Payne, C., Hedberg, E. C., Kozloski, M., Dale, W., & McClintock, M.K. (2014). Using and interpreting mental health measures in the National Social Life, Health, and Aging Project. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *69(8)*, S99–S116, doi:10.1093/geronb/gbu100

Using and Interpreting Mental Health Measures in the National Social Life, Health, and Aging Project

Carolyn Payne,[1] E. C. Hedberg,[1] Michael Kozloski,[1] William Dale,[2] and Martha K. McClintock[3]

[1]NORC at the University of Chicago, Illinois.
[2]Department of Medicine, and
[3]Departments of Comparative Human Development and Psychology and the Institute for Mind and Biology, The University of Chicago, Illinois.

Introduction. National Social Life, Health, and Aging Project (NSHAP) included five unique mental health measures in Waves 1 and 2 that researchers can use to measure the overall emotional health of participants: depressive symptoms, happiness–unhappiness, anxiety symptoms, perceived stress, and felt loneliness. For each, we detail the rationale for its development and explain how to score, analyze, and interpret results.

Method. NSHAP developed its measures by modifying traditional short-form scales to improve response efficiency and reduce respondent burden. Scoring protocols and interpretations were developed for each measure. U.S. population estimates for older adults born between 1920 and 1947 were generated using age-eligible samples from Waves 1 and 2.

Results. NSHAP's protocols yielded U.S. prevalence rates similar to other nationally representative studies of older adults and comparable between waves. Higher estimates of anxiety symptoms and perceived stress in Wave 2 compared with Wave 1 were explained by age, administration mode, and time period. Analytic strategies for longitudinal analyses are provided. In Wave 2, mental health generally was worse at older ages, with women having more symptoms at younger ages than men. Women had fewer anxiety symptoms at the oldest ages.

Discussion. NSHAP's mental health measures were successfully integrated into the project's survey and showed acceptable external reliability as well as moderately stable individual characteristics across the 5 years between Waves 1 and 2. Depressive symptoms and unhappiness may form a mental health cluster in the elderly, distinct from anxiety symptoms, perceived stress, and felt loneliness. Gender differences in age-specific patterns of mental health were evident using the exact age of participants rather than the traditional decade groupings. Administration mode and time period (between 2005–2006 and 2010–2011) were determined to be potential confounds that need to be accommodated in longitudinal analyses of aging, whereas sample composition was not an issue for interpreting mental health measures.

Key Words: Anxiety—Depressive symptoms—Emotion—Loneliness—Mental health—Stress—Unhappiness.

Overview

The National Social Life, Health, and Aging Project (NSHAP) is a multidisciplinary study that seeks to determine relationships between sociological, psychological, and biological health factors among community-dwelling, older adults. Mental health is a key component of psychological function, and NSHAP has developed several measures so that it can be included in interdisciplinary analyses. Because NSHAP focuses on measuring health factors among community-dwelling, older adults, NSHAP defines mental health more broadly than clinical investigators. Our goal is to measure common emotions in older adults to study the full range of symptoms experienced by community-dwelling older adults. We consider feelings of unhappiness, loneliness, and stress to contribute to overall mental health, along with the more well-studied depressive and anxiety symptoms.

Our preliminary analyses show that mental health symptoms are among the most significant indicators of health status and are strongly associated with both social networks and mortality. Therefore, analysts that focus on the interrelationships between social life and health during aging are urged to include mental health status as controls, if not mediating factors. Moreover, mental health data in combination with NSHAP's data on cognitive and sensory function, such as frailty, sexuality, comorbidities, and mortality as well as couples, households, and neighborhoods, provide a unique and crucial resource for understanding the impact of mental health on diverse aspects of aging trajectories.

NSHAP provides five mental health measures based on existing short-form scales: depressive symptoms, happiness–unhappiness, anxiety symptoms, perceived stress, and felt loneliness. Although primarily querying negative symptoms, low-frequency scores indicate positive mental health states. Descriptive statistics of longitudinal data are presented in Figure 1. The purpose of this article is to provide recommended scoring protocols for these measures as well as analytic methods (with details provided in the Electronic Supplement) and interpretation guidelines with a short literature review targeted for nonpsychologist users.

Received November 30, 2013; Accepted July 10, 2014
Decision Editor: Mark Hayward, PhD

Wave 1 and Wave 2 data are publicly available (NSHAP Wave 1: Linda J. Waite, Edward O. Laumann, Wendy Levinson, Stacy Tessler Lindau, and Colm A. O'Muircheartaigh. NSHAP: Wave 1. ICPSR20541-v6. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research [distributor], April 30, 2014. doi:10.3886/ICPSR20541.v6. NSHAP Wave 2: Linda J. Waite, Kathleen Cagney, William Dale, Elbert Huang, Edward O. Laumann, Martha K. McClintock, Colm A. O'Muircheartaigh, L. Phillip Schumm, and Benjamin Cornwell. NSHAP: Wave 2 and

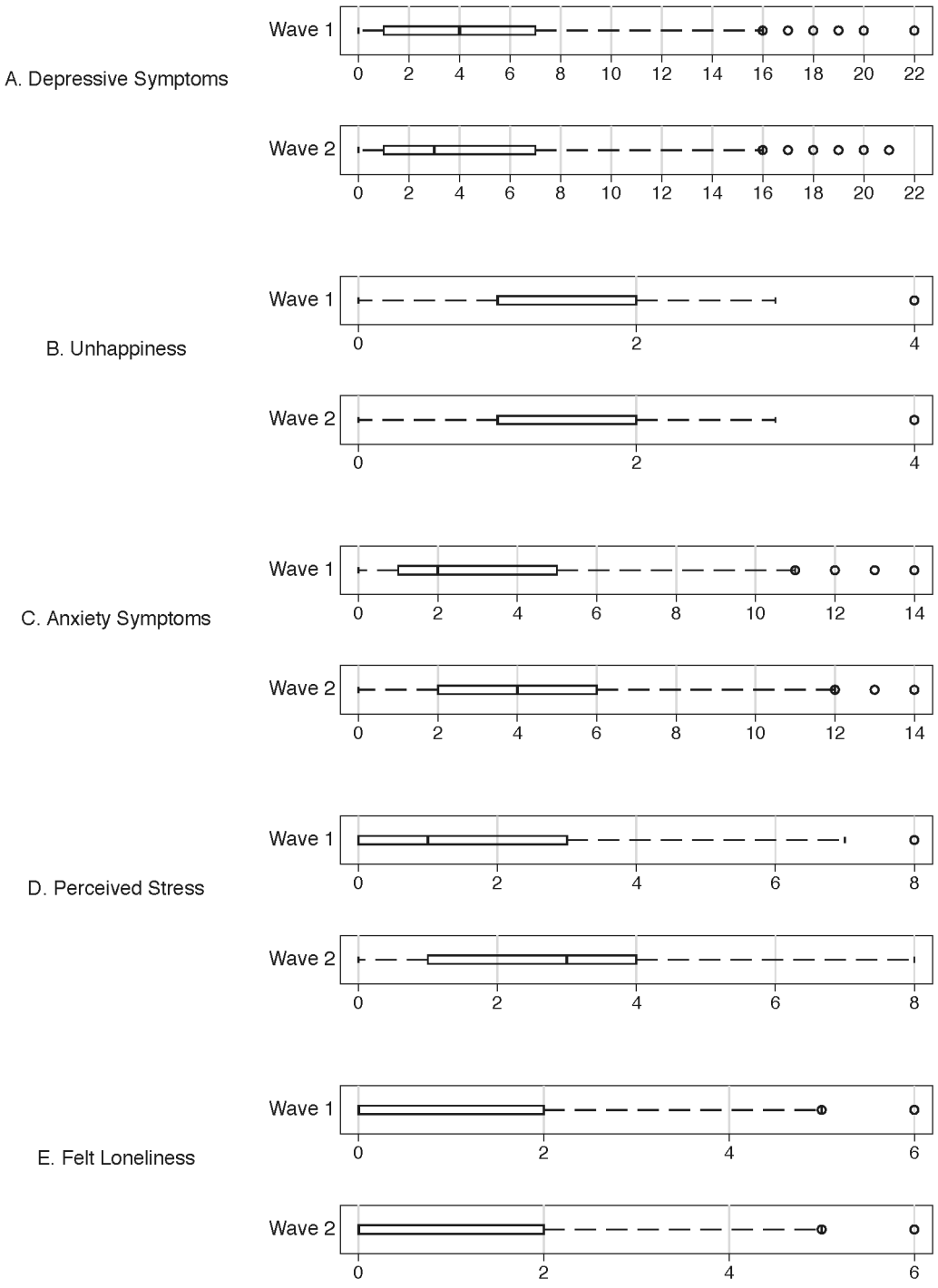

Figure 1. Distributions of scores on all five mental health measures estimated for the U.S. population of older adults from Waves 1 and 2. For each plot, the box area has three lines: the lower line represents the lower (25%) quartile, the middle line represents the median (50%), and the upper line represents the upper (75%) quartile. Two whiskers extend from the box to 1.5 times the lower quartile and 1.5 times the upper quartile. Dots are used for each observation that is outside of this range, with the exception of several extreme outliers that are excluded for presentation purposes only.

Partner Data Collection. ICPSR34921-v1. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research [distributor], April 29, 2014. doi:10.3886/ICPSR34921.v1.)

To facilitate the effective and accurate use of NSHAP's mental health measures, we present and solve three analytic problems. (a) By modifying well-established short-form scales to streamline the in-home interview, we created unique measures, which prevented the use of the original scales' scoring protocols and cutpoints. Here, we present the scoring protocols for NSHAP's measures that yield means and prevalences comparable to the literature. (b) We discovered that the traditional method of presenting tables of descriptive data by age decade categories was misleading and did not capture gender differences in aging patterns that would be of broad interest. Here, we present analytical strategies for age as a continuous variable that may be applied in future studies to determine potential gender differences in aging trajectories. (c) Although NSHAP used the same survey questions and response sets in Waves 1 and 2, longitudinal analyses of anxiety symptoms and perceived stress yielded significantly different results between waves (see Figure 1). Here, we identify the reasons for these wave differences so that analysts can adjust for them in their analyses.

One potential source of differences between Waves 1 and 2 might be their sample compositions. Such potential differences, however, are partially corrected by using the weighted values for each wave to adjust for sampling strategy and by restricting all participants in the analyses to be the same age as the returning respondents (see Tables 1 and 2). If future analysts use unweighted values, without age restrictions, they should be aware that the samples themselves do manifest significant differences in several characteristics (e.g., gender, age, race, cognitive function, and education), which are not reflective of the older U.S. population. In addition, not all of the original participants were reinterviewed in Wave 2. Those that went on to be reinterviewed in Wave 2 were similar in gender and race composition, but they were younger and had better self-rated physical and mental health as well as more education and better memory function (see Table 1). Finally, Wave 2 included the partners of the original participants as well as a few participants who had declined to be interviewed in Wave 1 but agreed in Wave 2 (Non-Interviewed Respondents [NIRs]; Table 2). These partners and NIRs were similar to the returning respondents but were slightly more likely to be white and have better cognitive function.

