e]	0.009	0.458[d**]	0.040	0.264[d***]	0.054	0.565[e]	0.011	0.524[d*]	0.016
Self-rated health (1. Poor - 5. Excellent)	0.383[e]	0.012	0.437[e]	0.041	0.176[§***]	0.056	0.373[e]	0.014	0.355[e]	0.020
Number of chronic conditions (0-8)	−0.172[e]	0.014	−0.283[§*]	0.047	−0.217[e]	0.055	−0.178[e]	0.015	−0.179[e]	0.023
Cognitive functioning score (0-20)	0.072[e]	0.014	0.118[e]	0.052	0.010	0.058	0.075[e]	0.016	0.046[e]	0.024
Mortality in 2014 (0: deceased; 1: alive)	0.096[e]	0.014	0.043	0.052	0.057	0.057	0.095[e]	0.016	0.076[e]	0.023
Wealth ($, log transformed)	0.222[e]	0.013	0.200[e]	0.050	−0.029***	0.058	0.207[e]	0.015	0.201[e]	0.022
Postitive affects[d]										
Single-item global life satisfaction	0.447[e]	0.011	0.385[e]	0.044	0.330[§*]	0.053	0.435[e]	0.013	0.440[e]	0.019
Self-rated health (1. Poor - 5. Excellent)	0.400[e]	0.012	0.450[e]	0.040	0.290[§*]	0.054	0.398[e]	0.013	0.404[e]	0.019
Number of chronic conditions (0-8)	−0.230[e]	0.013	−0.276[e]	0.048	−0.146[e]	0.059	−0.219[e]	0.015	−0.196[e]	0.023
Cognitive functioning score (0-20)	0.180[e]	0.014	0.198[e]	0.051	0.166[e]	0.059	0.157[e]	0.016	0.208[e]	0.023
Mortality in 2014 (0: deceased; 1: alive)	0.142[e]	0.014	0.061	0.052	0.022*	0.059	0.109[e]	0.016	0.131[e]	0.023
Wealth ($, log transformed)	0.173[e]	0.014	0.166[e]	0.050	−0.021***	0.059	0.162[e]	0.015	0.197[e]	0.023
Purpose in life[d]										
Single-item global life satisfaction	0.392[e]	0.013	0.244[§**]	0.056	0.219[§*]	0.074	0.344[e]	0.015	0.410[§*]	0.025
Self-rated health (1. Poor - 5. Excellent)	0.368[e]	0.013	0.318[e]	0.052	0.194[§*]	0.072	0.319[e]	0.015	0.367[e]	0.025
Number of chronic conditions (0-8)	−0.221[e]	0.015	−0.179[e]	0.056	−0.264[e]	0.070	−0.208[e]	0.016	−0.240[e]	0.027
Cognitive functioning score (0-20)	0.172[e]	0.015	0.127[e]	0.058	0.193[e]	0.074	0.148[e]	0.017	0.148[e]	0.028
Mortality in 2014 (0: deceased; 1: alive)	0.145[e]	0.015	0.006**	0.057	0.088	0.072	0.105[e]	0.017	0.177[§*]	0.027
Wealth ($, log transformed)	0.181[e]	0.015	0.092	0.059	−0.071**	0.072	0.140[e]	0.016	0.175[e]	0.027

[a] Reference group in the comparison by race/ethnicity/interview language [b] Reference group in the comparison by acquiescent response style

[c] Modified with correlated residuals and fixed residual variances. [d] Modified with correlated residuals.

[e] Estimated regression coefficient significant at $p < 0.05$

Estimated regression coefficient significantly different from the reference group at: * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

Err.=0.047) than for non-Hispanic Whites with −0.172 (Std. Err.=0.014).

3.3 Reliability, measurement invariance and construct validity with acquiescent response style

When using illogical response patterns on the seven PIL items, 31.1% of the respondents were classified as acquiescers. The rates varied across comparison groups at 29.8%, 31.5% and 51.8% for non-Hispanic Whites, Hispanics interviewed in English and Hispanics interviewed in Spanish, respectively. The proportion of acquiescers was significantly higher for Hispanics interviewed in Spanish than for Whites ($p < 0.001$). Note this was not a reflection of straight-lining as only 24 respondents chose either "strongly agree" "somewhat agree" across all PIL items.

The reliability of SWL and PAF did not vary between nonacquiescers and acquiescers with Cronbach's α at 0.900 for nonacquiescers and 0.890 for acquiescers on SWL and 0.930 for nonacquiescers and 0.910 for acquiescers on PAF. However, on PIL, Cronbach's α was particularly low for acquiescers at 0.600 while it was high 0.850 for nonacquiescers.

Table 2 provides series of measurement invariance models fitted separately to nonacquiescers ($n = 4,365$) and acquiescers ($n = 1,969$). The fit of non-modified configural invariance models of PAF and PIL were better when fitted to nonacquiescers than acquiescers. For example, the configural invariance model of PIL without modification fitted to the overall sample data showed a weak fit (CFI=0.866; RMSEA = 0.125; SRMR = 0.058). The same model was quite robust when fitted to nonacquiescers (CFI = 0.957; RMSEA = 0.085; SRMR=0.033) but particularly bad when fitted to acquiescers (CFI = 0.670; RMSEA = 0.157; SRMR = 0.083). A similar yet attenuated pattern was observed for PAF. This was not true for SWL: the non-modified configural invariance model was better for acquiescers (CFI = 0.970; RMSEA = 0.137; SRMR = 0.029) than nonacquiescers (CFI = 0.962; RMSEA = 0.158; SRMR = 0.031), although CFI and RMSEA within each group offered inconsistent information about the model fit.

