NORC
at the UNIVERSITY of CHICAGO

NORC WORKING PAPER SERIES

Loneliness in Older Adults in the USA and Germany: Measurement Invariance and Validation

WP-2015-004 | FEBRUARY 8, 2016

PRESENTED BY:
NORC at the University of Chicago
55 East Monroe Street
30th Floor
Chicago, IL 60603
Ph. (312) 759-4000

AUTHORS:
Louise C. Hawkley, Rebeccah Duvoisin, Johannes Ackva
NORC at the University of Chicago

James C. Murdoch
2M Research Services, LLC

Maike Luhmann
University of Cologne

AUTHOR INFORMATION

Louise C. Hawkley[1]
NORC at the University of Chicago
1155 East 60th Street
Chicago IL 60637
Office: (773) 256-6214
hawkley-louise@norc.org

Rebeccah Duvoisin
NORC at the University of Chicago
1155 East 60th Street
Chicago IL 60637
(773) 256-6162
duvoisin-rebeccah@norc.org

Johannes Ackva
Adelphi
Caspar-Theyss-Strasse 14a
14193 Berlin
+49 (30) 8900068-493
jackva@gmail.com

James C. Murdoch
2M Research Services, LLC
500 E. Border St Ste 680
Arlington, TX 76010
(817) 631-3876
cmurdoch@2mresearch.com

Maike Luhmann
Herbert-Lewin-Str. 2
50931 Cologne
GERMANY
+49- (0) 221-470 86 07
maike.luhmann@uni-koeln.de

Key words: loneliness, older adults, aging

[1] Direct correspondence to the first author at hawkley-louise@norc.org.

Acknowledgments

The Health and Retirement Study is supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration.

The National Social Life, Health, and Aging Project is supported by the National Institute on Aging and the National Institutes of Health (R37AG030481; R01AG033903).

The German Socio-Economic Panel is funded through the Joint Science Conference by the German Federal Government and the State of Berlin.

Support for the development of this manuscript was provided by the National Institute on Aging and the National Institutes of Health (R37AG030481; R01AG033903), and by the Working Paper Series group at NORC at the University of Chicago.

Table of Contents

Table of Tables

Abstract

Loneliness is a subjective experience characterized by the distress that accompanies a perceived deficit in social relationship quantity and especially quality. Feelings of loneliness increase risk for morbidity and mortality, and these health costs are particularly evident in older age and as lonely feelings persist over time. The scope of the problem of loneliness in older age is not yet well understood, however. To accelerate the development of a cumulative knowledge base on loneliness, we harmonize data from three nationally representative samples—the Health and Retirement Study and the National Social Life, Health, and Aging Project in the United States, and the Socio-Economic Panel in Germany—to investigate the measurement invariance of a 3-item loneliness measure and validate the loneliness construct across surveys. A series of multi-group confirmatory factor analyses provided support for strict measurement invariance. Validation was demonstrated with equivalent inverse associations of loneliness with self-rated health, marital status, and frequency of social activity across surveys and countries. We articulate a method to harmonize data across national surveys and encourage replication and extension to expand the scope of loneliness research.

Loneliness is a subjective experience that is distinct from social isolation per se (e.g., living alone, having a small social network) and is characterized by the distress that accompanies the perceived absence or deficit of satisfying social relationships (Peplau & Perlman, 1982). In essence, loneliness represents a mismatch between an individual's social needs and the provisions the social environment offers or is perceived to offer. Loneliness is prevalent among older adults in developed countries, although prevalences differ widely across countries. When asked directly, over 8% of older adults in the Netherlands and Germany, 10% in Sweden, almost 16% in Greece, Spain, and the United States of America (USA), and as many as 25% of older adults in Italy endorse feeling lonely (Fokkema, De Jong Gierveld, & Dykstra, 2012; Shiovitz-Ezra & Ayalon, 2012).

Loneliness increases risk for morbidity and mortality (Hawkley & Cacioppo, 2010; Luo, Hawkley, Waite, & Cacioppo, 2012; Perissinotto, Cenzer, & Covinsky, 2012), with mortality risk increased by 26 percent according to a recent meta-analysis (Holt-Lunstad, Smith, Baker, Harris, & Stephenson, 2015). The health costs of loneliness are particularly evident in older age and as lonely feelings persist over time (Hawkley & Capitanio, 2015). To the extent that loneliness increases in prevalence, intensity, and chronicity with age, rapidly aging populations raise the specter of a growing social, economic, and health burden for individuals and societies.

The scope of the problem of loneliness in older age is not yet well understood, however. In recognition of the public health risk posed by loneliness, national population-based surveys are increasingly including a standardized measure of loneliness in their survey instruments to help advance research on this topic. Research would benefit from combining data from various national surveys to test the replicability and generalizability of findings, increase sample heterogeneity and statistical power to facilitate comparisons among population subgroups, and broaden the psychometric assessment of the loneliness construct to show the strength of its measurement properties (Curran & Hussong, 2009). In this study, we combine data across three national surveys and two countries—USA and Germany—to explicitly examine the measurement qualities of a loneliness measure in older adults, and to test whether the measure exhibits comparable associations with criterion validation variables across all surveys.