Table 1. Effects of Sample Loss Between Wave 1 (2005–2006; ages 57–85) and Wave 2 (2010–2011; ages 62–91)

	Reinterviewed in Wave 2 (N = 2,261)	Died, too sick to interview, or lost (N = 744)	p Value
Women	52.2%	49.2%	.19
Race/ethnicity			.31
White	81.2%	78.9%	
Black	9.6%	11.4%	
Hispanic, nonblack	6.9%	6.7%	
Other	2.3%	3.0%	
Mean age (*SD*)	67.1 (7.2)	71.2 (8.7)	<.001
Mean SPMSQ (*SD*)	9.3 (0.9)	8.7 (1.7)	<.001
Education			<.001
<HS	15.8%	28.0%	
HS/equivalent	25.5%	32.0%	
VC/AD/some college	32.0%	23.3%	
Bachelors or more	26.8%	16.8%	
Self-rated physical health			<.001
Poor	4.4%	15.1%	
Fair	16.5%	23.3%	
Good	29.3%	30.6%	
Very good	35.5%	22.5%	
Excellent	14.3%	8.5%	
Self-rated mental health			<.001
Poor	1.3%	3.5%	
Fair	7.2%	12.8%	
Good	24.4%	29.9%	
Very good	39.8%	33.5%	
Excellent	27.4%	20.3%	

Notes. AD = associates degree; HS = high school diploma; SPMSQ = Short Portable Mental Status Questionnaire; VC = vocational certification. All variables as collected in Wave 1. Wave 1 weighted values estimate the U.S. population, comparing estimates based on those who went on to a Wave 2 interview with those who did not (died, too sick to interview, or lost to follow-up). Comparison of those who were reinterviewed in Wave 2 to those who were deceased, too sick to interview, or lost to follow-up.

Depressive Symptoms

Depressed people have feelings of sadness, isolation, irritability, worthlessness, hopelessness, agitation, and guilt. These symptoms affect how a person feels, thinks, and behaves. Depression is associated with many emotional and physical problems that make it difficult to perform normal everyday activities (Sharp & Lipsky, 2002).

The prevalence of depression is increasing and is the leading cause of disability worldwide, affecting 350 million people in 2012 (Marcus, Yasamy, van Ommeren, Chisholm, & Saxena, 2012; Kessler, McGonagle, Swartz, Blazer, & Nelson, 1993; Klerman & Weissman, 1989; Wauterickx & Bracke, 2005). Depression is prevalent among older adults (Mirowsky & Ross, 1992), as it is associated with factors such as underemployment (Dooley, Prause, & Ham-Rowbottom, 2000), economic hardship (Mirowsky & Ross, 2001), lack of social support (Cornwell, 2003; Lin & Ensel, 1984), and poor health (Farmer & Ferraro, 1997). Depression has been shown to reduce a person's physical health status more than having angina, arthritis, asthma, or diabetes (Moussavi et al., 2007). When compared with people with normal mental health, people with increased depressive symptoms display higher mortality rates following myocardial infarctions (Frasure-Smith & Lespérance, 2008), higher levels of traditional cardiac risk factors (i.e., smoking, high cholesterol, hypertension, diabetes, and obesity) (Pozuelo et al.,

Table 2. Estimated Traits of U.S. Older Adults Aged 62–91 in 2010–2011

	Returning respondents (N = 2,261)	Partners + NIRs (N = 935)	p Value
Women	52.1%	53.2%	.67
Race/ethnicity			.09
White	80.8%	83.6%	
Black	10.0%	6.8%	
Hispanic, nonblack	6.7%	6.7%	
Other	2.5%	2.9%	
Mean age (*SD*)	72.5 (7.3)	72.0 (7.6)	.13
Mean CCFM score (*SD*)	13.7 (3.8)	14.2 (4.1)	.02
Education			.84
<HS	16.6%	15.4%	
HS/equivalent	25.3%	25.7%	
VC/AD/some college	31.4%	33.0%	
Bachelors or more	26.7%	25.9%	
Self-rated physical health			.86
Poor	5.4%	6.9%	
Fair	19.7%	17.3%	
Good	31.6%	32.2%	
Very good	30.7%	31.7%	
Excellent	12.6%	11.9%	
Self-rated mental health			.91
Poor	1.6%	1.6%	
Fair	10.1%	9.4%	
Good	30.4%	30.3%	
Very good	36.0%	38.4%	
Excellent	22.0%	20.3%	

Notes. AD = associates degree; CCFM = Chicago Cognitive Function Measure; HS = high school diploma; VC = vocational certification. All variables as collected in Wave 2. Comparison of Wave 1 respondents who returned in Wave 2 to their partners and those identified in Wave 1 but interviewed only in Wave 2 (NIRs).

2009), lower medication adherence (Gehi, Haas, Pipkin, & Whooley, 2005; Lin et al., 2004), more disturbed sleep (Jackowska, Kumari, & Steptoe, 2013), and higher rates of developing dementia (Metti et al., 2013). To better understand the causes and consequences of depressive symptoms in the U.S. population of older adults living at home, NSHAP modified a standard screening tool for depression.

Defining Depressive Symptomatology

Measuring frequencies of depressive symptoms is essential for quantifying levels of depression in the U.S. population of community-dwelling older adults, which range from normal fluctuations in mood to major clinical depression, as distinguished later. Fully consistent definitions, however, do not exist in the literature; clinicians and clinical investigators disagree about the magnitude of depressive symptoms required to diagnose specific clinical depressive disorders. Additionally, the considerable overlap among depressive disorders, which vary by the intensity and duration of depressive symptoms, make it challenging to produce consistent definitions (Blazer, 2003).

The most severe form of depression is *clinical or major depression*, an extremely debilitating disorder that often includes recurrent thoughts of death or suicidal ideation. Diagnostic criteria are described in the Fourth Edition of the American Psychiatric Association's (2000) *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV). *Minor, subsyndromal, or subthreshold depression* is associated with symptoms that are similar to major depression but are less severe, including impaired physical functioning, disability days, poorer self-rated health, use of antidepressant medications, and perceived low social support. Diagnostic criteria are also included in the Appendix of DSM-IV. Those that score ≥16 on the original 20-item Center for the Epidemiologic Studies Depression Scale (CES-D) (Radloff, 1977) but do not meet additional clinical criteria for major depression are typically considered to have minor depression.

Both clinical and minor depression are more prevalent among women and the unmarried (see Blazer, 2003 for a concise summary of these definitions, risk factors, and related evaluations). The term *clinically relevant depressive symptoms* refers to the depressive symptoms experienced by people who score ≥16 on the 20-item CES-D, which is a clinically validated cutpoint that includes people who are suffering from either minor or major depression. These symptoms are termed "clinically relevant" because the probability of clinically validated diagnosis is high (Blazer, 2003; Kohout, Berkman, Evans, & Cornoni-Huntley, 1993). Far more common, and relatively understudied, is the natural variation in everyday moods characterized by having some depressive symptoms, with CES-D scores ranging from 0 to 15.

Methods for NDSM

Instruments.—In order to efficiently (a) determine how often community-dwelling older adults experience depressive symptoms and (b) identify the national prevalence of those suffering from Frequent Depressive Symptoms (FDS, defined later), NSHAP used an existing 11-item short form of the CES-D and response categories based of the original NIMH form, thereby creating the NSHAP Depressive Symptoms Measure (NDSM). NDSM asks participants to describe the frequencies of their depressive symptoms within the past week; thus, the established cutpoint formally identifies those with Frequent Depressive Symptoms (FDS), which warrants further clinical testing to determine its sensitivity and specificity for the different types of depression. Equally important for characterizing depressive symptoms NDSM quantifies the variation in symptom frequency throughout the normal range typical of the U.S. population of community-dwelling older adults.

NDSM was derived from the short form of the CES-D, which was designed as a screening instrument to identify people at significant risk for clinical depression (Radloff, 1977). The CES-D has been widely validated through comparison of results to full clinical diagnostic evaluation (Lawton, Brody, & Saperstein, 1989; Ritchey, La Gory, Fitzpatrick, & Mullis, 1990). Due to the length of the original CES-D measure (20

depressive symptom items each with four response categories for symptom frequency, abbreviated as 20x4), several short forms have been derived and shown to produce results consistent with those of the original measure (Kohout et al., 1993). NSHAP chose to base its measure of depressive symptoms on a short form of the CES-D, known as the Iowa form from the Established Populations for Epidemiological Studies of the Elderly (EPESE), to minimize respondent burden and the overall interview time of NSHAP's survey. Although NDSM is based on screening tools that have previously been shown to identify clinically relevant depressive symptoms, NSHAP's measure has not yet been anchored to other clinically validated diagnoses and, thus, high scores should not be labeled as clinically relevant.

NDSM Wording.—NDSM quantifies the frequency of 11 symptoms *during the past week* using symptom descriptions and a time frame identical to the original 20-item CES-D and the EPESE Iowa short form (Table 3).

NDSM Response Categories.—The response categories for NSHAP's measures of depressive symptoms, anxiety symptoms, and perceived stress were standardized as follows: *Rarely or none of the time*, *some of the time, occasionally, and most of the time.* These category labels are shortened versions of those from the original CES-D (Radloff, 1977; Table 4). For abbreviated versions of the CES-D, a consensus was achieved (Kohout et al., 1993) that the two most frequent response categories should be combined into one, termed *much or most of the time* (11 × 3). Therefore, combining NSHAP's responses *occasionally* and *most of the time* into a single category is necessary to achieve full comparability of the NDSM to the well-validated EPESE's Iowa 11×3 CES-D scale (Table 4).

NDSM scoring.—The three response categories for symptom frequency in the NDSM, *Rarely or none of the time*, *some of the time*, and *much or most of the time*, were scored as 0, 1, and 2, with higher scores reflecting more recurrent symptoms. Scores on each of the 11 items were summed to produce a total score ranging from 0 to 22.

Cutpoint or caseness score for frequent depressive symptoms (FDS).—For those analysts wanting a dichotomous variable comparable to the original scales and other surveys, we recommend using ≥9 as the cutpoint for individuals that have Frequent Depressive Symptoms (FDS), which yields a sample prevalence comparable to other epidemiological studies (see "Results for NDSM" and Blazer, 2003; Beekman et al., 1995; Berkman et al., 1986; Steffick et al., 2000). Our cutpoint is based on the simulation regression equation developed by Kohout and colleagues to convert the validated cutpoint for the original 20-item CES-D into a cutpoint for the EPESE Iowa's 11×3 CES-D short form (Table 2 of Kohout et al., 1993: $16 = 1.87X + 0.53$ where "X" is the adjusted cutpoint for an 11×3 CES-D). The solution to this equation is 8.3. Because the CES-D is an integer scale, the cutpoint is 9 or greater for NSHAP's 11×3 measure.

Table 3. Eleven Items Selected From the Original 20-Item CES-D Scale (Radloff, 1977)

Scale item	Original	EPESE Iowa; HRS Wave 1; NSHAP Waves 1 and 2
I was bothered by things that usually don't bother me	•	
I did not feel like eating; my appetite was poor	•	•
I felt that I could not shake off the blues even with help from my family or friends	•	
I felt that I was just as good as other people	•	
I had trouble keeping my mind on what I was doing	•	
I felt depressed	•	•
I felt that everything I did was an effort	•	•
I felt hopeful about the future	•	
I thought my life had been a failure	•	
I felt fearful	•	
My sleep was restless	•	•
I was happy	•	•
I talked less than usual	•	
I felt lonely	•	•
People were unfriendly	•	•
I enjoyed life	•	•
I had crying spells	•	
I felt sad	•	•
I felt that people disliked me	•	•
I could not get "going"	•	•

Notes. CES-D = Center for the Epidemiologic Studies Depression Scale; EPESE = Established Populations for Epidemiological Studies of the Elderly; HRS = Health and Retirement Study; NSHAP = National Social Life, Health, and Aging Project.