When modified with correlated residuals, the configural invariance model showed a good fit for both acquiescers and nonacquiescers in case of PIL, a modest fit in case of PAF; still, the goodness of fit was better for nonacquiescers than acquiescers across PIL and PAF. For SWL, the modified model fitted well nonacquiescers, but the fit statistics remained inconsistent for acquiescers with CFI indicating a good fit and RMSEA indicating a bad fit. In fact, the modification to SWL produced an improved fit for nonacquiescers but a worsened fit for acquiescers based on all three fit indices. With respect to measurement invariance of the modified models, our analysis indicated a higher level of invariance across all SWL, PAF and PIL with nonacquiescers than acquiescers. Focusing on nonacquiescers, strict invariance was observed for SWL and PIL and scalar invariance for PAF. With acquiescers, configural invariance was not established for SWL; and metric invariance for the modified PAF and configural invariance for the modified PIL were established.

The estimated regression coefficients of all validity measures in explaining latent scores of SWL, PAF and PIL in Table 3.B were universally significant at $p<0.05$ for both nonacquiescers and acquiescers. While the differences in these estimates between nonacquiescers and acquiescers were small, the relationships appeared somewhat stronger for nonacquiescers than for acquiescers on SWL, somewhat mixed between nonacquiescers and acquiescers on PAF and stronger for acquiescers than for nonacquiescers on PIL. For example, the coefficient of SWL on GLS was 0.565 (Std. Err.=0.011) for nonacquiescers, significantly larger than 0.524 (Std. Err.=0.016) for acquiescers. However, the coefficient of PAF on GLS was 0.435 (Std. Err.=0.013) for nonacquiescers, not significantly different from 0.440 (Std. Err.=0.019) for acquiescers. Yet, the coefficient of PIL on GLS was 0.344 (Std. Err.=0.015) for nonacquiescers, significantly smaller than 0.410 (Std. Err.=0.025) for acquiescers.

4 Discussion

We examined measurement properties of well-established SWB scales for older Whites and Hispanics in the U.S. Among SWL, PIL and PAF scales, PAF appeared most desirable on reliability; SWL appeared most desirable on measurement invariance across non-Hispanic Whites and Hispanics broken down by interview language; and PIL performed the worst on reliability and measurement invariance. Modifications improved the fit of configural invariance models of all scales. The modification to configural invariance models was informed by estimated parameters as well as modification indices for SWL and PAF; however, for PIL, it was informed by a methodological factor through adding correlated residuals among items written in the direction of low sense of eudaimonia.

However, when the predicted latent scores of SWL, PAF and PIL were examined in relation to validation measures, the observed relationship was much more attenuated for Hispanics, particularly those interviewed in Spanish, than for Whites across all three scales. This indicates that the level of concurrent validity of these scales is higher for Whites than Hispanics. In sum, these three SWB scales are different with respect to measurement invariance between Whites and Hispanics but offer the same implications for concurrent validity, with higher validity for Whites than Hispanics. Additionally, measurement reliability was largely irrelevant for measurement validity when acquiescent response style was present.

Our attempt to assess the role of acquiescent response style in testing measurement invariance suggested that mea-

surement invariance as well as the model fit were dramatically worse when the invariance models were fitted to respondents associated with acquiescent response style than their counterpart, except for the configural invariance model of SWL with no modification. However, concurrent validity of these scales was not necessarily worse for acquiescers than nonacquiescers. Between nonacquiescers and acquiescers, the validity appeared somewhat better for the former on SWL, about the same on PAF and yet better for the latter on PIL. It should be noted that the invariance model of PIL included modifications informed by the direction of measurement items, while the modification to SWL and PAF was informed by parameter estimates and modification indices.

Our analysis suggests that, under the presence of acquiescent response style, the traditional measurement invariance test may not be effective for SWB instruments that use the Likert-type responses. With the increase of cross-cultural research which often compares distinctive cultural groups, response style is already a well-recognized concern (e.g., Harzing, 2006; P. B. Smith, 2004). Particularly for the U.S., it is an issue for comparative research involving Hispanics (Davis et al., 2019). Standard measurement invariance testing does not consider response style and may confound invariance with acquiescent response style.

Let us focus on SWL in our analysis that uses five items written in the same direction of high satisfaction on the agree-disagree Likert response scale. As currently measured, SWL itself may measure respondents' true satisfaction combined with their tendency to simply agree with the items. Imagine two hypothetical groups of people: one who are truly satisfied with life and have no response style; and the other who are unsatisfied with life and have a high acquiescent response tendency. Analysis of scales such as SWL may describe these two groups as though they are equally satisfied and may suggest that the measurement of SWL is invariant between the two groups. Obviously, this scale suffers from lacking validity, as true traits differ vastly between these groups. Although not shown in this analysis, when focusing on nonacquiescers, the White-Hispanic differences in concurrent validity mostly disappeared.

What then do we do in order to develop comparable measures and test their comparability for SWB research with diverse population subgroups such as those in the U.S.? Following the framework of measurement equivalence by Scheuch (1968), Van de Vijver (1998), and Van de Vijver and Tanzer (2004) we elaborate on issues with the methods for analyzing SWB scales and developing measurement items for SWB scales and the conceptualization of SWB.

For the methods end, using simple sums or means of scale items (e.g., Pavot & Diener, 2008) for group comparisons assumes complete invariance across groups. Our analysis suggests that such an approach overlooks realities of not only response style artifacts but also measurement properties of these scales and may arrive at erroneous conclusions. Newly developed statistical methods of testing measurement invariance (see Davidov et al., 2014) may guide the remedy.