To date, most national surveys have assessed loneliness with a single face-valid item that asks people the frequency with which they feel lonely or loneliness of a particular intensity. This measure has been criticized because it is prone to response bias (Borys & Perlman, 1985; Shiovitz-Ezra & Ayalon, 2012). Multi-item validated loneliness measures that avoid the potentially stigmatizing terms "lonely" and "loneliness" were introduced in Europe (De Jong Gierveld Loneliness Scale) and the USA (UCLA Loneliness Scale) in the 1980s (De Jong Gierveld & Kamphuis, 1985; Russell, Peplau, & Cutrona, 1980),

but until the early 2000s, these measures were not included in population-based national surveys. A reduced 3-item version of the UCLA Loneliness Scale is included in two USA surveys—the Health and Retirement Study (HRS), and the National Social Life, Health, and Aging Project (NSHAP)—and recently in the German Socio-Economic Panel (SOEP). We integrate and harmonize the data where necessary by creating new variables that are consistent across surveys in order to evaluate the measurement invariance of the multi-item loneliness measure. Measurement invariance is necessary to infer that the construct of loneliness has a comparable meaning across populations and also that comparisons of loneliness severity among populations are meaningful.

To establish the validity of the measure, we examine correlations of loneliness with health, marital status, and social activity. Poor physical health is robustly associated with higher levels of loneliness (Ó Luanaigh & Lawlor, 2008). Conversely, meta-analyses have shown that marriage protects against loneliness (Dykstra & De Jong Gierveld, 2004; Pinquart & Sörensen, 2003), an effect that is comparable across many countries of the world (Stack, 1998). Similarly, more frequent social activity is associated with lower levels of loneliness (Pinquart & Sörensen, 2003). We test whether these correlations are of similar magnitude across surveys and countries. Generalizability would indicate that these associations reflect fundamentals of human social behavior (Burkhauser & Lillard, 2005).

Methods

Samples

The sample was constructed using data from two USA surveys—the 2010 waves of HRS and NSHAP—as well as the 2011 Innovation Sample of the German SOEP. Each is a longitudinal panel survey with a nationally representative sample. To focus on older adults, only respondents 50 years and older were included in the analyses.

HRS. The HRS is sponsored by the National Institute on Aging and is conducted by the University of Michigan. The HRS is a longitudinal panel study that began in 1992. A multi-stage, clustered area probability frame was used to select households, and age-eligible individuals and their spouses or partners, regardless of year of birth, were eligible to participate. The survey's sample is composed of six birth cohorts that entered the study in different calendar years. Black and Hispanic respondents were oversampled in most cohorts (Health and Retirement Study, 2014). Initial response rates in HRS ranged from 70 to slightly over 80 percent, and the rate for re-interviewing has been between 92 and 95 percent (Health and Retirement Study, 2011). Currently, more than 26,000 Americans are interviewed every two

years. At each interview, detailed information is collected about the respondent's health, family relationships, employment, income and wealth, and demographic background. In 2010, a random one-half of households was pre-selected to receive the leave-behind questionnaire of which the loneliness items are a part, with the other half of the sample selected for 2012 (Health and Retirement Study, 2014). Among those who were eligible for and received the leave-behind questionnaire in 2010 (N=10,744), 8,242 cases returned the questionnaire, indicating a 76.7% return rate overall. Using the subsample of respondents 50 years and older who received the leave-behind questionnaire (N=7,962), 7,917 cases provided at least partial loneliness data, and 7,775 provided complete loneliness data (i.e., values for all three loneliness items).

NSHAP. The NSHAP is funded by the National Institute on Aging and is conducted by the NORC at the University of Chicago. NSHAP is a population-based study of health, social life, and well-being among older Americans. In Wave 1, a national area probability sample of households, carried out jointly with the 2004 round of the HRS, was selected from households across the USA. One age-eligible individual in each household was randomly selected for recruitment. The focal age range was 57-85 years. Selection of respondents for NSHAP balanced age and gender subgroups and oversampled Black and Hispanics (O'Muircheartaigh, Ekman, & Smith, 2009). In-home interviews were conducted in English and Spanish between July 2005 and March 2006, yielding a total of 3,005 respondents (1,455 men and 1,550 women). Data for the present study come from a second wave conducted from August 2010 through May 2011, during which Wave 1 respondents were re-interviewed. An attempt was also made to interview individuals who were sampled in Wave 1 but declined to participate. In addition, spouses or co-resident partners were also interviewed using the same instruments as the main respondent. The resulting Wave 2 dataset contains 3,377 total respondents, including 957 couples. The overall unconditional response rate for Wave 2 was 74%; the conditional response rate of Wave 1 respondents was 89%; the conditional response rate of partners was 84%; and the conversion rate for Wave1 non-respondents was 26% (Waite et al., 2014). All respondents received a leave-behind questionnaire, of which the loneliness items were a part. The leave-behind questionnaire was returned by 88.6% of respondents (Hawkley, Kocherginsky, Wong, Kim, & Cagney, 2014). Using the subsample of Wave 2 respondents who were 50 years and older and returned the leave-behind questionnaire (N=3,363), 2,854 cases provided at least partial loneliness data, and 2,693 provided complete loneliness data.