We do not recommend the alternative simple proportional method. This consists of calculating a cutpoint from the ratio of maximum scores of the two scales multiplied by the full scale's cutpoint (e.g., Steffick et al., 2000; Zauszniewski & Bekhet, 2009). Using NSHAP's Wave 1 data, this method yields a cutpoint of 6, categorizing 37% of NSHAP's sample as experiencing FDS. This prevalence is much higher than prevalences reported elsewhere for either significant depressive symptoms or being at risk for clinical depression (Blazer, 2003; Beekman et al., 1995; Berkman et al., 1986; Steffick et al., 2000).

Results for NDSM

Based on the calculated cutpoint for having Frequent Depressive Symptoms (FDS), NSHAP's Wave 2 protocol yielded 19% as the U.S. prevalence among older adults, after controlling for age and gender. This value is close to the 18% national prevalence of significant depressive symptoms estimated by the Health and Retirement Study (HRS Wave 1; a

Table 4. Response Categories for Symptom Frequency Within the Past Week on Five Modified Forms of the CES-D Scale

Original 20-item scale	EPESE Iowa Scale	HRS Scale	NDSM Waves 1 and 2	NDSM analysis Waves 1 and 2
Rarely or none of the time (less than 1 day).	Hardly ever or never.	None or almost none of the time.	Rarely or none of the time.	Rarely or none of the time.
Some or little of the time (1–2 days).	Some of the time.	Some of the time.	Some of the time.	Some of the time.
Occasionally or a moderate amount of time (3–4 days).	Much or most of the time.	Most of the time.	Occasionally.	Much or most of the time.
Most or all the time (5–7 days).		All or almost all of the time.	Most of the time.	

Notes. CES-D = Center for the Epidemiologic Studies Depression Scale; EPESE = Established Populations for Epidemiological Studies of the Elderly; HRS = Health and Retirement Study; NDSM = NSHAP's Depressive Symptoms Measure; NSHAP = National Social Life, Health, and Aging Project.

study that applied a comparable instrument within a similar population). NSHAP's Wave 1 prevalence estimate of 20% is similar to both values, and score distributions in Wave 1 and Wave 2 are similar (Figure 1A). In addition, NDSM's reliability coefficients were high and comparable between waves (0.79 in Wave 1 and 0.78 in Wave 2; Table 5).

When considering the entire range of symptom frequency as a continuous variable (0–22 symptoms), the average Wave 1 respondent reinterviewed in Wave 2 had fewer depressive symptoms 5 years later, although relative individual differences remained stable. However, across chronological ages, the symptoms increased (Table 6). This apparent contradiction suggests that there was a time period effect between 2005–2006 and 2010–2011 that reduced the number of depressive symptoms for the average participant, offsetting the increase expected with being 5 years older. Thus, future analyses of age trajectories and time period effects will benefit from focusing on the continuous variable of symptom frequency across the entire range of symptom frequencies and recognizing the relative stability of depressive symptom frequency within a person over a five year period.

We took another approach to understanding age effects by evaluating age as a continuous variable solely among Wave 2 participants. Doing so revealed gender differences in changes across ages. Graphically, the U.S. age-specific NDSM scores increased linearly with each year of age between 62 and 91, while comparable increases among men only occurred around 80 years of age. Thus, the gender difference was evident primarily between the ages of 67 and 79, when women appeared to experience depressive symptoms more frequently than men (Figure 2A). Initial regression analyses did not confirm statistically significant interactions between gender and age (gender, age, age^2 linear regression, .411 ≤ all p-values ≤ .503). Nevertheless, future age-targeted analyses are warranted, particularly because unhappiness results shared strikingly similar gender differences in symptom frequency at increasingly older ages.

Happiness–Unhappiness

Despite the importance of detecting and addressing unhappiness for improving quality of life, unhappiness has been only a minor research focus. In contrast, studies of happiness are abundant. Happiness, a state ranging from contentment to joy, is associated with positively assessing life quality and has been linked to race, social participation, and socioeconomic status (Argyle, 1999; Clemente & Sauer, 1976). It is also frequently treated as an outcome in evaluations of the benefits of income growth (Veenhoven & Hagerty, 2006), marriage (Schnittker, 2008; Stack & Eshleman, 1998; Waite & Gallagher, 2000), and health (Dockray & Steptoe 2010; Yang, 2008; Kirby, Coleman, & Daley, 2004). Relationships between unhappiness and important social and economic factors are much less studied; although, several reports suggest that unhappiness is not simply the absence of happiness (Cacioppo & Berntson, 1994; Russell & Carroll, 1999; Tay, 2011). Thus, NSHAP developed a measure to evaluate frequencies of both happy and unhappy feelings, providing the opportunity to explore the relationships of both with important health and social factors.

Table 5. Reliability Coefficients for NSHAP's Mental Health Measures

Facet of mental health	Wave 1	Wave 2
Depressive symptoms (NDSM, 11 items)	0.7886	0.7828
Anxiety symptoms (NASM, 7 items)	0.7185	0.7372
Perceived stress (NPSM, 4 items)	0.6463	0.6262
Felt loneliness (NFLM, 3 items)	0.8048	0.7881

Notes. NASM = NSHAP's Anxiety Symptoms Measure; NDSM = NSHAP's Depressive Symptoms Measure; NFLM = NSHAP's Felt Loneliness Measure; NPSM = NSHAP's Perceived Stress Measure; NSHAP = National Social Life, Health, and Aging Project.

The NSHAP Happiness–Unhappiness Measure (NHUM) combined the unipolar General Social Survey's Single-Item Subjective Happiness Scale (SISHS) (Lee & Bulanda, 2005) with two additional "unhappy" response categories to create a single item measure that detects the frequency of both happy and unhappy feelings. By itself, the SISHS produces strong and stable associations with measures of general well-being and other aspects of life satisfaction, including the Affect Balance Scale (Bradburn, 1969), Index of General Affect and Well-Being (Campbell, Converse, & Rodgers, 1976), and Life Satisfaction Scale (Andrews & Withey, 1976). Paralleling NSHAP's other mental health measures that focus on negative health symptoms, we

Table 6. Within-Person Stability and Change in the Mental Health of Participants Studied in Waves 1 and 2: Effects of Age, Gender, and Wave (ANCOVA, Repeated Measures Wave 1 and Wave 2, Data Are Centered so That the Wave 2 Change From Wave 1 Is for the Average Participant)

Mental health measures	Depressive symptoms (NDSM)	Unhappiness (NHUM)	Anxiety symptoms (NPSM)	Perceived stress (NPSM)	Felt loneliness (NFLM)
N	2,200	2,249	1,655	1,706	1,558
ANCOVA coefficients					
Within-person W1–W2 correlation	0.51***	0.45***	0.44***	0.41***	0.54***
Wave 2 change from Wave 1	0.41*	0.01	0.96***	1.09***	0.06
Women (men = referent)	0.26	0.06	0.02	0.05	0.05
Age (per year)	0.06*	<0.01	0.04*	<0.01	0.01
Age^2	<0.01	<0.01	<0.01	<0.01	<0.01

Notes. ANCOVA = analysis of covariance; NDSM = NSHAP's Depressive Symptoms Measure; NFLM = NSHAP's Felt Loneliness Measure; NHUM = NSHAP's Happiness–Unhappiness Measure; NPSM = NSHAP's Perceived Stress Measure; NSHAP = National Social Life, Health, and Aging Project.
*p < .05. ***p < .001.

quantified unhappiness by reverse coding the SISHS and adding two explicit response options for unhappiness.

Methods for NHUM

The NSHAP Happiness-Unhappiness Measure (NHUM) asked respondents, "*If you were to consider your life in general these days, how happy or unhappy would you say you are, on the whole…extremely happy, very happy, pretty happy, unhappy sometimes, or unhappy usually?*" Response categories were scored from 1 (*extremely happy*) to 5 (*unhappy usually*). In addition to reverse coding, NSHAP's measure differs from the SISHS, both in its slightly different stem, "*Taken all together, how would you say things are these days–would you say that you are...,"* and its response categories, *"very happy, pretty happy, or not too happy?*" (Lee & Bulanda, 2005). In addition, because the SISH is a measure of happiness, higher scores indicate higher levels of happiness. Only a full comparison between the two scales and a discussion of the limitations of NSHAP's scale is provided in the "Discussion" section.

Results for NHUM

On average, Waves 1 and 2 yielded the same population estimate of the mean NSHAP Happiness-Unhappines Measure (NHUM) score in the U.S. population of older adults (2.38 ± .02 and 2.39 ± .02, respectively). Moreover, there was a significant within-person correlation across waves (Table 6) and the distributions for Waves 1 and 2 were nearly identical (Figure 1B), as were response category distributions (Table 7).

Similar to the results for depressive symptoms, the estimated age-specific nationally representative scores on NHUM increased linearly with age for women, indicating greater unhappiness, while similar increases among men did not occur until age 80 and above (Figure 2B). This apparent gender difference in aging patterns was not statistically significant in Wave 2 in a simple model (gender, age, and age^2 linear regression, .185 ≤ all p-values ≤ .300). Nevertheless, the pattern of gender differences across ages for unhappiness was strikingly similar to the pattern for depressive symptoms (Figure 2A and B), indicating its robustness and worthiness for further investigation.

Anxiety Symptoms

Anxiety is a mood characterized by apprehension, worry, or foreboding out of proportion to a specific situation or not about anything in particular. Anxiety is also associated with significant physical symptoms, including hyperactivity, poor concentration, and autonomic arousal (Rosenbaum et al., 1997).

Anxiety symptoms are quite prevalent and can have significant health consequences. Anxiety is correlated with decreased health and well-being (Denollet, Maas, Knottnerus, Keyzer, & Pop, 2009; Sherbourne, Wells, Meredith, Jackson, & Camp, 1996). It has been linked to increased mortality in both men (Kawachi, Sparrow, Vokonas, & Weiss, 1994; Van Hout et al., 2004) and women (Denollet et al., 2009). Additionally, people who suffer from anxiety are more likely to utilize health services than those who are not (Frazier & Waid, 1999; Simon, Ormel, VonKorff, & Barlow, 1995; Wiltink et al., 2009). In the Epidemiological Catchment Area Studies, the prevalence of anxiety disorders outranked the prevalence of depressive disorders and dementia (Regier et al., 1988). Anxiety conditions are also common in older adults (Mehta et al. 2003; Flint, 1994; Vermeulen, Beekman, & Stek, 1994) as they are associated with stressful life events, deteriorating physical health, cognitive decline, lower socioeconomic status, and reduced social networks (Beekman et al., 1998). Thus, determining whether people suffer from anxiety is critical for studying their overall health and well-being.

Several screening instruments have been developed using specific definitions of anxiety conditions, which vary by the intensity, type, and duration of symptoms. Severe forms of anxiety disorders are defined according to the DSM-IV criteria and diagnosed using the Diagnostic Interview Schedule. Severe anxiety disorders include generalized anxiety disorder, phobias, panic disorder, obsessive-compulsive disorder, and posttraumatic stress disorder (Beekman et al., 1998). Screening instruments identify people who are at risk for anxiety disorders by detecting high levels of clinically relevant anxiety symptoms.