Particularly for the response style issues, it is worthwhile to discuss SWL and PIL. Both use a Likert response scale. However, the differences in freely estimated loadings between Spanish-interviewed Hispanics and Whites in Supplemental Table B1 were much larger for PIL than for SWL. Some of the items in PIL showed particularly low loadings for Spanish-interviewed Hispanics than other groups (e.g., Q7). These items are worded in a direction (high PIL) opposite to the rest. When residuals of low PIL items were modelled as correlated, not only the model fit improves but also loadings of PIL items became more comparable across groups. However, loadings of the items whose residuals were set to be correlated became smaller, more so for Hispanic groups (e.g., Q6). Acquiescent response style may, in fact, have been at play with Spanish-interviewed Hispanics who were shown to acquiesce more than the other two groups. If they chose responses, "strongly agree", "somewhat agree", or "slightly agree", regardless of item content, then 1) on directionally balanced scales (e.g., PIL), items worded in a different direction would show a low relationship with the construct; but 2) on unbalanced scales (e.g., SWL), this may not emerge. Further, with directionally balanced scales, some of this measurement artifact can be addressed, for example, by including correlated residuals (Brown, 2015) as done with the modified PIL in this study. However, such remedies cannot be applied for scales with unbalanced items.

This may add evidence of how directionally balanced scales may reduce the effect of acquiescent response style biases. In fact, statistical methods for detecting and controlling for acquiescent response style often requires more than one scale with multiple, directionally balanced items (e.g., Billiet & McClendon, 2000; Liu et al., 2018) or greatly benefit from such data structure (e.g., Javaras & Ripley, 2007).

On the conceptualization part, this study showed that measurement issues were more prevalent with Hispanics interviewed in Spanish than English. Spanish-interviewed Hispanics are likely to have low English proficiency and to hold Hispanic-specific cultural values. For them, PIL may posit problems. The concept of PIL highlights the future (e.g., "I enjoy making plans for the future and working to make them a reality."). However, Hispanic cultural values presentismo, a tendency to place emphasis on the present time (G. Marín & Marín, 1991). Furthermore, none of the SWB scores was associated with subsequent mortality for Hispanics and with wealth for Spanish-interviewed Hispanics. Perhaps, to Hispanics, particularly, Spanish-interviewed Hispanics, the concept of PIL may not be directly applicable in understanding their WB; the measurement items may not translate in Spanish in the way that the items function equivalently between languages; and, further, mortality and wealth may not be per-

tinent to the concept of WB.

This leads to the questionnaire development and translation that consider construct and item bias, when involving culturally and linguistically distinct groups in the analysis. In particular for the U.S., as Spanish questionnaires are an essential tool for Hispanic data collection (Korey & Lascher, 2006; Lee & Schwarz, 2014), special attention is needed in considering constructs and item bias for developing and translating questionnaires. Traditional and Web-based cognitive interviews (e.g., Behr et al., 2012; Willis & Gordon, 2005) on the topic of SWB, combined with experiments with questionnaire design and translation of SWB measures (Lee et al., 2019; Lee et al., 2016; Lee & Schwarz, 2014), may provide fruitful insights into ways to improving our ability to understand how people perceive their own WB and to design instruments that disentangles substantive constructs of SWB from measurement artifacts. Transparency in translation and questionnaire designs will further improve our ability to increase measurement equivalence, which encompasses the concept of measurement invariance as well as comparable measurement validity, suitable for cross-cultural research (Davidov & De Beuckelaer, 2010; Harkness et al., 2010).

Acknowledgements

This work was supported partially by grants (Grant #U01 AG009740 and #R01 AG040635) from the National Institute on Aging at the U.S. National Institutes of Health.

References

Alonzo, F. (2018). Hispanic heritage month 2018: Profile america facts for features: Cb18-ff.07. https://www.census.gov/content/dam/Census/library/visualizations/2018/comm/hispanic-fff-2018.pdf

Arbuckle, J. (1996). Full information estimation in the presence of incomplete data. In G. Marcoulides & R. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 243–277). Mahwah, NJ: Erlbaum Publishers.

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, *38(2)*, 143–156. https://doi.org/10.1509/jmkr.38.2.143.18840

Behr, D., Kaczmirek, L., Bandilla, W., & Braun, M. (2012). Asking probing questions in web surveys. *Social Science Computer Review*, *30(4)*, 487–498. https://doi.org/10.1177/0894439311435305

Berg, A. I., Hassing, L. B., McClearn, G. E., & Johansson, B. (2006). What matters for life satisfaction in the oldest-old? *Aging & Mental Health*, *10(3)*, 257–264. https://doi.org/10.1080/13607860500409435

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling: A Multidisciplinary Journal*, *7(4)*, 608–628. https://doi.org/10.1207/S15328007SEM0704_5

Brown, T. A. (2015). *Confirmatory factor analysis for applied research (2nd ed.)* New York; NY, Guilford Press.

Bugliari, D., Campbell, N., Chan, C., Hayden, O., Hurd, M., Main, R., & Emanuel, N. (2016). RAND HRS Data Documentation, Version P. http://hrsonline.isr.umich.edu/modules/meta/rand/randhrsp/randhrs_P.pdf

Byrne, B. M. (2009). *Structural equation modeling with amos: Basic concepts, applications, and programming*.

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). The quality of american life: Perceptions, evaluations, and satisfactions. https://www.russellsage.org/publications/quality-american-life

Cantril, H. (1966). The pattern of human concerns. https://www.worldcat.org/title/pattern-of-human-concerns/oclc/2082808

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously. a theory of socioemotional selectivity. *The American Psychologist*, *54(3)*, 165–181. http://www.ncbi.nlm.nih.gov/pubmed/10199217

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14(3)*, 464–504. https://doi.org/10.1080/10705510701301834

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9(2)*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, *6(3)*, 281–294. https://doi.org/10.1007/BF02138940

Davidov, E., & De Beuckelaer, A. (2010). How harmful are survey translations? a test with schwartz's human values instrument. *International Journal of Public Opinion Research*, *22(4)*, 485–510. https://doi.org/10.1093/ijpor/edq030

Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40(1)*, 55–75. https://doi.org/10.1146/annurev-soc-071913-043137

Davis, R. E., Johnson, T. P., Lee, S., & Werner, C. (2019). Why do latino survey respondents acqui-

esce? respondent and interviewer characteristics as determinants of cultural patterns of acquiescence among latino survey respondents. *Cross-Cultural Research*, *53(1)*, 87–115. https://doi.org/10.1177/1069397118774504