SOEP. The SOEP is funded by the German Federal Government and the State of Berlin and is managed by the German Institute for Economic Research, DIW Berlin. The SOEP is a nationally representative longitudinal study of private households in Germany, and all adults age 16 and over in these households are invited to participate. The SOEP was started in the Federal Republic of Germany in 1984, and added

the German Democratic Republic in 1990. With successive additions to the panel, the SOEP now surveys almost 11,000 households annually. Initial response rates for the SOEP ranged from about 60 to 70 percent (depending on the subsample), and the response rate has exceeded 90 percent at follow-ups (Haisken-DeNew & Frick, 2005). The SOEP provides a wide range of data for research in the social, economic, and behavioral sciences. In 2011, a nationally representative sample of 1,008 households was recruited for an "Innovation Sample," and 1,625 household members age 16-92 years were tested (Richter & Schupp, 2012). The loneliness items were part of the in-person interview; using the subsample of respondents 50 years and older (N=940), 936 cases provided at least partial loneliness data, and 927 provided complete loneliness data.

Measures

Loneliness

All three surveys based their loneliness measure on the well-validated 20-item UCLA Loneliness Scale (Russell, 1996). The UCLA Loneliness Scale does not ask direct questions about feelings of loneliness and instead assesses loneliness indirectly, a strategy that minimizes the stigma and associated response bias elicited by a direct approach (Borys & Perlman, 1985; Shiovitz-Ezra & Ayalon, 2012). The UCLA Loneliness Scale has been shown to be comprised of three dimensions: an Intimate Connectedness facet, reflecting experiences of social value as an individual; a Relational Connectedness facet, reflecting experiences of social value in dyadic friendships; and a Collective Connectedness facet, reflecting experiences of collective identity and belonging to a group (Hawkley, Browne, & Cacioppo, 2005; Hawkley, Gu, Luo, & Cacioppo, 2012). The three facets are highly correlated but exhibit discriminant validity in correlations with criterion variables. The first facet captures the largest proportion of the variance in the latent loneliness factor and reflects the social isolation that best typifies loneliness as conceived by individual differences researchers.

The HRS was the first large national USA survey to introduce and validate a standardized loneliness measure. Three items from the first facet were chosen to represent loneliness: "How often do you feel… (1) left out, (2) isolated, (3) that you lack companionship?" Respondents are provided with a 3-point response scale, from 1 (often), 2 (sometimes), to 3 (rarely ever or never). Responses are averaged to generate a loneliness score. The three-item scale was validated using pilot data from the 2002 wave of data collection in the HRS by demonstrating correlations with criterion variables that were comparable in magnitude to those obtained in the Chicago Health, Aging, and Social Relations Study that used the validated 20-item scale and a 4-point response scale (Hughes, Waite, Hawkley, & Cacioppo, 2004). An 11-item version of the loneliness scale is now routinely administered to HRS respondents in a leave-

behind packet. For consistency across surveys, only the validated three-item version is used in the present study.

The same three-item scale is administered as a leave-behind questionnaire in the NSHAP, but in Wave 2 (2010), respondents were provided with a 4-point response scale: 1 (never), 2 (hardly ever), 3 (some of the time), and 4 (often). For implementation in the SOEP, the three items were translated into German by a native German speaker and, for validation, back-translated into English by a native English speaker. In the SOEP, the three-item scale was administered using a computer-assisted personal interview (CAPI) methodology and a 4-point response scale comparable to that used in NSHAP, but reversed in direction: 1 (often), 2 (sometimes), 3 (rarely), and 4 (never). These items were harmonized across surveys by collapsing the two infrequent categories of the 4-point response scale and by reversing the direction of responses where necessary so that higher values indicate greater loneliness frequency (see Table 1). The distribution of responses to each item is displayed in Table 2.

Validation Variables

Following the example set by Bath et al. (Bath, Deeg, & Poppelaars, 2010), and to facilitate replication by other researchers, Table 1 identifies the source variables in each survey used to generate harmonized measures of health, marital status, and social activity, and summarizes how harmonization was accomplished.

Health. In each survey, respondents were asked to rate their health on a scale from 1 to 5. Differences in the direction of the response scale (poor to excellent versus excellent to poor) were recoded so that higher values indicated better health status.

Marital status. Three dummy variables were created to compare individuals who were (1) married or living with a partner in a stable relationship (*married or live-in partner*), (2) *divorced*, or (3) *widowed*, with individuals in other marital status groups (reference group).

Social activity. Socializing frequency was measured in categories that were harmonized across surveys by creating a dichotomous variable to compare weekly or more frequent socializing with less frequent socializing.

Table 1. Harmonization of data across HSR, NSHAP, and SOEP surveys.