To efficiently identify the national prevalence of community-dwelling older adults who suffer from recurring anxiety symptoms, NSHAP used the items from the Hospital

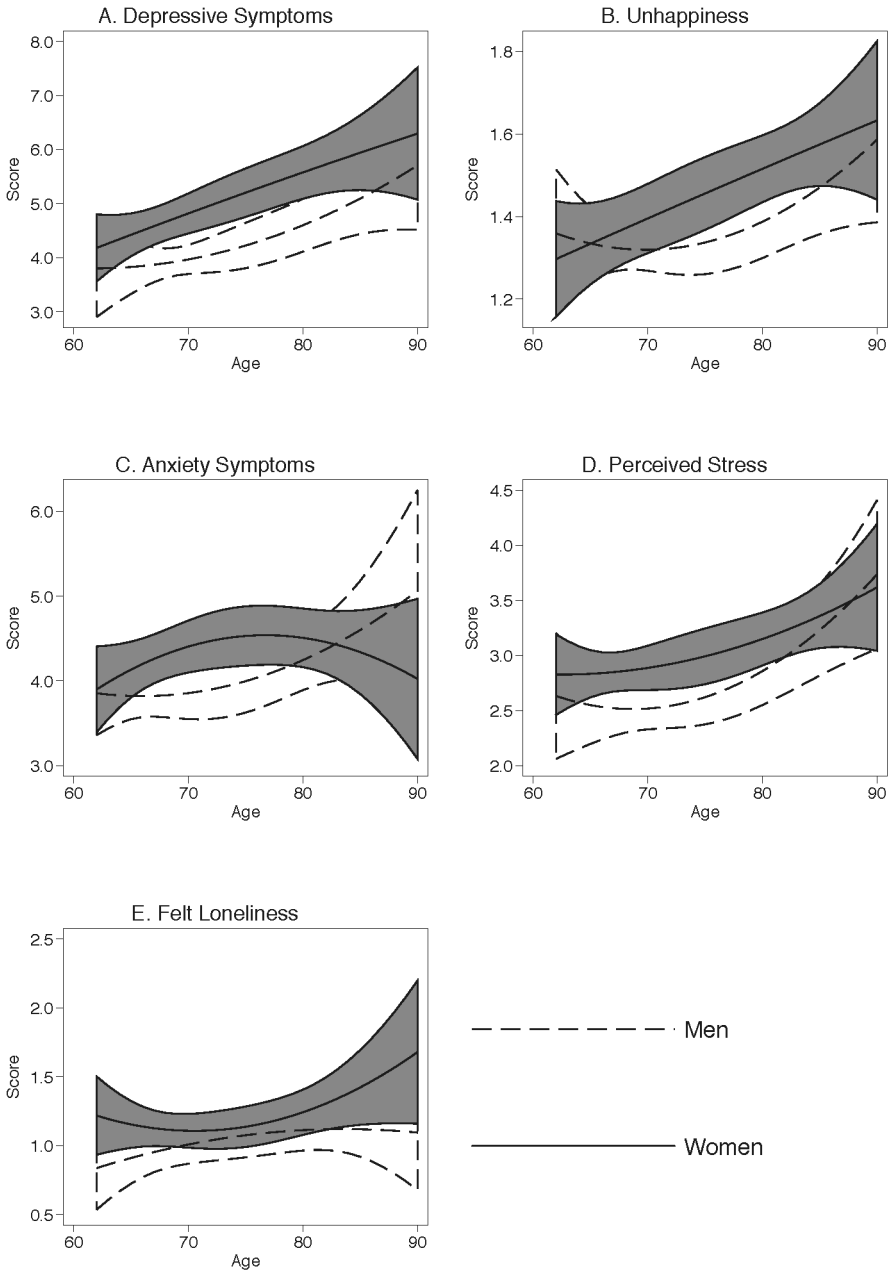

Figure 2. Plots of estimated U.S. age-specific scores on NSHAP's Wave 2 mental health measures by age and gender. Quadratic regression lines of outcome on age and prediction confidence intervals are displayed. Although confidence intervals overlap, general patterns are summarized by the regression lines. However, because these models do not include key control variables, these figures are only for descriptive purposes, laying the foundation for future analyses.

Anxiety and Depression Scale's Anxiety Subscale (HADS-A) and frequency response categories that match those of the NDSM and NPSM, creating the NSHAP Anxiety Symptoms Measure (NASM). The HADS-A was developed by Zigmond and Snaith in 1983 and is a well-validated scale that detects possible cases of anxiety disorders in nonpsychiatric populations (Bjelland, Dahl, Haug, & Neckelmann, 2002; Herrmann, 1997). HADS-A was designed for use in a hospital setting; however, several studies have confirmed that it is valid when applied in community settings or primary care practices (Snaith, 2003). Although HADS-A has been clinically validated and assesses the intensity of anxiety symptoms to predict possible anxiety disorder cases, NASM focuses solely on the frequency of symptoms to accelerate administration as a survey module aimed to measure variation in anxiety symptoms among home-dwelling older adults in the United States.

Methods for NASM

Wording and response categories.—NSHAP's Anxiety Symptoms Measure (NASM) assesses the frequency of respondents' anxiety symptoms *during the past week* using

seven items that are identical to those of the HADS-A, with the exception of being stated in past tense rather than present tense (Zigmond & Snaith, 1983). NSHAP modified the phrasing of the HADS-A symptom frequency response categories to maintain consistency between mental health measures, with the goal of reducing interview time and respondent burden. The NASM symptom frequency response categories are *rarely or none of the time, some of the time, occasionally,* and *most of the time.*

Mode of Administration.— Note that the NASM questionnaire was part of the in-person, computer-assisted personal interview (CAPI) for some participants (N = 1,993) and the leave-behind questionnaire (LBQ) for others (N = 763) in Wave 1, but it was included solely in the LBQ in Wave 2 (Table 8). In addition, Waves 1 and 2 sample populations overlapped considerably, but they were not identical (see Tables 1 and 2), an issue that can be addressed by using weighted scores and comparable age ranges as we do here.

NASM scoring.—The four response categories for symptom frequency in NSHAP's 7×4 measure were scored from 0 to 3, with 3 corresponding to the highest frequency category (*most of the time* for all items except 4, which is reverse coded). Scores corresponding to the participant responses on each of the seven items were summed to produce a final score ranging from 0 to 21, with higher scores reflecting more recurring anxiety symptoms. *Occasionally* and *most of the time* response categories were not collapsed as they were for NDSM and NPSM, because NASM scoring assignments and ranges were identical to those of the well-validated HADS-A (Zigmond & Snaith, 1983).

Cutpoint or caseness score for frequent anxiety symptoms (FAS).—We recommend using ≥8 as a cutpoint to identify participants have anxiety symptoms significantly often, Frequent Anxiety Symptoms (FAS). This cut point based on the literature yields a sample prevalence comparable to other epidemiological studies. Scores of 8–10 on the original HADS-A scale were described by its developers (Zigmond & Snaith, 1983) as detecting "doubtful cases of anxiety disorders" while scores of 11 or higher described participants who were likely suffering from clinically significant anxiety. Applying a cutpoint of ≥8 provides a subsample of participants who frequently suffer from anxiety symptoms and who should undergo further clinical testing to determine if severe anxiety disorders are present (Zigmond & Snaith, 1983).

Table 7. Stable Prevalence of Responses for Happiness–Unhappiness Estimated From Waves 1 and 2

Response category	Wave 1 percent	Wave 2 percent
Extremely happy	15	15
Very happy	42	42
Pretty happy	34	33
Unhappy sometimes	8	9
Unhappy usually	1	1

Results for NASM

Prevalence rates for significantly FAS are largely unknown among community-dwelling older adults, with reports ranging from 17% to 52% (Potvin, Forget, Grenier, Préville, & Hudon, 2011). NSHAP's protocol yielded an overall prevalence of FAS of 13% in Wave 1 and 21% in Wave 2, with comparable reliability coefficients for NASM in each wave (.72 in Wave 1 and .74 in Wave 2; Table 5). The prevalence of frequent symptoms increased among women between waves from 15% to 22% and among men from 11% to 19%, both of which are similar to the prevalence of "sufficient anxiety symptoms" reported by Himmelfarb and Murrell (1984; 21.5% among women and 17.1% among men).

When measured as a continuous variable, the frequency of anxiety symptom increased significantly between Waves 1 and 2 (Figure 1). The increase was significant both when using adjusted values for all participants in Waves 1 and 2 and just among those Wave 1 respondents returning after 5 years (all p-values < .00001 for repeated measures using a centered analysis of covariance (ANCOVA); stacked regression, paired-t, and quasi-independent t tests).

In addition to the effect of aging 5 years (demonstrated by repeated measures analyses), the increase in estimates of anxiety symptom frequency between all participants of Waves 1 and 2 could have arisen from three sources: mode of questionnaire administration, time period, and sample composition. We evaluate each and provide analytic strategies to tailor different types of studies accordingly. It is noteworthy, nonetheless, that despite these factors, there was significant stability in FAS within individuals (Table 6).

First, in Wave 1, participants were randomly assigned to complete NASM during the in-person, CAPI or the LBQ

Table 8. Mode of Administration and Completion Rates (in Parentheses) of NSHAP's Mental Health Measures in Waves 1 and 2

	Wave 1 mode		Wave 2 mode	
Mental health measure	In-person interview	Leave-behind questionnaire[a]	In-person interview	Leave-behind questionnaire[a]
Depressive symptoms	100% (98%)	0%	100% (99%)	0%
Unhappiness	100% (100%)	0%	100% (100%)	0%
Anxiety symptoms[b]	67% (67%)	33% (27%, 25%)	0%	100% (87%, 78%)
Perceived stress[b]	67% (66%)	33% (27%, 26%)	0%	100% (87%, 80%)
Felt loneliness	0%	100% (84%, 79%)	0%	100% (87%, 80%)

Notes. NSHAP = National Social Life, Health, and Aging Project.
[a]% returned leave-behind questionnaire, % completed mental health measure.
[b]In-person interview respondents completed Module A only; leave-behind questionnaire respondents received Version 2 only.

(one-third randomly assigned to LBQ; see Table 8). Those who completed NASM in the LBQ manifested a higher prevalence of FAS (18% ± .02) than those who completed it in the CAPI (11% ± .01). This is consistent with the finding that higher prevalences are reported when questionnaires about negative, private, or possibly embarrassing traits are conducted anonymously rather than face-to-face with an interviewer (Bowling, 2005). In Wave 2, NASM was administered only in the LBQ (Table 8), which contributed to the significantly higher Wave 2 FAS prevalence ($p < .001$). To avoid this confound, the effect of questionnaire mode administration can be modeled and used to adjust the Wave 1 CAPI scores. Alternatively, analysts can case restrict the Wave 1 sample to include only those who completed the LBQ.

Second, among original Wave 1 participants interviewed in 2005–2006 and again in 2010–2011 (Wave 2), the prevalence of FAS almost doubled (12% ± .01 to 20% ± 0.01; $p < .001$) and anxiety symptom frequency, measured continuously, increased significantly (3.42 ± .10 to 4.60 ± .10; $t = 8.45$; $p < .001$). In an ANCOVA, we eliminated the potential for sample bias by restricting cases to only those Wave 1 participants reinterviewed in Wave 2. The increase was primarily due to wave differences, in addition to aging 5 years (Table 6). Analysts can test the hypothesis that the wave differences are due to a time period effect, perhaps caused by the recession beginning just before 2010–2011 and increasing the frequency of anxiety symptoms experienced in Wave 2 among both returning and new participants in Wave 2.

Third, in Wave 2, the newly recruited partners and Wave 1 nonrespondents had a higher prevalence of FAS (23% ± .02) than the original participants (20% ± .01). We considered the possibility that this higher prevalence could result from differences in personal characteristics between the two sample populations (see Tables 1 and 2). However, there was no significant difference between the two populations in self-rated mental or physical health, and the only distinctive characteristics of the newly recruited partners and Wave 1 nonrespondents are associated with lower, not higher, anxiety, including better cognitive function and a tendency to be white. Thus, sample characteristics appear unlikely to explain the increase in anxiety symptoms from Wave 1 to Wave 2, although analysts may wish to investigate further, include these characteristics in their models, or use longitudinal analyses to study the effect of aging 5 years.