Davis, R. E., Resnicow, K., & Couper, M. P. (2011). Survey response styles, acculturation, and culture among a sample of mexican american adults. *Journal of Cross-Cultural Psychology*, *42(7)*, 1219–1236. https://doi.org/10.1177/0022022110383317

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11(4)*, 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Diener, E. (2009). Culture and well-being. https://doi.org/10.1007/978-90-481-2352-0

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49(1)*, 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Diener, E., Kahneman, D., & Helliwell, J. (2010). International differences in well-being. https://doi.org/10.1093/acprof:oso/9780199732739.001.0001

Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and open questions in the science of subjective well-being. *Collabra: Psychology*, *4(1)*, 15. https://doi.org/10.1525/collabra.115

Diener, E., Lucas, R. E., & Scollon, C. N. (2006). Beyond the hedonic treadmill: Revising the adaptation theory of well-being. *American Psychologist*, *61(4)*, 305–314. https://doi.org/10.1037/0003-066X.61.4.305

Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97(2)*, 143–156. https://doi.org/10.1007/s11205-009-9493-y

Enders, C. K. (2001). The performance of the full information maximum likelihood estimator in multiple regression models with missing data. *Educational and Psychological Measurement*, *61(5)*, 713–740. https://doi.org/10.1177/0013164401615001

Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P., & Smith, T. W. (2010). Survey methods in multinational, multiregional, and multicultural contexts. https://www.wiley.com/en-us/Survey+Methods+in+Multinational%5C%2C+Multiregional%5C%2C+and+Multicultural+Contexts-p-9780470177990

Harzing, A.-W. (2006). Response styles in cross-national survey research. *International Journal of Cross Cultural Management*, *6(2)*, 243–266. https://doi.org/10.1177/1470595806066332

Helliwell, J., Layard, R., & Sachs, J. (2013). World happiness report. https://s3.amazonaws.com/happiness-report/2012/World_Happiness_Report_2012.pdf

Hirschfeld, G., & von Brachel, R. (2014). Multiple-group confirmatory factor analysis in r–a tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research & Evaluation*, *19(7)*. https://pareonline.net/getvn.asp?v=19&n=7

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6(1)*, 1–55. https://doi.org/10.1080/10705519909540118

Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, *20(3)*, 296–309. https://doi.org/10.1177/0022022189203004

Javaras, K. N., & Ripley, B. D. (2007). An "unfolding" latent variable model for likert attitude data. *Journal of the American Statistical Association*, *102(478)*, 454–463. https://doi.org/10.1198/016214506000000960

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36(4)*, 409–426. https://doi.org/10.1007/BF02291366

Kagawa-Singer, M., Dressler, W. W., George, S. M., & Elwood, W. N. (2014). *The cultural framework for health: An integrative approach for research and program design and evaluation*. Bethesda, MD.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of rmsea in models with small degrees of freedom. *Sociological Methods & Research*, *44(3)*, 486–507. https://doi.org/10.1177/0049124114543236

Kobau, R., Sniezek, J., Zack, M. M., Lucas, R. E., & Burns, A. (2010). Well-being assessment: An evaluation of well-being scales for public health and population estimates of well-being among us adults. *Applied Psychology: Health and Well-Being*, *2(3)*, 272–297. https://doi.org/10.1111/j.1758-0854.2010.01035.x

Korey, J. L., & Lascher, E. L. (2006). Macropartisanship in california. *Public Opinion Quarterly*, *70(1)*, 48–65. https://doi.org/10.1093/poq/nfj011

Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when rmsea and cfi disagree. *Multivariate Behavioral Research*, *51(2–3)*, 220–239. https://doi.org/10.1080/00273171.2015.1134306

Lee, S., Alvarado-Leiton, F., Vásquez, E., & Davis, R. E. (2019). Impact of the terms “regular” or “pasable”

as spanish translation for "fair" of the self-rated health question among us latinos: A randomized experiment. *American Journal of Public Health*, *e1–e8*. https://doi.org/10.2105/AJPH.2019.305341

Lee, S., McClain, C., Webster, N., & Han, S. (2016). Question order sensitivity of subjective well-being measures: Focus on life satisfaction, self-rated health, and subjective life expectancy in survey instruments. *Quality of Life Research*, *25(10)*, 2497–2510. https://doi.org/10.1007/s11136-016-1304-8

Lee, S., Nguyen, H. A., & Tsui, J. (2011). Interview language: A proxy measure for acculturation among asian americans in a population-based survey. *Journal of Immigrant and Minority Health*, *13(2)*. https://doi.org/10.1007/s10903-009-9278-z

Lee, S., & Schwarz, N. (2014). Question context and priming meaning of health: Effect on differences in self-rated health between hispanics and non-hispanic whites. *American Journal of Public Health*, *104(1)*, 179–185.

Liu, M., Suzer-Gurtekin, Z. T., Keusch, F., & Lee, S. (2018). Response styles in cross-cultural surveys. In *Advances in comparative survey methods* (pp. 477–499). https://doi.org/10.1002/9781118884997.ch22

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2004). Unemployment alters the set point for life satisfaction. *Psychological Science*, *15(1)*, 8–13. https://doi.org/10.1111/j.0963-7214.2004.01501002.x

Marín, G., Gamba, R. J., & Marín, B. V. (1992). Extreme response style and acquiescence among hispanics. *Journal of Cross-Cultural Psychology*, *23(4)*, 498–509. https://doi.org/10.1177/0022022192234006

Marín, G., & Marín, B. V. (1991). *Research with hispanic populations*. Sage Publications.

National Research Council. (2013). Subjective well-being: Measuring happiness, suffering, and other dimensions of experience (A. A. Stone & C. Mackie, Eds.). https://doi.org/10.17226/18548

OECD. (2008). Quality framework for oecd statistical activities. https://www.oecd.org/sdd/qualityframeworkforoecdstatisticalactivities.htm

OECD. (2013). OECD guidelines on measuring subjective well-being. Organisation for Economic Cooperation; Development.