Domain	Question	Variable names and response categories			New Variables	New Response Categories
		HRS	NSHAP	SOEP		
Loneliness	"How often do you feel… (1) left out, (2) isolated, (3) that you lack companionship?"	*MLB020A, MLB020B, MLB020C*: 1=often, 2=sometimes, 3=hardly ever or never. MODE: leave-behind questionnaire.	*leftout2, isolated2, companion2*: 1=never, 2=hardly ever, 3=some of the time, 4=often. MODE: leave-behind questionnaire.	*plh0268, plh0269, plh0270*: 1=often, 2=sometimes, 3=rarely, 4=never. MODE: computer-assisted personal interview.	leftout, isolated, companion	1=rarely or never, 2=sometimes, 3=often
Health	Self-rated health	*MC001* (physical health): 1=very good, 2=good, 3=satisfactory, 4=poor, 5=bad.	*health* (physical health): 1=poor, 2=fair, 3=good, 4=very good, 5=excellent.	*ple0008* (current health): 1=very good, 2=good, 3=satisfactory, 4=poor, 5=bad.	health	1=poor to 5=excellent
Social roles	Marital status	*MB063* (marital status): married, annulled, separated, divorced, widowed, never married. *MLB004* (live with spouse or partner): yes/no.	*maritlst* (marital status): married, living with a partner, separated, divorced, widowed, never married.	*pld0131* (marital status): 1=married, living together, 2=married, living apart, 3=unmarried, 4=divorced, 5=widowed. *pld0133* (living with partner in household): yes/no.	(1) married or living with partner, (2) divorced, (3) widowed. Ref=all other categories.	0=no, 1=yes
Social activity	Frequency of socializing with friends, relatives or neighbors	*Variable:* MLB013A (meet up with family members not living with you), MLB017A (meet up with friends). *Six categories:* 1=3 or more times/week, 2=once or twice/week, 3=once or twice/month, 4=every few months, 5=once or twice/year, 6=less than once a year or never. Take maximum of family or friend socializing.	*Variable:* social (frequency of socializing with friends or relatives in past year). *Seven response categories:* 0=never, 1=less than once/year, 2=about once or twice/year, 3=several times/year, 4=about once a month, 5=every week, 6=several times/week.	*pli0094* (meet with friends, relatives or neighbors): 1=never, 2=rarely, 3=every month, 4=every week.	socialize weekly or more	0=no, 1=yes

Table 2. Means, percentages, and standard deviations of harmonized variables in HRS, NSHAP, and SOEP.[1]

Variable	HRS 2010	NSHAP 2010	SOEP 2011
Age (years)	67.3 (SD=10.8) range=50-101	72.3 (SD=7.7) range=50-99	64.7 (SD=9.3) range = 50-92
Female (%)	57.6 (49.4)	54.4 (SD=49.8)	51.3 (SD=50.0)
Self-rated health (range=1-5)	3.19 (1.08)	3.24 (SD=1.05)	3.15 (SD=0.98)
Marital status			
Married or living with partner (%) Divorced Widowed	66.0 (SD=47.4) 13.3 (SD=33.9) 17.7 (SD=38.1)	72.2 (SD=44.8) 7.1 (SD=25.8) 17.8 (SD=38.2)	74.7 (SD=43.5) 11.8 (SD=32.2) 13.0 (SD=33.7)
Socialize with friends, family, neighbors			
At least monthly (%) At least weekly (%)	78.5 (SD=41.1) 52.0 (SD=50.0)	73.8 (SD=44.0) 51.6 (SD=50.0)	74.3 (SD=43.7) 30.7 (SD=46.1)
Loneliness item: Lack companionship			
Never	55.0	61.6	33.0	70.7	41.2
Hardly ever/rarely			28.6		29.5
Sometimes	34.1	29.4		23.2	
Often	10.9	9.0		6.1	
Loneliness item: Left out					
Never	58.0	69.6	28.7	80.6	49.0
Rarely			40.9		31.6
Sometimes	35.2	26.8		14.0	
Often	6.8	3.7		5.4	
Loneliness item: Isolated					
Never	66.6	73.8	38.3	86.5	63.1
Rarely			35.5		23.4
Sometimes	27.0	23.2		8.8	
Often	6.4	3.1		4.7	

[1] Statistics are based on cases with at least partial loneliness data (see text); within surveys, cell sizes vary depending on rates of missing data for each variable.

Statistical Analysis

Multi-group confirmatory factor analysis was used to evaluate the equivalence of the harmonized loneliness measure across surveys. A set of nested models was conducted to test for configural invariance (i.e., same number of factors and a similar pattern of loadings), weak invariance (i.e., equal loadings), strong invariance (i.e., equal intercepts), and strict invariance (i.e., equal residuals) (Ender, 2013). For the tests of measurement invariance, model identification was achieved by assigning the "left-out" item a fixed intercept of 0, a fixed loading of 1 on the latent loneliness factor, and a residual variance of 0. *Configural* invariance signifies that the same construct is measured across groups, and was modeled by constraining the model identification parameters to equality across surveys while leaving all other parameters free to vary. *Weak* invariance implies that the loneliness factor has the same meaning across groups. Specifically, weak invariance indicates that a one-unit increase in an item score corresponds to an equal unit change in the latent loneliness factor across groups, a relationship that is necessary for valid comparisons of correlates of loneliness across surveys and countries. Weak invariance was modeled by constraining the "left-out" item intercept to 0 and its residual variance to 0, and constraining the latent loneliness factor loading to 1 across all groups. *Strong* invariance indicates that not only factor loadings but also intercepts are equal across groups so that the latent loneliness factor is scaled identically across groups; this permits loneliness scores to be compared across surveys. To test the assumptions of the strong invariance (also known as scalar invariance) requirement, the intercepts (or means) of the "isolated" and "companion" items are constrained to equality. *Strict* invariance is necessary to ensure that residual variances are attributable to measurement reliability and not to bias; unequal error variances suggest that an unmodeled source of variance is systematically captured by item responses that may bias test scores differentially across groups (Wu, Li, & Zumbo, 2007). To test the assumptions of strict invariance, the residuals for the "isolated" and "companion" items are constrained to equality across groups.