Furthermore, evaluating age as a continuous variable in Wave 2 analyses revealed not only gender differences across chronological ages in a cross-sectional analysis, but also a pattern quite different compared with those of depressive symptoms and unhappiness. The estimated U.S. age-specific NASM scores among women did not increase linearly across chronological ages. Rather, anxiety scores peaked around 75 years of age and then were nearly as low again by age 90 as at age 62 (see Figure 2C). Although men and women manifested anxiety symptoms with the same estimated frequency around age 62, men's scores increased linearly, if not exponentially, through age 90. This gender difference in age patterns tended to be significant in an initial analysis (gender × age, coef = .73, $t = 1.67$, $p = .100$; gender × age^2 coef = .01, $t = –1.72$, $p = .091$, gender coef = –25.91, $t = –1.61$, $p = .115$, age and age^2 terms ≥ .38). Moreover, the overall pattern in Figure 2C differs from that shared by depressive symptoms, unhappiness, and perceived stress, a topic that will be interesting to explore in future analyses.

Perceived Stress

Perceived stress is the feeling that problems are piling up too high to manage. Thus, perceived stress differs from stress, which is defined by a physiological state. Perceived stress can profoundly impact a person's overall health and well-being (Cohen, Kamarck, & Mermelstein, 1983). Elevated levels have been linked to increased risks of cardiovascular disease (Baum & Posluszny, 1999; Hamarat et al., 2001; McDade, Hawkley, & Cacioppo, 2006; Rosengren, Tibblin, & Wilhelmsen, 1991), cancer diagnosis (Baum & Posluszny, 1999; Hamarat et al., 2001), strokes (Hamarat et al., 2001; Jood, Redfors, Rosengren, Blomstrand, & Jern, 2009), anxiety disorders, and depression (Kelly, Tyrka, Anderson, Price, & Carpenter, 2008). Perceived stress also affects health indirectly, as it has been associated with higher fat diets, smoking, and less exercise (Ng & Jeffery, 2003). In older people who have reduced physical health and social support, stress from diverse sources is particularly prevalent. NSHAP aimed to measure how much stress respondents perceived in their daily lives.

Methods for NPSM

The NSHAP's Perceived Stress Measure (NPSM) was derived from the only empirically established index of general stress appraisal, The Perceived Stress Scale (PSS) (Cohen et al., 1983). The orginal PSS measures the frequency of 14 perceived stressors in the past month using five frequency response categories. Short versions of the PSS are recognized as valid indices, including a 4-item scale developed by Cohen & Williamson (1988). NPSM was derived from this 4-item scale, which was similar to the scale used by the HRS in 2002.

Wording and response categories.—NSHAP changed the wording of the symptom items in the original 4-item PSS in two ways: (a) the items were phrased as a first-person declarative statement rather than a question and (b) the referent time frame was shorter (*during the past week* rather than *month*) (see Table 9 for phrasing). These modifications increased comparability to other NSHAP mental health measures. For ease of completion and interview efficiency, the response categories also matched those of NDSM and

Table 9. Rephrasing of the Perceived Stress Scale Short Form

Cohen and Williamson (1988) HRS 2002	NPSM Waves 1 and 2
In the last month, how often have you felt that you were unable to control the important things in your life?	During the past week, I was unable to control important things in my life.
In the last month, how often have you felt confident about your ability to handle your personal problems?	During the past week, I felt confident about my ability to handle personal problems.
In the last month, how often have you felt that things were going your way?	During the past week, I felt that things were going my way.
In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	During the past week, I felt that difficulties were piling up so high I could not overcome them.

Notes. HRS = Health and Retirement Study; NPSM = NSHAP's Perceived Stress Measure; NSHAP = National Social Life, Health, and Aging Project.

Table 10. Comparison of Symptom Frequency Response Categories Utilized by Four Forms of the Perceived Stress Scale

Original 4-item scale	HRS 2002	NPSM questionnaire	NPSM analyses
Never	Hardly ever	Rarely or none	Rarely or none of
Almost never	(or never)	of the time	the time
Sometimes	Some of the time	Some of the time	Some of the time
Fairly often	Often	Occasionally	Much or most
Very often		Most of the time	of the time

Notes. HRS = Health and Retirement Study; NPSM = NSHAP's Perceived Stress Measure; NSHAP = National Social Life, Health, and Aging Project.

NASM, rather than the five response categories in the original PSS or the three in HRS (Table 10). To increase comparability between NSHAP and HRS measures, we combined NSHAP's two highest symptom frequency categories (*occasionally* and *most of the time)* into a single category termed *much or most of the time*, as we described for depressive symptoms (as mentioned earlier and in Table 4).

Mode of administration.—Also note that NPSM was completed by themselves after the interview in the LBQ in Wave 2, whereas it was randomly assigned in Wave 1 to be administered by the interviewer in the CAPI (N = 1,984) or the LBQ (N = 767) (one third randomly assigned to LBQ; see Table 8).

NPSM scoring.—The three response categories for symptom frequency were scored from 0 to 2 (with 2 indicating the highest frequency) and then summed to yield a NPSM score ranging from 0 to 8. This score represents the frequency of perceived stress symptoms in the past week (Cohen et al., 1983). For prevalence estimates of frequently feeling stressed, we recommend reporting the percentage of participants with a NPSM score of 1 or higher (i.e., those who report experiencing stress symptoms more frequently than *rarely or none of the time*), termed FPS.

Results for NPSM

The reliability coefficients for NPSM were comparable between waves (.65 in Wave 1 and .63 in Wave 2; Table 5), and individuals participating in Waves 1 and 2 maintained their frequency of perceived stress symptoms relative to the population across the intervening five years (i.e., correlated scores indicating stable individual differences; Table 6). Nonetheless, the mean perceived stress score was higher for the Wave 2 cohort than the Wave 1 cohort (2.85 ± .06 vs. 1.59 ± .04, respectively) as was the prevalence of frequently perceived stress (77% vs. 52%, respectively). Among longitudinal respondents, the prevalence of perceived stress also increased from Wave 1 to Wave 2 (paired t and quasi-independent tests, all ps < .0001). The 2002 HRS estimate (1.61 ± .04; 63%) lies between NSHAP's estimates based on Waves 1 and 2.

The increase in perceived stress for the average participant between waves (1.09 ± .12) was independent of aging 5 years, which was not significant in this simple statistical model (Table 6). We tested and confirmed this hypothesis through analyses of effects related to administration mode and time period of the interviews. First, Wave 1 participants had higher NPSM scores if they received the measure in the LBQ (1.91 ± .08) than if they received it in the CAPI (1.46 ±.05). Second, the mean among longitudinal participants increased significantly from Wave 1 to 2 (1.45 ±.05 to 2.76 ±.07). Third, new Wave 2 participants, Wave 1 partners, and NIRs, had a significantly higher mean (2.96 ± .10) than did returning participants in Wave 2 (2.77 ± .07). This is not likely a simple effect of sample composition because the new participants did not have worse mental and physical health, which leads to increased perceptions of stress. Instead, theirs was similar to the returning respondents (see Table 2).

Focusing on Wave 2, men and women perceived stress more frequently at older ages (Figure 2D). There was no statistically significant gender difference in these higher values at older ages (gender, age, age^2 linear regression, .080 ≤ all p-values ≤ .509). Nevertheless, the graphical pattern in Figure 2D closely resembles that of depressive symptoms and unhappiness, which calls for further analyses of age-specific gender differences and aging trajectories.

Felt Loneliness

Loneliness is the feeling of social isolation, arising from perceived deficits in either the number or quality of social relationships (Peplau & Perlman, 1982). It may be the subjective experience of an objectively small social network, i.e., actual isolation, or it can be felt even within large networks (Stack, 1998). Felt loneliness is associated with several health problems, including severe depressive symptoms (Cacioppo, Hughes, Waite, Hawkley, & Thisted, 2006), cardiovascular disease, elevated blood pressure, poorer sleep quality (Cacioppo, Hawkley, & Thisted, 2010), Alzheimer's

disease (Wilson et al., 2007), and decreased physical activity (Hawkley, Thisted, & Cacioppo, 2009). Loneliness is a problem that affects many older adults as it is exacerbated by the loss of spouses and close friends as well as by worsening health (Peters & Liefbroer, 1997; Essex & Sunghee, 1987).

The NSHAP Felt Loneliness Measure (NFLM) is nearly identical to the loneliness scale used in HRS in 2002. HRS's scale was adapted from the well-established Revised UCLA Loneliness Scale (R-UCLA) (Russell, Peplau, & Cutrona, 1980). HRS's 3-item loneliness scale asked respondents to rate the frequency of their feelings of loneliness using three frequency response categories (*hardly ever, some of the time,* and *often*). Internal reliability found in the HRS module is .72 (Hughes, Waite, Hawkley, & Cacioppo, 2004). The correlation among the 3-item scale and the larger UCLA loneliness scale is .82 (Hughes et al., 2004).

Methods for NFLM

As in HRS (2002), NFLM asked respondents how often they felt that they lacked companionship, felt left out, or felt isolated from others during the past week. Following HRS, Wave 1 utilized three frequency response categories, from which Wave 2 expanded to four response categories. We recommend collapsing the Wave 2 response categories *never* and *hardly ever* to create one category, *never or hardly ever*, which is identical in Waves 1 and 2 and in HRS (Table 11).

In both waves, NFLM was administered in the LBQ. Scores of 0, 1, and 2 were assigned to each response category, producing a range of 0 to 6 that is comparable to the HRS scale. To determine the prevalence of feeling lonely frequently (Frequently Felt Loneliness [FFL]), we recommend reporting the percentage of participants who scored ≥1 (i.e., who reported feeling lonely more frequently than *hardly ever*).

Results for NFLM

Similar mean estimates of scores for the NFLM scores were observed in Waves 1 and 2 (.99 ± .03 vs. 1.09 ± .03; 44% vs. 49% had FFL). These results were comparable to those observed in HRS in 2002 (.90 ± .03; 42.1% FFL), indicating reasonable external validity. The distributions of results in Waves 1 and 2 were nearly identical (Figure 1), and scores were significantly associated within individuals across 5 years without an effect of being 5 years older (Table 6). Reliability coefficients for NFLM were also comparable between waves (.80 in Wave 1 and .79 in Wave 2; Table 5).

Based on inspection of Wave 2 graphical results (Figure 2E), any gender differences in NFLM scores would likely occur between the ages of 62 and 69 and/or 80 and 90, when women had slightly higher levels. Overall, however, the interaction between gender and age was not statistically significant in this preliminary linear regression (gender × age coef = –.32 ± .20, t = –1.55 p = .128, gender ×age^2 coef = .002 ± .001, t = 1.57, p = .123, gender coef = 11.76 ± 7.58, t = 1.55, p = .127; age and age^2 $p \geq .510$).

Table 11. Comparison of Symptom Frequency Response Categories Used by Four Short-Form Felt Loneliness Scales

HRS 2002	NFLM	Wave 1	NFLM Wave 2 questionnaire	NFLM Wave 2 analyses
Hardly ever	Hardly ever (or never)		Never Hardly ever	Never or hardly ever
Some of the time	Some of the time		Some of the time	Some of the time
Often	Often		Often	Often

Notes. HRS = Health and Retirement Study; NFLM = NSHAP's Felt Loneliness Measure; NSHAP = National Social Life, Health, and Aging Project.