Ofstedal, M. B., Fisher, G. G., & Herzog, A. R. (2005). Documentation of cognitive functioning measures in the health and retirement study. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-006.pdf

Ofstedal, M. B., & Weir, D. R. (2011). Recruitment and retention of minority participants in the health and retirement study. *The Gerontologist*, *51*(Suppl. 1), S8–S20. https://doi.org/10.1093/geront/gnq100

Oishi, S. (2010). Culture and well-being: Conceptual and methodological issues. In *International differences in well-being* (pp. 34–69). https://doi.org/10.1093/acprof:oso/9780199732739.003.0003

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes. measures of social psychological attitudes, vol. 1* (pp. 17–59). San Diego, CA, Academic Press.

Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology*, *3(2)*, 137–152. https://doi.org/10.1080/17439760701756946

Russell, J. A., & Pratt, G. (1980). A description of the affective quality attributed to environments. *Journal of Personality and Social Psychology*, *38(2)*, 311–322.

Ryan, L. H. (2016). Subjective well-being. In S. K. Whitbourne (Ed.), *Encyclopedia of adulthood and aging*. https://doi.org/10.1002/9781118521373

Ryff, C. D. (1989). Happiness is everything, or is it? explorations on the meaning of psychological well-being. *Journal of Personality and Social Psychology*, *57(6)*, 1069–1081. https://doi.org/10.1037/0022-3514.57.6.1069

Sacks, D. W., Stevenson, B., & Wolfers, J. (2010). Subjective well-being, income, economic development and growth (no. w16441). https://ssrn.com/abstract=1689373

Scheuch, E. K. (1968). The cross-cultural use of sample surveys: Problems of comparability. historical social research – historische sozialforschung, *18*, 104–138. https://doi.org/10.2307/20755744

Schneider, S. (2016). Extracting response style bias from measures of positive and negative affect in aging research. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *73(1)*, gbw103. https://doi.org/10.1093/geronb/gbw103

Schwarz, N., Strack, F., Kommer, D., & Wagner, D. (1987). Soccer, rooms, and the quality of your life: Mood effects on judgments of satisfaction with life in general and with specific domains. *European Journal of Social Psychology*, *17(1)*, 69–79. https://doi.org/10.1002/ejsp.2420170107

Smith, J., Fisher, G. G., Ryan, L. H., Clarke, P. J., House, J., & Weir, D. R. (2013). Psychosocial and lifestyle questionnaire 2006–2010. documentation report core section lb. https://hrs.isr.umich.edu/publications/biblio/8515

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-*

Cultural Psychology, *35(1)*, 50–61. https://doi.org/10.1177/0022022103260380

Sonnega, A., Faul, J. D., Ofstedal, M. B., Langa, K. M., Phillips, J. W., & Weir, D. R. (2014). Cohort profile: The health and retirement study (HRS). *International journal of epidemiology*, *43*(2), 576–585. 10.1093/ije/dyu067

Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, *53(1)*, 80–93. https://doi.org/10.1037/0022-0167.53.1.80

United Nations. (2010). World population ageing 2009. united nations publication, sales no. e. e.10.xiii.5. United Nations Department of Economic; Social Affairs, Population Division.

Van de Vijver, F. (1998). Towards a theory of bias and equivalence (J. Harkness, Ed.), 41–65. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49731-1

Van de Vijver, F., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, *54(2)*, 119–135. https://doi.org/10.1016/j.erap.2003.12.004

Van Vaerenbergh, Y., & Thomas, T. D. (2013). Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research*, *25(2)*, 195–217. https://doi.org/10.1093/ijpor/eds021

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology*, *54(6)*, 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Willis, G. B., & Gordon, B. (2005). *Cognitive interviewing: A tool for improving questionnaire design*. Sage Publications.

Appendix A
Question Wording of Selected Subjective Well-Being Measures

Single-Item Global Life Satisfaction

Please think about your life-as-a-whole. How satisfied are you with it? Are you completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied?
Response scale:
(a) Completely satisfied (b) Very satisfied (c) Somewhat satisfied (d) Not very satisfied (e) Not at all satisfied

Single-Item Momentary Life Satisfaction

Please think about your life and situation right now. How satisfied are you with your life as a whole these days?
Response scale:
(a) Completely satisfied (b) Very satisfied (c) Somewhat satisfied (d) Not very satisfied (e) Not at all satisfied

Satisfaction with Life (SWL)

Please say how much you agree or disagree with the following statements.
Q1. In most ways my life is close to ideal.
Q2. The conditions of my life are excellent.
Q3. I am satisfied with my life.
Q4. So far, I have gotten the important things I want in life.
Q5. If I could live my life again, I would change almost nothing.
Response scale:
(a) Strongly disagree (b) Somewhat disagree (c) Slightly disagree (d) Neither agree nor disagree (e) Slightly agree (f) Somewhat agree (g) Strongly agree

Positive Affect (PAF)

During the past 30 days, to what degree did you feel . . .
Q1. Determined?
Q2. Enthusiastic?
Q3. Active?
Q4. Proud?
Q5. Interested?
Q6. Happy?
Q7. Attentive?
Q8. Content?
Q9. Inspired?
Q10. Hopeful?
Q11. Alert?
Q12. Calm?
Q13. Excited?
Response scale:
(a) Very much (b) Quite a bit (c) Moderately (d) A little (e) Not at all

Purpose in Life (PIL)

Please say how much you agree or disagree with the following statements.
Q1. I enjoy making plans for the future and working to make them a reality.
Q2. My daily activities often seem trivial and unimportant to me. (-)
Q3. I am an active person in carrying out the plans I set for myself
Q4. I don't have a good sense of what it is I'm trying to accomplish in life. (-)
Q5. I sometimes feel as if I've done all there is to do in life. (-)
Q6. I live life one day at a time and don't really think about the future. (-)

Q7. I have a sense of direction and purpose in my life.
Response scale:
(a) Strongly disagree (b) Somewhat disagree (c) Slightly disagree (d) Slightly disagree (e) Somewhat agree (f) Strongly agree

Appendix B

Estimated Configural Invariance Model Parameters for the Multigroup Confirmatory Factor Analysis for Non-Hispanic Whites, Hispanics Interviewed in English and Hispanics Interviewed in Spanish, 2010 Health and Retirement Study.