Several model fit indices were obtained, including χ^2, RMSEA (Root Mean Square Error of Approximation), and CFI (Comparative Fit Index). Values of RMSEA $\leq$0.08 are typically used to indicate reasonable model fit and <=0.05 as good model fit. We note, however, that when the degrees of freedom are small, RMSEA values can be quite large even when the model fits well (Kenny, Kaniskan, & McCoach, 2014). We therefore preferentially evaluate model fit using values of CFI over 0.90 as indicative of acceptable fit and values over 0.95 as good model fit (Hu & Bentler, 1999; Ortuño-Sierra et al., 2013). To compare models, we report results of the likelihood ratio test (i.e., difference in chi-squared), but we note that this test is overly sensitive to large sample sizes and may too readily reject the hypothesis of equivalent fit. We therefore preferentially use criteria recommended by Wu et al. (2007),

namely ΔCFI, where a ΔCFI less than 0.01 is used to indicate measurement invariance (Cheung & Rensvold, 2002; Dimitrov, 2010).

Correlational analyses were conducted separately for each survey to examine associations of loneliness with validation variables. A fully interactive regression model was run using the integrated data to test whether the associations of the validation variables with loneliness differ across surveys. All analyses were conducted with Stata, version 13 (StataCorp, 2013).

Results

Measurement Invariance

Initial separate principal factor analyses revealed high item loadings (≥0.60) on a single loneliness factor in each survey (see Table 3). These findings support the assertion that the items are good indicators of the latent loneliness factor in each survey. Table 4 displays model fit for each successively stringent test of invariance. The configural invariance model (Model 1) exhibited poor model fit using the RMSEA criterion, but the CFI (0.942) indicated adequate fit of the model to the data. We therefore proceeded with a test of weak invariance (i.e., equal loadings across groups). In Model 2, Table 4, the modest ΔCFI of -0.002 indicates support for weak invariance. Model 3 shows a minimal change in the CFI that again did not exceed -0.01, indicating support for strong invariance. Finally, in Model 4, the ΔCFI of -0.009 indicates support for strict invariance. The measurement invariance of the loneliness factor therefore permits valid cross-survey comparison of associations with loneliness and differences in mean loneliness.

Table 3. Standardized factor loadings from principal factor analyses conducted separately by group.

Item	HRS	NSHAP	SOEP
Feel left out	0.79	0.78	0.74
Lack companionship	0.68	0.65	0.60
Feel isolated	0.75	0.73	0.74

Table 4. Tests of loneliness measurement invariance across HRS, NSHAP, and SOEP surveys.

Model #	Invariance Model	X^2	*df*	ΔX^2	Δdf	*p*	RMSEA	CFI	ΔCFI
1	*Configural* (equal number of factors and patterns of loadings)	667.79	3	-	-	-	0.242	0.942	-
2	*Weak* (equal loadings)	694.86	7	13.54	4	0.009	0.161	0.940	0.002
3	*Strong* (equal intercepts)	725.68	11	30.82	4	<0.001	0.131	0.938	0.002
4	*Strict* (equal residuals)	834.89	15	109.21	4	<0.001	0.120	0.929	0.009

For subsequent analyses, loneliness scores were calculated by averaging responses across all items for which data were available. Higher scores represent greater loneliness. The Cronbach's alpha for the three-item loneliness scale was 0.81 in HRS, 0.79 in NSHAP, and 0.77 in SOEP. Mean values for loneliness (range=1-3) were highest in HRS at 1.48 (*SD*=0.54), followed by 1.37 (*SD*=0.48) in NSHAP, and 1.26 (*SD*=0.45) in SOEP. All group differences were significant at $p<.01$.

Validation

Correlations of loneliness with the validation variables are displayed in Table 5. The general pattern of correlations is comparable across surveys and replicates prior research in showing that loneliness is inversely correlated with self-rated health, *r*'s = -0.18 to -0.24, being married or partnered (versus being unmarried), *r*'s = -0.14 to -0.23, and socializing weekly or more often, *r*'s = -0.07 to -0.10; and is positively correlated with being divorced or widowed, *r*'s = 0.09 to 0.16. The one exception was a near-zero correlation between widowhood and loneliness in the SOEP sample, $r = 0.01$. Interaction terms were calculated between these variables and the HRS and NSHAP samples and employed in a regression model that adjusted for the main effects of survey (reference group=SOEP), marital status, and social activity. This model revealed no significant interactions, indicating that the associations between loneliness and these validation variables is equivalent in magnitude across surveys and countries.

Table 5. Correlations of loneliness with covariates in HRS, NSHAP, and SOEP.