Discussion

Validity and Stable Individual Differences

After applying the recommended revisions in scoring and cutpoints described in this article, NSHAP's estimates of national prevalences and averages for each facet of mental health are comparable to those observed in other similar samples, including other nationally representative survey studies conducted in the home and related to aging in older adults. The legitimacy of NSHAP's mental health measures was further evidenced by the consistency of their reliability coefficients in Waves 1 and 2 (Table 5). Finally, for all five mental health measures, those who participated in both waves displayed frequencies of mental health symptoms that were significantly correlated across waves, indicating relatively stable individual differences during 5 years.

Longitudinal Analyses

Our longitudinal analyses were designed to test the efficacy of NSHAP's mental health measures, but they also enabled us to assess broad age-related changes over 5 years among participants who were evaluated in both waves. Aging 5 years between waves did not significantly change the symptom frequencies of unhappiness, perceived stress, and felt loneliness for the average participant reinterviewed in Wave 2. In contrast, the frequencies of anxiety and depressive symptoms were more prevalent at older ages. Future analyses can profitably determine whether these changes for the average participant can be generalized to diverse subgroups distinguished by such traits as gender, education, and social context.

In addition to age, there were significant wave effects for anxiety symptoms, perceived stress, and depressive symptoms that analysts need to consider. Anxiety symptoms and perceived stress became more frequent in Wave 2 compared with Wave 1, independent of age. In contrast, the frequency

of depressive symptoms decreased slightly from Wave 1 to Wave 2, despite an increase with age (Table 6). There was evidence to evaluate at least three explanations for these wave effects.

First, the mode of NASM and NPSM administration differed between waves. In Wave 2, each measure was administered in the LBQ, whereas participants were randomly selected to receive the questions in either the LBQ or the in-person, CAPI in Wave 1 (Table 8). Analyses of administration mode effects in Wave 1 responses revealed that the LBQ measures were associated with more frequent anxiety and stress symptoms than the CAPI measures. Participants may be more willing to reveal poor mental health conditions on confidential questionnaires than to field interviewers with whom they have a social relationship (Bowling, 2005). Thereby, the mode of administration likely contributed to reports of more frequent anxiety and stress symptoms in Wave 2. In contrast, the drop in depressive symptoms between waves cannot be attributed to mode effects, as this measure was administered to all respondents in the CAPI in both waves (Table 8).

Second, longitudinal participants who reported higher frequencies of anxiety and stress symptoms in Wave 2 than Wave 1 may have been affected by the time periods in which the surveys were administered. For instance, there was an economic recession between waves. Were anxiety and stress symptoms more frequent during 2010–2011 (Wave 2) because the American society was in the early, uncertain stages of the housing recession and people were anxious about its consequences? If so, why were depressive symptoms not also more frequent in Wave 2? It may be because people had yet to experience the prolonged helplessness and loss that can increase depressive symptoms (Dooley et al. 2000; Mirowsky & Ross, 2001). Another possible explanation deserving analysis is retirement. Do anxiety and stress symptoms become more frequent as people approach retirement? If so, participants 57–62 years of age in Wave 1 likely reported more frequent symptoms when they were approaching retirement age 5 years later in Wave 2.

Third, Waves 1 and 2 had different samples, which is critical to consider, especially in analyses that included all participants in both waves (e.g., quasi-independent *t* tests) and in those focused on variables not affected by adjusting sample weights. Participants in Wave 1 that went on to be reinterviewed in Wave 2 had better mental health in Wave 1 than those who were not reinterviewed because they were deceased or too sick to interview (Table 1). Therefore, the increase in anxiety and perceived stress in Wave 2 is unlikely to have been caused by selecting those with better mental health. Moreover, the mental health of returning participants was indistinguishable from the new Wave 2 participants, namely partners of returning participants and Wave 1 nonrespondents added to Wave 2 (Table 2). Again, the sampling hypothesis cannot account for the increases in anxiety symptoms and perceived stress.

Taken together, there is evidence that several factors altered the symptom frequencies of anxiety, perceived stress, and depression between waves. These same factors most likely influenced other NSHAP measures as well. Thus, analysts may need to control for them in future studies, and researchers should explore the questions presented earlier when using NSHAP's nationally representative survey of community-dwelling older adults to further elucidate the aging process.

Age and Gender in Wave 2

Our Wave 2 analyses of age differences revealed important patterns that future researchers are encouraged to investigate, given that this article's purpose is primarily descriptive. First, our results show that grouping participants by decade age categories can be misleading. For instance, while the well-known gender difference in depressive symptoms (Gove & Tudor, 1973; Kohn, Dohrenwend, & Mirotznik, 1998; Link & Dohrenwend, 1980; Mirowsky & Ross, 2003; Steffick et al., 2000) was reproduced in NSHAP's Wave 2 sample, with women experiencing symptoms significantly more frequently than men, the gender difference appeared to occur particularly between 67 and 79 years of age (Figure 2A). The classic age groups based on decade—such as 60–69, 70–79, and 80–89—presume a linear process and also do not necessarily correspond with the ages at which change occurs. Thus, these age categories can obscure significant gender differences in aging patterns, and if one has questions about gender differences in mental health with age, the answers may be different at different ages, e.g., at 62, 75, or 90 years of age.

Our second interesting finding was that nearly all aspects of mental health were worse at older ages. There is a literature establishing an increase in depressive symptoms with age (Blazer, Burchett, Service, & George, 1991; Ferraro & Wilkinson, 2013; Yang, 2007), and NSHAP extends this association to other aspects of mental health, broadly conceptualized as unhappiness, anxiety symptoms, perceived stress, and felt loneliness. The exception was a decrease in the prevalence of anxiety symptoms among the oldest women (Figure 2).

Gender differences in these age patterns were seen graphically for all measures and were statistically significant or clear trends. Our cross-sectional data showed that the age-specific frequencies of depressive symptoms and unhappiness increased steadily among women while men's symptoms did not become more prevalent until ages in the mid-70s. This pattern is consistent with a robust literature that shows that women experience depressive symptoms more frequently than men (Gove & Tudor, 1973; Kohn et al., 1998; Link & Dohrenwend, 1980; Mirowsky & Ross, 2003; Steffick et al., 2000).

Gender differences were most distinct for anxiety symptoms. Men showed the same age-specific pattern as they did in other mental health components, with the prevalence of symptoms beginning to increase among men in their mid-70s. In contrast, women of this age had levels of anxiety

symptoms that were higher than those at older or younger ages. Ninety-year-old women reported nearly the same frequencies of symptoms as 65-year-old women (Figure 2C).

The different age-specific patterns strongly suggest that there are separate components of mental health. Depressive symptoms and unhappiness share such strikingly similar age-specific patterns that they may make up a single component of mental health (Figure 2A and B). Depression and an unhappy mood have also been found to be associated in younger adults (Williams, Teasdale, Segal, & Kabat-Zinn, 2007). Further, the distinct gender differences in age-specific patterns for anxiety symptoms suggest that anxiety symptoms are a part of a mental health component distinct from depressive symptoms; however, like depressive symptoms, anxiety symptoms became more prevalent among longitudinal participants over a 5-year period. Both perceived stress and felt loneliness were similar graphically to depressive symptoms and unhappiness, so they may be part of the same facet of mental health. On the other hand, perceived stress had wave effects similar to those for anxiety symptoms whereas felt loneliness did not, indicating that the components are likely separable.

In sum, these results provide robustness for a construct underlying depressive symptoms and unhappiness and, thereby, provide power for determining their associations with physical health and social networks. Determining whether the mental health measures in NSHAP describe different facets of mental health requires further analysis. Mental health is a property of an individual at a particular age. Therefore, social and population analyses may profit by considering the specific age of each individual, characterizing individuals in terms of the relative frequencies of all facets of mental health, and entering the measures separately into a statistical model to test the working hypotheses described here.

Cautions regarding reporting and interpretation

NSHAP cautions users of its mental health measures to avoid overinterpreting results. For instance, NSHAP's measures are not identical to the original scales from which they were derived, although we are confident that the scoring systems detailed here do provide measures comparable to the literature and to other large survey studies of older adults. NSHAP's measures focus exclusively on symptom frequency rather than intensity or morbidity. Because they were designed for a streamlined interview, some of NSHAP's measures include numbers of items and response categories that are different from the original scales. The adverbs and phrases used to describe each symptom and frequency response category may differ from the original scales, along with reference time frames or lead ins. Therefore, it is important not to refer to NSHAP's measures simply with the title of the original, well-validated scales, although the variable labels in the NSHAP data set do use these acronyms (e.g., *cesd, hads, pss,* and *uclalonely*).

In addition, NSHAP's measures and the scales on which they are based were not designed to diagnose individuals with major or minor depression or with DSM-IV anxiety disorders; therefore, these terms should not be used when interpreting data. The original CES-D and HADS-A are well-validated as screening instruments to help identify people with clinically relevant symptoms that put them at risk for important disorders and warrant further psychiatric or neuropsychological testing for diagnosis. In contrast, NSHAP selected a nationally representative sample of home-dwelling older adults, in which clinical mental health problems are less common than in a clinical population. Our goal was to assess variation in frequency of symptomatology, primarily within the normal range.

Although NDSM and NASM have support for their external validity, as our recommended cutpoints yield prevalences similar to literature reports, it is not appropriate to report NSHAP's prevalences as clinically relevant symptoms of depression or anxiety. Rather, NSHAP's measures estimate the prevalence of people who frequently suffer from depressive or anxiety symptoms (FDS and FAS, respectively). Only a subset of participants who report frequent symptoms within the past week are likely to have diagnosable mental health disorders. Thus, we do not recommend that NDSM or NASM results be interpreted in terms of specific depressive or anxiety disorders.

During the development of NDSM, we identified common inaccuracies in the CES-D literature that can be avoided. For example, we found that using a proportional method to convert cutpoints from the 20-item CES-D scale to a shortened form is not accurate, confirming Kohout and colleagues (1993), who provided a more accurate regression formula. Moreover, the literature interpreting the CES-D scale uses inconsistent depression terminology and often overstates results found through the application of modified forms of the CES-D. Here we provided guidelines for interpretation that are consistent with psychiatric diagnoses of different types of depression, as distinct from a depressed mood.

When considering NHUM and comparing our findings to the extant literature, it is imperative to recognize that any measure related to unhappiness, happiness, or both is controversial. Researchers do not agree if different types of positive and negative affect are bipolar (Russell, 1980; Russell & Carroll, 1999) or bivariate constructs (Cacioppo & Berntson, 1994; Larsen, McGraw, & Cacioppo, 2001); thus, it is unclear whether a bipolar scale (happiness–unhappiness) or two unipolar scales (one for happiness and another for unhappiness) are appropriate. Many survey studies use a single measure of happiness. For instance, the SISHS is a unipolar scale that measures the intensity of happy feelings and is often interpreted as a measure of life satisfaction (Lee & Bulanda, 2005). Despite extensive validations of the SISHS scale, it is only one item, and even those who argue that unhappiness and happiness are bipolar variables agree that two questions are needed to validate unipolar scales by avoiding having respondents

misinterpret the univariate happiness scale as bipolar (Russell & Carroll, 1999). Two questions are also ideal because happiness is often defined differently, especially by men and women and by members of different ethnic groups.

In contrast, NSHAP intended to measure both unhappy and happy feelings, so we developed a single bipolar scale that helped conserve interview time. One cautionary note, however, is that the referent time of NHUM may be considered contradictory (an extended period of time (*[consider] your life in general…[and] on the whole*) vs. recently (*these days)*). Also, the scale measures intensity of happiness (*extremely happy, very happy, pretty happy*) but frequency of unhappiness (*unhappy sometimes, unhappy usually*). Thus, the scale is not truly bipolar because it does not measure the two variables in the same context or through parallel response category structure. In addition, NHUM differs from most literature scales that use "sad" as the contrast to "happy" (Larsen et al., 2001; Russell & Carroll, 1999; Watson & Tellegen, 1985). Nevertheless, inversely scored, NHUM manifests the same gender differences and age patterns seen for depressive symptomatology (NDSM), supporting its use as an associated measure of dysphoria. This does not detract from using NHUM in other contexts as a measure of a positive state, especially because it is NSHAP's only measure of overall life satisfaction.