Table B1
Standardized Factor Loadings and Standard Errors

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Loading	Std. Err.	Loading	Std. Err.	Loading	Std. Err.
SWL	Q1. CLOSETOIDEAL	0.870	0.004	0.752	0.025	0.680	0.035
	Q2. CONDEXCEL	0.891	0.004	0.827	0.020	0.733	0.031
	Q3. SATISFLIFE	0.885	0.004	0.904	0.014	0.880	0.019
	Q4. IMPORTTHING	0.744	0.007	0.824	0.020	0.846	0.022
	Q5. CHANGENOT	0.620	0.009	0.583	0.036	0.595	0.040
SWL[a]	Q1. CLOSETOIDEAL	0.835	0.003	0.845	0.012	0.857	0.013
	Q2. CONDEXCEL	0.843	0.003	0.855	0.011	0.864	0.012
	Q3. SATISFLIFE	0.923	0.004	0.855	0.018	0.766	0.026
	Q4. IMPORTTHING	0.784	0.005	0.805	0.015	0.811	0.017
	Q5. CHANGENOT	0.635	0.010	0.565	0.038	0.518	0.045
PAF	Q1. DETERMINED	0.612	0.009	0.544	0.038	0.633	0.039
	Q2. ENTHUSIASTIC	0.770	0.006	0.656	0.031	0.745	0.029
	Q3. ACTIVE	0.648	0.008	0.682	0.029	0.663	0.035
	Q4. PROUD	0.647	0.008	0.670	0.030	0.565	0.042
	Q5. INTERESTED	0.772	0.006	0.807	0.020	0.557	0.043
	Q6. HAPPY	0.756	0.006	0.688	0.029	0.646	0.036
	Q7. ATTENTIVE	0.651	0.008	0.663	0.031	0.586	0.041
	Q8. CONTENT	0.736	0.007	0.706	0.028	0.726	0.031
	Q9. INSPIRED	0.746	0.007	0.716	0.027	0.707	0.032
	Q10. HOPEFUL	0.781	0.006	0.712	0.027	0.666	0.035
	Q11. ALERT	0.685	0.008	0.709	0.027	0.599	0.040
	Q12. CALM	0.647	0.008	0.646	0.032	0.610	0.039
	Q13. EXCITED	0.669	0.008	0.700	0.028	0.715	0.031
PAF[b]	Q1. DETERMINED	0.588	0.010	0.503	0.040	0.603	0.041
	Q2. ENTHUSIASTIC	0.748	0.007	0.638	0.034	0.699	0.034
	Q3. ACTIVE	0.649	0.009	0.696	0.029	0.641	0.038

Continues on next page

Continued from previous page

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Loading	Std. Err.	Loading	Std. Err.	Loading	Std. Err.
	Q4. PROUD	0.630	0.009	0.653	0.032	0.541	0.045
	Q5. INTERESTED	0.768	0.006	0.801	0.021	0.535	0.045
	Q6. HAPPY	0.747	0.007	0.684	0.030	0.614	0.040
	Q7. ATTENTIVE	0.661	0.008	0.658	0.032	0.591	0.042
	Q8. CONTENT	0.729	0.007	0.698	0.029	0.714	0.032
	Q9. INSPIRED	0.723	0.007	0.692	0.030	0.706	0.034
	Q10. HOPEFUL	0.766	0.006	0.695	0.029	0.678	0.036
	Q11. ALERT	0.699	0.008	0.712	0.028	0.614	0.040
	Q12. CALM	0.659	0.008	0.660	0.031	0.624	0.039
	Q13. EXCITED	0.664	0.008	0.709	0.028	0.724	0.032
PIL	Q1. PLANFUTURE	0.639	0.010	0.486	0.049	0.234	0.067
	Q2. ACTTRIVIAL	−0.543	0.012	−0.545	0.043	−0.558	0.054
	Q3. OWNPLANS	0.611	0.011	0.460	0.050	0.209	0.068
	Q4. NOSENSE	−0.627	0.011	−0.763	0.033	−0.678	0.050
	Q5. DONEALL	−0.591	0.011	−0.647	0.038	−0.670	0.048
	Q6. ONEDAY	−0.489	0.012	−0.606	0.040	−0.601	0.051
	Q7. DIRECTION	0.646	0.010	0.445	0.051	0.156	0.067
PIL[b]	Q1. PLANFUTURE	0.700	0.010	0.716	0.038	0.646	0.075
	Q2. ACTTRIVIAL	−0.432	0.013	−0.284	0.054	−0.095	0.072
	Q3. OWNPLANS	0.686	0.010	0.722	0.038	0.631	0.073
	Q4. NOSENSE	−0.507	0.013	−0.399	0.051	−0.123	0.072
	Q5. DONEALL	−0.437	0.013	−0.318	0.053	−0.182	0.071
	Q6. ONEDAY	−0.385	0.014	−0.274	0.055	−0.283	0.070
	Q7. DIRECTION	0.695	0.010	0.674	0.039	0.451	0.067

SWL=Satisfaction with life; PAF=Positive affects; PIL= Purpose in life. [a] Modified with correlated residuals and fixed residual variances. [b] Modified with correlated residuals.