	HRS 2010[1] N=7,101	NSHAP 2010[2] N=2,742	SOEP 2011[3] N=756
Self-rated health	-0.23	-0.18	-0.24
Married or living with partner	-0.23	-0.22	-0.14
Divorced	0.11	0.09	0.12
Widowed	0.13	0.16	0.01
Socialize weekly or more	-0.09	-0.10	-0.07

[1] With sample size of 7,101, correlations $\geq|0.025|$ are significant at $p<.05$, two-tailed.
[2] With sample size of 2,742, correlations $\geq|0.038|$ are significant at $p<.05$, two-tailed.
[3] With sample size of 756, correlations $\geq|0.072|$ are significant at $p<.05$, two-tailed.

Discussion

This is the first study to explicitly examine the measurement qualities of a three-item loneliness scale that is finding its way into a growing number of national surveys. One of these surveys is the German SOEP, and the present study serves to validate its use in that context among adults 50 years and older. Results provide evidence that this very brief loneliness measure captures a single loneliness construct and exhibits strict measurement invariance across surveys. The surveys differ in the language and country of origin, in the response scales offered to respondents, and in the mode of administration. None of these differences appeared to jeopardize measurement invariance in an older adult population.

Strict measurement invariance permits a valid and meaningful comparison of mean loneliness levels across countries and surveys. Results revealed higher levels of loneliness in the HRS than the NSHAP and SOEP surveys. The fact that NSHAP and SOEP used a 4-point response scale that was subsequently collapsed to three categories raises the question whether this contributed to loneliness differences. For instance, are respondents more likely to endorse higher values on the loneliness items if, instead of four response options, they have only three, as was the case in HRS? Research has shown that respondents tend to respond in the middle of the scale range because they assume that only extreme circumstances warrant using the extreme ends of the scale (Schwarz, 1990). This fact would suggest that respondents in HRS, who were administered a scale with only a 3-point range should have lower loneliness values on average than respondents in NSHAP and SOEP who received a 4-point response scale and could respond more extremely without endorsing the most extreme (i.e., most lonely) end of the scale. Item-level responses are consistent with this conjecture, but only in the prevalence of low frequency responses; as shown in Table 2, respondents in HRS were less likely to endorse the extreme, “never,” but *more* likely to

endorse the extreme, “often,” than respondents in NSHAP and SOEP. However, the scale administered to HRS respondents was also directionally opposite to that administered to NSHAP and SOEP respondents; specifically, the HRS scale ran from often (“1”) to never (“3”), whereas the NSHAP and SOEP scales ran from never (“1”) to often (“3”). Although reversing the direction of the scale may not in itself affect responses, the intuitive incongruence between the numerical value and its label (1=often, rather than 1=never) has been shown to significantly affect means and variances (Rammstedt & Krebs, 2007). This phenomenon may help to explain why HRS respondents were more likely to endorse frequent (i.e., often) than infrequent isolation experiences; a value of “1” may have intuitively signaled low frequency experiences for at least some respondents. Also consistent with the incongruence hypothesis, mean loneliness values for the HRS sample had greater variance despite the fact that the HRS sample was larger than the NSHAP and SOEP samples. Whether this is attributable to response scale differences requires additional empirical testing.

Despite the similar response scale in NSHAP and SOEP, loneliness levels in SOEP were lower than in NSHAP. Differential return rates and item nonresponses in the leave-behind questionnaire are a concern, but in this case, lonelier individuals in NSHAP were less likely to return the leave-behind questionnaire, as suggested by lower return rates from separated, divorced, and widowed respondents (Hawkley et al., 2014). Thus, any discrepancy between the NSHAP and SOEP loneliness scores is underestimated in this study. The SOEP was the only survey to administer the items in a CAPI format rather than in a leave-behind questionnaire, and this could have contributed to an underreporting of loneliness since respondents may be less likely to admit feeling left out and isolated and lacking companionship when asked in person. However, using data from the World Values Survey, an explicit question about loneliness revealed that Germans reported feeling very lonely less frequently than those in the USA, even though respondents in both countries were interviewed by telephone (Van Tilburg & Dykstra, 2008). Thus, it seems likely that the loneliness difference between NSHAP and SOEP represents a difference between countries, not a difference between modes of questionnaire administration. Identification of cultural and policy factors that contribute to the country difference in loneliness intensity could be useful in understanding how to alleviate the burden of loneliness in more affected countries.

Associations between loneliness and criterion variables provided evidence that the three-item scale is a valid measure of the loneliness construct. This is particularly relevant for the SOEP survey as present evidence constitutes validation of the recently introduced measure in the German survey context. Self-rated health and loneliness have been reliably inversely correlated in prior research (Hawkley & Capitanio, 2015; Nummela, Seppänen, & Uutela, 2011), and results reported here replicate and extend prior work by showing that the magnitude of the association is equivalent in American and German

samples. The protective effect of marriage is well documented and generalizes across diverse nations included in the World Values Survey (Stack, 1998). We, too, found a sizeable association between being married and loneliness in each sample and country and, moreover, that the association was of a similar magnitude across countries. The protective effect of marriage may be qualified by the quality of the marriage, however. Prior research has shown that marriage is protective against loneliness only to the extent that the marital partner is a confidant or, in other words, a close relationship partner (Hawkley et al., 2008). Similarly, frequent socializing, although protective, is no guarantee against loneliness because not all socializing is desirable. Relationships with critical or demanding others, or interactions that fall short of the desired quality, may even exacerbate feelings of isolation and loneliness (Hawkley et al., 2008). Relationship quality, whether with a spouse or the broader social network, is information worth obtaining to better understand how the social context is translated into feelings of loneliness. Available data did not permit testing the impact of marital quality on risk for loneliness across surveys because only NSHAP and HRS, and not SOEP, collect data on marital closeness, happiness, or satisfaction.