Conclusion

The recommended scoring protocols for NSHAP's unique mental health measures based on symptom frequency yield prevalences and means of mental health symptoms comparable to similar national studies of older adults, which demonstrate the external validity of NSHAP's measures. Although scores on each mental health measure were correlated within individuals between Waves 1 and 2, indicating stability of individual differences over time, future longitudinal analysts of mental health symptoms should be alert to sources of differences between waves. For example, NASM and NPSM detected higher symptom frequencies in Wave 2 as a result of mode of administration and time period. Differences in sample composition should certainly be considered, although they proved to be an unlikely explanation in the case of mental health measures.

Wave 2 analyses showed that gender differences and age-specific patterns are more accurate and interesting when age is defined as a continuous variable, despite how many studies focus on analyzing data in terms of decade age categories. Analyzing depressive symptoms scores with age as a continuous variable revealed that the commonly reported gender difference in depressive symptoms may not exist at all ages among older U.S. adults. In addition, depressive symptoms and unhappiness had similar age patterns in Wave 2, indicating that future analyses may determine whether they are aspects of one mental health component. Perceived stress showed an age-specific pattern similar to that of depression, yet wave effects similar to anxiety. Anxiety symptoms and felt loneliness showed gender differences in age patterns and may profitably be treated as different mental health components. Taken together, the NSHAP mental health measures allow a rich characterization of the mental health of older U.S. adults and may potentially mediate the reciprocal dynamic between social life and physical health during aging.

Key Points

- Evidenced-based cutpoints for "frequent symptoms" are established for NSHAP's depressive symptoms and anxiety symptoms measures as ≥ 9 and ≥ 8, respectively.
- NSHAP's reported average scores and prevalences for its mental health measures are comparable to other nationally representative studies of older adults, which supports the external validity of NSHAP's unique mental health measures. Significant within-person correlations across waves demonstrate moderate stability of mental health characteristics across five years.
- The increases in anxiety symptoms and perceived stress from Wave 1 to Wave 2 largely reflect changes in administration mode and time period, in addition to aging five years. Analytic strategies for handling these issues are presented.
- Wave 2 scores on all five mental health measures are higher for older participants among both men and women, with the exception of women having less frequent anxiety symptoms at older than younger ages. Women also generally scored higher than men.
- Gender differences in age-specific patterns are evident when age is used as a continuous variable and are obscured by using the traditional decade age categories.

Supplementary Material

Supplementary material can be found at: http://psychsocgerontology.oxfordjournals.org/

Funding

NSHAP was funded by the National Institutes of Health, including the Institute on Aging (R01AG033903, R37AG030481, and R01AG021487, Linda J. Waite, principal investigator), the Office of Women's Health Research, the Office of AIDS Research, and the Office of Behavioral and Social Sciences Research and by NORC, which was responsible for the data collection. M. Kozloski was also supported by the National Institute on Aging Postdoctoral Training Program (T32 AG00243) through the Center on the Demography and Economics of Aging.

Acknowledgements

We thank Kristen E. Wroblewski for the analyses in Tables 1 and 2.

Correspondence

Correspondence should be addressed to Martha K. McClintock, PhD, 940 East 57th Street, The Institute for Mind and Biology, The University of Chicago, Chicago, Illinois 60637. Email: marthasecret2003@yahoo.com.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: American Psychiatric Association.

Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: America's perception of life quality*. New York, NY: Plenum Press.

Argyle, M. (1999). Causes and correlates of happiness. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 353–372). New York, NY: Russell Sage Foundation.

Baum, A., & Posluszny, D. M. (1999). Health psychology: Mapping biobehavioral contributions to health and illness. *Annual Review of Psychology*, *50*, 137–163. doi:10.1146/annurev.psych.50.1.137

Beekman, A. T. F., Bremmer, M. A., Deeg, D. J. H., Van Balkom, A. J. L. M., Smit, J. H., Beurs, E. D., … Van Tilburg, W. (1998). Anxiety disorders in later life: A report from the longitudinal aging study Amsterdam. *International Journal of Geriatric Psychiatry*, *13*, 717–726. doi:10.1002/(SICI)1099-1166(1998100)13:10<717::AID-GPS857>3.0.CO;2-M

Beekman, A. T., Deeg, D. J., van Tilburg, T., Smit, J. H., Hooijer, C., & van Tilburg, W. (1995). Major and minor depression in later life: A study of prevalence and risk factors. *Journal of Affective Disorders*, *36*, 65–75. doi:10.1016/0165-0327(95)00061-5

Berkman, L., Berkman, C., Kasl, S., Freeman, D. H., Leo, L., Ostfeld, A. M., … Brody, J. A. (1986). Depressive symptoms in relation to physical health and functioning in the elderly. *American Journal of Epidemiology*, *124*, 372–388.

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale. An updated literature review. *Journal of Psychosomatic Research*, *52*, 69–77.

Blazer, D. G. (2003). Depression in late life: Review and commentary. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *58*, 249–265. doi:10.1093/gerona/58.3.M249

Blazer, D., Burchett, B., Service, C., & George, L. K. (1991). The association of age and depression among the elderly: An epidemiologic exploration. *Journal of Gerontology*, *46*, M210–M215. doi:10.1093/geronj/46.6.M210

Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health*, *27*, 281–291. doi:10.1093/pubmed/fdi031

Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago, IL: Aldine.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*, 401–423. doi:10.1037/0033-2909.115.3.401

Cacioppo, J. T., Hawkley, L. C., & Thisted, R. A. (2010). Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychology and Aging*, *25*, 453–463. doi:10.1037/a0017216

Cacioppo, J. T., Hughes, M. E., Waite, L. J., Hawkley, L. C., & Thisted, R. A. (2006). Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses. *Psychology and Aging*, *21*, 140–151. doi:10.1037/0882-7974.21.1.140

Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). The situation of women. In A. Campbell, P. E. Converse, & W. L. Rodgers (Eds.), *The finality of American life* (pp. 395–492). New York, NY: Russell Sage Foundation.

Clemente, F., & Sauer, W. J. (1976). Life satisfaction in the United States. *Social Forces*, *54*, 621–631. doi:10.2307/2576286

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*, 385–396. doi:10.2307/2136404

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The social psychology of health: Claremont Symposium on Applied Social Psychology* (pp. 31–67). Newbury Park, CA: Sage.

Cornwell, B. (2003). The dynamic properties of social support: Decay, growth, and staticity, and their effects on adolescent depression. *Social Forces*, *81*, 953–978. doi:10.1353/sof.2003.0029

Denollet, J., Maas, K., Knottnerus, A., Keyzer, J. J., & Pop, V. J. (2009). Anxiety predicted premature all-cause and cardiovascular death in a 10-year follow-up of middle-aged women. *Journal of Clinical Epidemiology*, *62*, 452–456. doi:10.1016/j.jclinepi.2008.08.006

Dockray, S., & Steptoe, A. (2010). Positive affect and psychobiological processes. *Neuroscience & Biobehavioral Reviews*, *35*, 69–75. Retrieved fromhttp://www.sciencedirect.com/science/article/pii/S0149763410000072

Dooley, D., Prause, J., & Ham-Rowbottom, K. A. (2000). Underemployment and depression: Longitudinal relationships. *Journal of Health and Social Behavior*, *41*, 421–436. doi:10.2307/2676295

Essex, M. J., & Sunghee, N. (1987). Marital status and loneliness among older women: The differential Importance of close family and friends. *Journal of Marriage and Family*, *49*, 93–106. doi:10.2307/352674

Farmer, M. M., & Ferraro, K. F. (1997). Distress and perceived health: Mechanisms of health decline. *Journal of Health and Social Behavior*, *38*, 298–311. doi:10.2307/2955372

Ferraro, K. F., & Wilkinson, L. R. (2013). Chapter 10: Age, aging, and mental health. In C. S. Aneshensel, J. C. Phelan, & A. Bierman (Eds.), *Handbook of the sociology of mental health (2nd ed.,* pp. 183–203). New York, NY: Springer.

Flint, A. J. (1994). Epidemiology and comorbidity of anxiety disorders in the elderly. *The American Journal of Psychiatry*, *151*, 640–649.

Frasure-Smith, N., & Lespérance, F. (2008). Depression and anxiety as predictors of 2-year cardiac events in patients with stable coronary artery disease. *Archives of General Psychiatry*, *65*, 62–71. doi:10.1001/archgenpsychiatry.2007.4

Frazier, L. D., & Waid, L. D. (1999). Influences on anxiety in later life: The role of health status, health perceptions, and health locus of control. *Aging & Mental Health*, *3*, 213–220. doi:10.1080/13607869956163

Gehi, A., Haas, D., Pipkin, S., & Whooley, M. A. (2005). Depression and medication adherence in outpatients with coronary heart disease: Findings from the Heart and Soul Study. *Archives of Internal Medicine*, *165*, 2508–2513. doi:10.1001/archinte.165.21.2508

Gove, W. R., & Tudor, J. (1973). Adult sex roles and mental illness. *American Journal of Sociology*, *78*, 50–73. doi:10.1086/225404

Hamarat, D., Thompson, K. M., Zabrucky, D., Steele, K. B., Matheny, F., & Aysan, E. (2001). Perceived stress and coping resource availability as predictors of life satisfaction in young, middle-aged, and older adults. *Experimental Aging Research*, *27*, 181–196.

Hawkley, L. C., Thisted, R. A., & Cacioppo, J. T. (2009). Loneliness predicts reduced physical activity: Cross-sectional & longitudinal analyses. *Health Psychology*, *28*, 354–363. doi:10.1037/a0014400

Herrmann, C. (1997). International experiences with the Hospital Anxiety and Depression Scale—A review of validation data and clinical results. *Journal of Psychosomatic Research*, *42*, 17–41.

Himmelfarb, S., & Murrell, S. A. (1984). The prevalence and correlates of anxiety symptoms in older adults. *The Journal of Psychology*, *116*, 159–167. doi:10.1080/00223980.1984.9923632

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. doi:10.1177/0164027504268574

Jackowska, M., Kumari, M., & Steptoe, A. (2013). Sleep and biomarkers in the English Longitudinal Study of Ageing Associations with C-reactive protein, fibrinogen, dehydroepiandrosterone sulfate and hemoglobin. *Psychoneuroendocrinology, 8*, 1484–1493. doi:10.1016/j.psyneuen.2012.12.015. Retrieved from http://www.sciencedirect.com/science/article/pii/S0306453012004283

Jood, K., Redfors, P., Rosengren, A., Blomstrand, C., & Jern, C. (2009). Self-perceived psychological stress and ischemic stroke: A case–control study. *BMC Medicine*, *53*. Retrieved from http://www.biomedcentral.com/1741–7015/7/53. doi:10.1186/1741-7015-7-53

Kawachi, I., Sparrow, D., Vokonas, P. S., & Weiss, S. T. (1994). Symptoms of anxiety and risk of coronary heart disease. The Normative Aging Study. *Circulation*, *90*, 2225–2229. doi:10.1161/01.CIR.90.5.2225

Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*, 87–98. doi:10.1016/j.jbtep.2007.02.003

Kessler, R. C., McGonagle, K. A., Swartz, M., Blazer, D. G., & Nelson, C. B. (1993). Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. *Journal of Affective Disorders*, *29*, 85–96. doi:10.1016/0165-0327(93)90026-G

Kirby, S. E., Coleman, P. G., & Daley, D. (2004). Spirituality and well-being in frail and nonfrail older adults. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *59*, 123–129. doi:10.1093/geronb/59.3.P123

Klerman, G. L., & Weissman, M. M. (1989). Increasing rates of depression. *The Journal of the American Medical Association*, *261*, 2229–2235. doi:10.1001/jama.261.15.2229

Kohn, R., Dohrenwend, B. P., & Mirotznik, J. (1998). Epidemiological findings on selected psychiatric disorders in the general population. In B. P. Dohrenwend (Ed.), *Adversity, stress, and psychopathology* (pp. 235–284). New York, NY: Oxford University Press.