Table B2
Standardized Intercept Estimates and Standard Errors

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Intercept	Std. Err.	Intercept	Std. Err.	Intercept	Std. Err.
SWL	Q1. CLOSETOIDEAL	2.590	0.028	2.288	0.095	2.432	0.113
	Q2. CONDEXCEL	2.570	0.028	2.302	0.095	2.421	0.113
	Q3. SATISFLIFE	3.024	0.032	2.649	0.106	2.879	0.126
	Q4. IMPORTTHING	3.330	0.034	2.930	0.115	2.888	0.128

Continues on next page

Continued from previous page

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Intercept	Std. Err.	Intercept	Std. Err.	Intercept	Std. Err.
	Q5. CHANGENOT	2.186	0.025	2.193	0.092	2.414	0.110
SWL[a]	Q1. CLOSETOIDEAL	2.625	0.028	2.365	0.097	2.511	0.116
	Q2. CONDEXCEL	2.571	0.027	2.298	0.094	2.464	0.114
	Q3. SATISFLIFE	3.024	0.032	2.664	0.106	2.929	0.128
	Q4. IMPORTTHING	3.404	0.034	3.080	0.118	3.171	0.138
	Q5. CHANGENOT	2.187	0.025	2.202	0.092	2.436	0.111
PAF	Q1. DETERMINED	3.166	0.033	2.899	0.115	2.326	0.111
	Q2. ENTHUSIASTIC	3.046	0.032	2.652	0.107	2.381	0.110
	Q3. ACTIVE	3.028	0.032	2.751	0.109	2.731	0.123
	Q4. PROUD	3.016	0.032	2.949	0.116	1.973	0.097
	Q5. INTERESTED	3.757	0.038	3.225	0.125	1.909	0.096
	Q6. HAPPY	3.872	0.039	3.413	0.131	2.855	0.126
	Q7. ATTENTIVE	3.432	0.036	2.755	0.111	2.748	0.125
	Q8. CONTENT	3.321	0.034	2.850	0.113	2.877	0.127
	Q9. INSPIRED	2.786	0.030	2.495	0.103	2.097	0.101
	Q10. HOPEFUL	3.336	0.035	3.076	0.121	2.819	0.125
	Q11. ALERT	4.016	0.041	3.460	0.133	2.824	0.126
	Q12. CALM	3.609	0.037	3.160	0.123	2.820	0.126
	Q13. EXCITED	2.827	0.030	2.591	0.105	2.327	0.108
PAF[b]	Q1. DETERMINED	3.166	0.033	2.900	0.115	2.324	0.110
	Q2. ENTHUSIASTIC	3.044	0.032	2.669	0.107	2.386	0.110
	Q3. ACTIVE	3.028	0.032	2.751	0.109	2.728	0.123
	Q4. PROUD	3.017	0.032	2.954	0.116	1.975	0.097
	Q5. INTERESTED	3.758	0.038	3.224	0.125	1.911	0.096
	Q6. HAPPY	3.872	0.039	3.413	0.131	2.858	0.126
	Q7. ATTENTIVE	3.432	0.036	2.757	0.111	2.750	0.125
	Q8. CONTENT	3.322	0.034	2.854	0.113	2.876	0.127
	Q9. INSPIRED	2.788	0.030	2.500	0.103	2.098	0.101
	Q10. HOPEFUL	3.335	0.034	3.077	0.121	2.818	0.125
	Q11. ALERT	4.016	0.041	3.460	0.133	2.825	0.126
	Q12. CALM	3.609	0.037	3.162	0.123	2.821	0.126
	Q13. EXCITED	2.827	0.030	2.591	0.105	2.329	0.108
PIL	Q1. PLANFUTURE	−3.988	0.040	−3.547	0.135	−3.635	0.157
	Q2. ACTTRIVIAL	−1.792	0.022	−1.622	0.076	−1.494	0.083

Continues on next page

Continued from previous page

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Intercept	Std. Err.	Intercept	Std. Err.	Intercept	Std. Err.
	Q3. OWNPLANS	−3.863	0.039	−3.581	0.137	−4.595	0.194
	Q4. NOSENSE	−1.619	0.020	−1.558	0.075	−1.535	0.084
	Q5. DONEALL	−1.459	0.019	−1.459	0.072	−1.495	0.082
	Q6. ONEDAY	−1.623	0.020	−1.620	0.076	−1.512	0.083
	Q7. DIRECTION	−3.586	0.037	−3.582	0.137	−4.552	0.193
PIL[b]	Q1. PLANFUTURE	−3.987	0.040	−3.551	0.135	−3.634	0.157
	Q2. ACTTRIVIAL	−1.792	0.022	−1.620	0.076	−1.494	0.083
	Q3. OWNPLANS	−3.862	0.039	−3.581	0.137	−4.604	0.194
	Q4. NOSENSE	−1.619	0.020	−1.558	0.075	−1.537	0.084
	Q5. DONEALL	−1.459	0.019	−1.459	0.072	−1.495	0.082
	Q6. ONEDAY	−1.623	0.020	−1.620	0.076	−1.512	0.083
	Q7. DIRECTION	−3.586	0.037	−3.589	0.137	−4.534	0.194

SWL=Satisfaction with life; PAF=Positive affects; PIL= Purpose in life. [a] Modified with correlated residuals and fixed residual variances. [b] Modified with correlated residuals.