The present study indicates that data harmonization and measure validation across national surveys can be used to good effect to increase understanding of the universal and culture-specific predictors of loneliness. This is increasingly relevant in rapidly aging societies that share a concern for a growing older adult population that is particularly vulnerable to loneliness and its health consequences. Data harmonization is not without challenges, however. For instance, the loneliness items were identical in content, but the response scale and survey mode differed across surveys. The social activity variable differed in content and response scale options, and information was lost in collapsing different response categories to a dichotomous measure that could be considered equivalent across surveys. Even then, equivalence can be questioned given that dichotomization may result in variables that do not behave comparably if they arise from different response scales. This is a measurement issue worthy of study in its own right. Where questions are placed in the survey may also affect results and should be considered. In addition, survey differences not examined here—including differences in sample, study design, and response rates—could affect how well results generalize to the wider populations of American and German older adults (Bath, Deeg, & Poppelaars, 2010).

Measurement equivalence across surveys in the present study indicates that efforts to harmonize loneliness data are worthwhile. Several other national surveys are using the three-item or a longer version of the original 20-item UCLA Loneliness Scale, including the English Longitudinal Study of Ageing, the Irish Longitudinal Study of Ageing, and other national surveys in the HRS family of studies. Methods described here are offered to assist other researchers in extending the present work to include other countries and surveys, with the caveat that extensions to other surveys should include a formal evaluation

of the measurement equivalence of the scale in different contexts and consideration of other methodological differences among surveys. Vast amounts of data—surveys, cases, and variables—are already available; with no data collection costs, resources can be applied directly to data compilation, preparation, and analysis. This represents a unique opportunity to conduct cross-country research to examine whether loneliness is experienced comparably across a variety of political and social contexts and whether country differences in contextual factors that explain loneliness might have policy implications that can be developed and implemented for the good of older adults everywhere.

References

Bath, P. A., Deeg, D., & Poppelaars, J. (2010). The harmonisation of longitudinal data: A case study using data from cohort studies in The Netherlands and the United Kingdom. *Ageing & Society, 30,* 1419-1437.

Borys, S., & Perlman, D. (1985). Gender differences in loneliness. *Personality and Social Psychology Bulletin, 11,* 63-74.

Burkhauser, R. B., & Lillard, D. R. (2005). The contribution and potential of data harmonization for cross-national comparative research. *Journal of Comparative Policy Analysis: Research and Practice, 7(4),* 313-330.

Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management, 25,* 1-27.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233-255.

Curran, P. J., & Hussong, A. M. (2009). Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods, 14,* 81-100.

De Jong Gierveld, J., & Kamphuis, F. (1985). The development of a Rasch-type loneliness scale. *Applied Psychological Measurement, 9,* 289-299.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43,* 121-149.

Dykstra, P. A., & De Jong Gierveld, J. (2004). Gender and marital-history differences in emotional and social loneliness among Dutch older adults. *Canadian Journal on Aging, 23,* 141-155.

Ender, P. (2013). The hierarchy of factor invariance. Retrieved December 15, 2014. http://www.stata.com/meeting/new-orleans13/abstracts/materials/nola13-ender.pdf.

Fokkema, T., De Jong Gierveld, J., & Dykstra, P. A. (2012). Cross-national differences in older adult loneliness. *Journal of Psychology, 146,* 201-228.

Haisken-DeNew, J. P., & Frick, J.R., Eds. (2005). DTC Desktop Companion to the German Socio-Economic Panel (SOEP), Version 8.0 - Dec 2005. Updated to Wave 21 (U). Retrieved December 19, 2014. http://www.diw.de/documents/dokumentenarchiv/17/diw_01.c.38951.de/dtc.409713.pdf#page=15.

Hawkley, L. C., Browne, M. W., & Cacioppo, J. T. (2005). How can I connect with thee? Let me count the ways. *Psychological Science, 16,* 798-804.

Hawkley, L. C., & Cacioppo, J. T. (2010). Loneliness matters: A theoretical and empirical review of consequences and mechanisms. *Annals of Behavioral Medicine, 40,* 218-227. PMC Journal – In process.

Hawkley, L. C., & Capitanio, J. P. (2015). Perceived social isolation, evolutionary fitness, and health outcomes: A lifespan approach. *Philosophical Transactions of the Royal Society B, 370,* 2014114.

Hawkley, L. C., Gu, Y., Luo, Y.-J., & Cacioppo, J. T. (2012). The mental representation of social connections: Generalizability extended to Beijing adults. *PLoS ONE 7(9),* e44065. doi:10.1371/journal.pone.0044065.

Hawkley, L. C., Hughes, M. E., Waite, L. J., Masi, C. M., Thisted, R. A., & Cacioppo, J. T. (2008). From social structural factors to perceptions of relationship quality and loneliness: The Chicago health, aging, and social relations study. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 63B,* S375-S384.