Kohout, F. J., Berkman, L. F., Evans, D. A., & Cornoni-Huntley, J. (1993). Two shorter forms of the CES-D depression symptoms index. *Journal of Aging and Health*, *5*, 179–193. doi:10.1177/089826439300500202

Larsen, J. T., McGraw, A. P., & Cacioppo, J. T. (2001). Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, *81*, 684–696. doi:10.1037//0022-3514.81.4.684

Lawton, M. P., Brody, E. M., & Saperstein, A. R. (1989). A controlled study of respite service for caregivers of Alzheimer's patients. *The Gerontologist*, *29*, 8–16.

Lee, G. R., & Bulanda, J. R. (2005). Change and consistency in the relation of marital status to personal happiness. *Marriage and Family Review*, *38*, 69–84. doi:10.1300/J002v38n01_06

Lin, N., & Ensel, W. M. (1984). Depression-mobility and its social etiology: The role of life events and social support. *Journal of Health and Social Behavior*, *25*, 176–188. doi:10.2307/2136667

Lin, E. H. B., Katon, W., Von Korff, M., Rutter, C. R., Simon, G. E., Oliver, M., … Young, B. (2004). Relationship of depression and diabetes self-care, medication adherence, and preventive care. *Diabetes Care*, *27*, 2154–2160. doi:10.2337/diacare.27.9.2154

Link, B., & Dohrenwend, B. (1980). Formulation of hypotheses about the true prevalence of demoralization in the United States. In B. P. Dohrenwend, B. S. Dohrenwend, M. S. Gould, B. Link, R. Neugebauer, & R. Wunsch-Hitzig (Eds.), *Mental illness in the United States: Epidemiological estimates* (pp. 114–132). New York, NY: Praeger.

Marcus, M., Yasamy, M. T., van Ommeren, M., Chisholm, D., & Saxena, S. (2012). *Depression: A global public health concern*. World Health Organization Department of Mental Health and Substance Abuse. Retrieved from http://www.who.int/ mental_health/management/depression/who_paper_depression_wfmh_2012.pdf

McDade, T. W., Hawkley, L. C., & Cacioppo, J. T. (2006). Psychosocial and behavioral predictors of inflammation in middle-aged and older adults: the Chicago health, aging, and social relations study. *Psychosomatic Medicine*, *68*, 376–381. doi:10.1097/01.psy.0000221371.43607.64

Mehta, K. M., Simonsick, E. M., Penninx, B. W., Schulz, R., Rubin, S. M., Satterfield, S., & Yaffe, K. (2003). Prevalence and correlates of anxiety symptoms in well-functioning older adults: Findings from the health aging and body composition study. *Journal of the American Geriatrics Society*, *51*, 499–504. doi:10.1046/j.1532-5415.2003.51158.x

Metti, A. L., Cauley, J. A., Newman, A. B., Ayonayon, H. N., Barry, L. C., Kuller, L. M., … Yaffe, K. (2013). Plasma beta amyloid level and depression in older adults. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *68*, 74–79. doi:10.1093/gerona/gls093

Mirowsky, J., & Ross, C. E. (1992). Age and depression. *Journal of Health and Social Behavior*, *33*, 187–205. doi:10.2307/2137349

Mirowsky, J., & Ross, C. E. (2001). Age and the effect of economic hardship on depression. *Journal of Health and Social Behavior*, *42*, 132–150. doi:10.2307/3090174

Mirowsky, J., & Ross, C. (2003). *Social causes of psychological distress* (2nd ed.). New York, NY: Aldine de Gruyter.

Moussavi, S., Chatterji, S., Verdes, E., Tandon, A., Patel, V., & Ustun, B. (2007). Depression, chronic diseases, and decrements in health: Results from the World Health Surveys. *Lancet*, *370*, 851–858. doi:10.1016/S0140-6736(07)61415-9

Ng, D. M., & Jeffery, R. W. (2003). Relationships between perceived stress and health behaviors in a sample of working adults. *Health Psychology*, *22*, 638–642. doi:10.1037/0278-6133.22.6.638

Peplau, L. A., & Perlman, D. (1982). *Loneliness: A sourcebook of current theory, research and therapy*. New York, NY: Wiley Interscience.

Peters, A., & Liefbroer, A. C. (1997). Beyond marital status: Partner history and well-being in old age. *Journal of Marriage and Family*, *59*, 687–699. doi:10.2307/353954

Potvin, O., Forget, H., Grenier, S., Préville, M., & Hudon, C. (2011). Anxiety, depression, and 1-year incident cognitive impairment in community-dwelling older adults. *Journal of the American Geriatrics Society*, *59*, 1421–1428. doi:10.1111/j.1532-5415.2011.03521.x

Pozuelo, L., Tesar, G., Zhang, J., Penn, M., Franco, K., & Jiang, W. (2009). Depression and heart disease: What do we know, and where are we headed? *Cleveland Clinic Journal of Medicine*, *76*, 59–70. doi:10.3949/ccjm.75a.08011

Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*, 385–401. doi:10.1177/014662167700100306

Regier, D. A., Boyd, J. H., Burke, J. D., Rae, D. S., Myers, J. K., Kramer, … Locke, B. Z. (1988). One-month prevalence of mental disorders in the United States. *Archives of General Psychiatry*, *45*, 977–986. doi:10.1001/archpsyc.1988.01800350011002

Ritchey, F. J., La Gory, M., Fitzpatrick, K. M., & Mullis, J. (1990). A comparison of homeless, community-wide, and selected distressed samples on the CES-Depression Scale. *American Journal of Public Health*, *80*, 1384–1386. doi:10.2105/AJPH.80.11.1384

Rosenbaum, J. F., Pollack, M. H., Otto, M. W., et al. (1997). In N. H. Cassem, T. A. Stern, J. F. Rosenbaum, et al. (Eds.), *The Massachusetts general handbook of general hospital psychiatry* (4th ed.). St. Louis, MO: Mosby-Year Book.

Rosengren, A., Tibblin, G., & Wilhelmsen, L. (1991). Self-perceived psychological stress and incidence of coronary artery disease in middle-aged men. *The American Journal of Cardiology*, *68*, 1171–1175. doi:10.1016/0002-9149(91)90189-R

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178. doi:10.1037/h0077714

Russell, J. A., & Carroll, J. M. (1999). On the bipolarity of positive and negative affect. *Psychological Bulletin*, *125*, 3–30. doi:10.1037/0033-2909.125.1.3

Russell, D., 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. doi:10.1037/0022-3514.39.3.472

Schnittker, J. (2008). Diagnosing our national disease: Trends in income and happiness, 1973 to 2004. *Social Psychology Quarterly*, *71*, 257–280. doi:10.1177/019027250807100307

Sharp, L. K., & Lipsky, M. S. (2002). Screening for depression across the lifespan: A review of measures for use in primary care settings. *American Family Physician*, *66*, 1001–1008.

Sherbourne, C. D., Wells, K. B., Meredith, L. S., Jackson, C. A., & Camp, P. (1996). Comorbid anxiety disorder and the functioning and well-being of chronically ill patients of general medical providers. *Archives of General Psychiatry*, *53*, 889–895. doi:10.1001/archpsyc.1996.01830100035005

Simon, G., Ormel, J., VonKorff, M., & Barlow, W. (1995). Health care costs associated with depressive and anxiety disorders in primary care. *The American Journal of Psychiatry*, *152*, 352–357.

Snaith, R. P. (2003). The Hospital Anxiety and Depression Scale. *Health and Quality of Life Outcomes*. Retrieved from http://www.hqlo.com/content/1/1/29

Stack, S. (1998). Marriage, family and loneliness: A cross-national study. *Sociological Perspectives*, *41*, 415–432. doi:10.2307/1389484

Stack, S., & Eshleman, J. R. (1998). Marital status and happiness: A 17-nation study. *Journal of Marriage and Family*, *60*, 527–536. doi:10.2307/353867

Steffick, D. E., Wallace, R. B., Herzog, A. R., Ofstedal, M. B., Steffick, D., Fonda, S., & Langa, K. (2000). *HRS/AHEAD documentation report: Documentation of affective functioning measures in the Health and Retirement Study*. Ann Arbor, MI: Survey Research Center at the University of Michigan.

Tay, S. C. (2011). *The psychometric principles of affect: Are they ideal? (Unpublished doctoral dissertation)*. University of Illinois at Urbana-Champaign, Urbana, IL.

Van Hout, H. P. J., Beekman, A. T. F., De Beurs, E., Comijs, H., Van Marwijk, H., De Haan, M., … Deeg, D. J. H. (2004). Anxiety and the risk of death in older men and women. *The British Journal of Psychiatry*, *185*, 399–404. doi:10.1192/bjp.185.5.399

Veenhoven, R., & Hagerty, M. (2006). Rising happiness in nations 1946–2004: A reply to Easterlin. *Social Indicators Research*, *79*, 421–436. doi:10.1007/s11205-005-5074-x

Vermeulen, A. W. A., Beekman, A. T. F., & Stek, M. L. (1994). The prevalence of anxiety disorders in later life: A review. *Tijdschr Psychiatr*, *36*, 657–668.

Waite, L. J., & Gallagher, M. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially*. New York, NY: Doubleday.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*, 219–235. doi:10.1037/0033-2909.98.2.219

Wauterickx, N., & Bracke, P. (2005). Unipolar depression in the Belgian population. *Social Psychiatry and Psychiatric Epidemiology*, *40*, 691–699. doi:10.1007/s00127-005-0928-8

Williams, M., Teasdale, J., Segal, Z., & Kabat-Zinn, J. (2007). *Mindful way through depression (enhanced): Freeing yourself from chronic unhappiness*. New York, NY: The Guilford Press.

Wilson, R. S., Krueger, K. R., Arnold, S. E., Schneider, J. A., Kelly, J. F., Barnes, L. L., … Bennett, D. A. (2007). Loneliness and risk of Alzheimer's disease. *Archives of General Psychiatry*, *64*, 234–240.

Wiltink, J., Tschan, R., Michal, M., Subic-Wrana, C., Eckhardt-Henn, A., Dieterich, M., & Beutel, M. E. (2009). Dizziness: Anxiety, health care utilization and health behavior—Results from a representative German community survey. *Journal of Psychosomatic Research*, *66*, 417–424.

Yang, Y. (2007). Is old age depressing? Growth trajectories and cohort variations in late-life depression. *Journal of Health and Social Behavior*, *48*, 16–32. doi:10.1177/002214650704800102

Yang, Y. (2008). Social inequalities in happiness in the United States, 1972 to 2004: An age-period-cohort analysis. *American Sociological Review*, *73*, 204–226. doi:10.1177/000312240807300202

Zauszniewski, J. A., & Bekhet, A. K. (2009). Depressive symptoms in elderly women with chronic conditions: Measurement issues. *Aging & Mental Health*, *13*, 64–72. doi:10.1080/13607860802154481

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370. doi:10.1111/j.1600-0447.1983.tb09716.x