Table B3
Standardized Error Variance Estimates and Standard Errors

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Error Var	Std. Err.	Error Var	Std. Err.	Error Var	Std. Err.
SWL	Q1. CLOSETOIDEAL	0.243	0.007	0.435	0.038	0.538	0.048
	Q2. CONDEXCEL	0.207	0.007	0.316	0.033	0.462	0.045
	Q3. SATISFLIFE	0.216	0.007	0.183	0.026	0.226	0.034
	Q4. IMPORTTHING	0.446	0.010	0.321	0.033	0.285	0.036
	Q5. CHANGENOT	0.616	0.011	0.660	0.042	0.646	0.048
SWL[a]	Q1. CLOSETOIDEAL	0.302	0.006	0.287	0.020	0.266	0.022
	Q2. CONDEXCEL	0.290	0.005	0.269	0.019	0.254	0.020
	Q3. SATISFLIFE	0.385	0.007	0.352	0.024	0.342	0.027
	Q4. IMPORTTHING	0.148	0.008	0.268	0.031	0.413	0.040
	Q5. CHANGENOT	0.597	0.012	0.680	0.043	0.732	0.046
	Q1~Q2	0.357	0.012	0.026	0.061	−0.231	0.077
	Q3~Q4	0.001	0.023	0.335	0.046	0.383	0.040
	Q3~Q5	0.127	0.026	0.004	0.063	0.190	0.057
	Q4~Q5	0.140	0.015	0.194	0.047	0.202	0.046

Continues on next page

Continued from previous page

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Error Var	Std. Err.	Error Var	Std. Err.	Error Var	Std. Err.
PAF	Q1. DETERMINED	0.625	0.011	0.704	0.041	0.599	0.049
	Q2. ENTHUSIASTIC	0.407	0.009	0.570	0.041	0.445	0.043
	Q3. ACTIVE	0.580	0.011	0.535	0.040	0.560	0.047
	Q4. PROUD	0.581	0.011	0.551	0.040	0.681	0.048
	Q5. INTERESTED	0.404	0.009	0.348	0.033	0.690	0.048
	Q6. HAPPY	0.429	0.010	0.527	0.040	0.583	0.047
	Q7. ATTENTIVE	0.576	0.011	0.561	0.041	0.657	0.048
	Q8. CONTENT	0.458	0.010	0.501	0.039	0.472	0.044
	Q9. INSPIRED	0.443	0.010	0.488	0.039	0.501	0.046
	Q10. HOPEFUL	0.390	0.009	0.493	0.039	0.557	0.046
	Q11. ALERT	0.530	0.011	0.497	0.039	0.642	0.048
	Q12. CALM	0.582	0.011	0.582	0.041	0.628	0.048
	Q13. EXCITED	0.553	0.011	0.509	0.039	0.489	0.045
PAF[b]	Q1. DETERMINED	0.655	0.011	0.747	0.041	0.637	0.050
	Q2. ENTHUSIASTIC	0.440	0.010	0.593	0.043	0.512	0.047
	Q3. ACTIVE	0.578	0.011	0.515	0.041	0.589	0.049
	Q4. PROUD	0.603	0.011	0.574	0.042	0.708	0.048
	Q5. INTERESTED	0.410	0.010	0.358	0.034	0.714	0.048
	Q6. HAPPY	0.442	0.010	0.533	0.041	0.623	0.049
	Q7. ATTENTIVE	0.563	0.011	0.566	0.042	0.651	0.049
	Q8. CONTENT	0.468	0.010	0.512	0.040	0.490	0.046
	Q9. INSPIRED	0.477	0.011	0.521	0.041	0.502	0.048
	Q10. HOPEFUL	0.413	0.010	0.517	0.041	0.541	0.048
	Q11. ALERT	0.511	0.011	0.493	0.040	0.624	0.049
	Q12. CALM	0.566	0.011	0.564	0.041	0.611	0.049
	Q13. EXCITED	0.560	0.011	0.498	0.040	0.476	0.046
	Q4~Q5	0.074	0.013	0.071	0.048	0.222	0.054
	Q4~Q7	0.092	0.014	0.020	0.047	0.041	0.057
	Q4~Q11	0.370	0.012	0.568	0.036	0.457	0.048
	Q5~Q7	0.369	0.013	0.382	0.048	0.162	0.065
	Q5~Q11	0.290	0.014	0.201	0.053	0.460	0.049
	Q7~Q11	0.278	0.014	0.281	0.053	0.435	0.048
PIL	Q1. PLANFUTURE	0.592	0.013	0.764	0.047	0.945	0.031
	Q2. ACTTRIVIAL	0.706	0.013	0.703	0.046	0.689	0.060
	Q3. OWNPLANS	0.627	0.013	0.789	0.046	0.956	0.028

Continues on next page

Continued from previous page

		Non-Hispanic Whites (n=5,620)		Hispanics, English Interviews (n=406)		Hispanics, Spanish Interviews (n=330)	
		Error Var	Std. Err.	Error Var	Std. Err.	Error Var	Std. Err.
	Q4. NOSENSE	0.607	0.013	0.417	0.051	0.541	0.067
	Q5. DONEALL	0.651	0.013	0.581	0.049	0.552	0.064
	Q6. ONEDAY	0.761	0.012	0.633	0.048	0.639	0.062
	Q7. DIRECTION	0.582	0.013	0.802	0.045	0.976	0.021
PIL[b]	Q1. PLANFUTURE	0.510	0.014	0.488	0.055	0.583	0.096
	Q2. ACTTRIVIAL	0.814	0.012	0.919	0.031	0.991	0.014
	Q3. OWNPLANS	0.530	0.014	0.479	0.055	0.602	0.092
	Q4. NOSENSE	0.743	0.013	0.841	0.040	0.985	0.018
	Q5. DONEALL	0.809	0.012	0.899	0.034	0.967	0.026
	Q6. ONEDAY	0.852	0.011	0.925	0.030	0.920	0.039
	Q7. DIRECTION	0.517	0.014	0.545	0.053	0.796	0.060
	Q2~Q4	0.241	0.014	0.399	0.044	0.471	0.046
	Q2~Q5	0.262	0.013	0.274	0.048	0.331	0.052
	Q2~Q6	0.111	0.014	0.290	0.047	0.282	0.055
	Q4~Q5	0.294	0.013	0.473	0.040	0.445	0.047
	Q4~Q6	0.176	0.014	0.446	0.042	0.349	0.052
	Q5~Q6	0.293	0.013	0.402	0.043	0.422	0.049

SWL=Satisfaction with life; PAF=Positive affects; PIL= Purpose in life. [a] Modified with correlated residuals and fixed residual variances.
[b] Modified with correlated residuals.