Hawkley, L. C., Kocherginsky, M., Wong, J., Kim, J., & Cagney, K. A. (2014). Missing data in Wave 2 of NSHAP: Prevalence, predictors, and recommended treatment. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 69(S2),* S38-S50.

Health and Retirement Study. (2011). Sample sizes and response rates (2002 and beyond). Retrieved January 29, 2014. http://hrsonline.isr.umich.edu/sitedocs/sampleresponse.pdf.

Health and Retirement Study. (2014). 2010 Core Final, Version 5.0. Data Description and Usage. Retrieved December 8, 2014. http://hrsonline.isr.umich.edu/modules/meta/2010/core/desc/h10dd.pdf.

Holt-Lunstad, J., Smith, T. B., Baker, M., Harris, T., & Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: A meta-analytic review. *Perspectives on Psychological Science, 10,* 227-237.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6(1),* 1-55.

Hughes, M. E., Waite, L. J., Hawkley, L. C., & Cacioppo, J. T. (2004). A short scale for measuring loneliness in large surveys: Results from two population-based studies. *Research on Aging, 26,* 655-672.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research.* Online early.

Luo, Y., Hawkley, L. C., Waite, L. J., & Cacioppo, J. T. (2012). Loneliness, health, and mortality in old age: A national longitudinal study. *Social Science & Medicine, 74,* 907-914.

Nummela, O., Seppänen, M., & Uutela, A. (2011). The effect of loneliness and changes in loneliness on self-rated health (SRH): A longitudinal study among aging people. *Archives of Gerontology and Geriatrics, 53*, 163-167.

Ó Luanaigh, C., & Lawlor, B. A. (2008). Loneliness and the health of older people. *International Journal of Geriatric Psychiatry, 23*, 1213-1221.

O'Muircheartaigh, C., Eckman, S., & Smith, S. (2009). Statistical design and estimation for the National Social Life, Health, and Aging Project. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 64B(S1)*, i12-i19. doi:10.1093/geronb/gbp045.

Ortuño-Sierra, J., Badoud, D., Knecht, F., Paino, M., Eliez, S., Fonseca-Pedrero, E., & Debbané, M. (2013). Testing measurement invariance of the Schizotypal Personality Questionnaire-Brief scores across Spanish and Swiss adolescents. *PLoS ONE 8(12)*, e82041. doi:10.1371/journal.pone.0082041.

Peplau, L. A., & Perlman, D. (1982). Perspectives on loneliness. In L. A. Peplau & D. Perlman (Eds.), *Loneliness: A sourcebook of current theory, research and therapy* (pp. 1-20). New York: John Wiley & Sons.

Perissinotto, C. M., Cenzer, I. S., & Covinsky, K. E. (2012). Loneliness in older persons: A predictor of functional decline and death. *Archives of Internal Medicine, 172*, 1078-1083.

Pinquart, M., & Sörensen, S. (2003). Risk factors for loneliness in adulthood and old age—A meta-analysis. *Advances in Psychology Research, 19*, 111-143.

Rammstedt, B., & Krebs, D. (2007). Does response scale format affect the answering of personality scales? Assessing the Big Five dimensions of personality with different response scales in a dependent sample. *European Journal of Psychological Assessment, 23(1)*, 32-38.

Richter, D., & Schupp, J. (2012). SOEP Innovation Sample (SOEP-IS)—Description, Structure and Documentation. SOEP papers on Multidisciplinary Panel Data Research, #463. Berlin: Socio-Economic Panel (SOEP) at German Institute for Economic Research (DIW Berlin).

Russell, D. W. (1996). *UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. Journal of Personality Assessment*, 66, 20-40.

Russell, D. W., Peplau, L. A., & Cutrona, C. E. (1980). *The Revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. Journal of Personality and Social Psychology*, 39, 472-480.

Schwarz, N. (1990). Assessing frequency reports of mundane behaviors: Contributions of cognitive psychology to questionnaire construction. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology. Review of personality and social psychology*, Vol. 11 (pp. 98-119). Thousand Oaks, CA: Sage Publications, Inc.

Shiovitz-Ezra, S., & Ayalon, L. (2012). Use of direct versus indirect approaches to measure loneliness in later life. *Research on Aging, 34*, 572-591.

Stack, S. (1998). Marriage, family and loneliness: A cross-national study. *Sociological Perspectives, 41*, 415-432.

StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.

Van Tilburg, T. G., & Dykstra, P. A. (2008). Differences in older adults' loneliness and depression across Europe. In P.A. Dkystra (Ed.), *Ageing, intergenerational solidarity and age-specific vulnerabilities* (pp. 141-153). The Hague: NIDI.

Waite, L. J., Cagney, K., Dale, W., Huang, E., Laumann, E. O., McClintock, M., O'Muircheartaigh, C. A., Schumm, L. P., & Cornwell, B. (2014). National Social Life, Health, and Aging Project (NSHAP): Wave 2 and Partner Data Collection. ICPSR34921-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-29. doi:10.3886/ICPSR34921.v1.

Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment Research & Evaluation*, *12*, 1-26.