



10-16-2014

## Dimensionality and Reliability of the Self-Care of Heart Failure Index Scales: Further Evidence from Confirmatory Factor Analysis

Claudio Barbaranelli


Christopher S. Lee

Ercole Vellone

Barbara Riegel

*University of Pennsylvania*, [briegel@nursing.upenn.edu](mailto:briegel@nursing.upenn.edu)

Follow this and additional works at: <https://repository.upenn.edu/nrs>

 Part of the [Cardiology Commons](#), [Cardiovascular Diseases Commons](#), [Circulatory and Respiratory Physiology Commons](#), [Health Services Research Commons](#), [Medical Humanities Commons](#), [Nursing Commons](#), and the [Preventive Medicine Commons](#)

---

### Recommended Citation

Barbaranelli, C., Lee, C. S., Vellone, E., & Riegel, B. (2014). Dimensionality and Reliability of the Self-Care of Heart Failure Index Scales: Further Evidence from Confirmatory Factor Analysis. *Research in Nursing & Health*, 37 (6), 524-537. <http://dx.doi.org/10.1002/nur.21623>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/nrs/112>  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

## Dimensionality and Reliability of the Self-Care of Heart Failure Index Scales: Further Evidence from Confirmatory Factor Analysis

### Abstract

The Self-Care of Heart Failure Index (SCHFI) is used widely, but issues with reliability have been evident. Cronbach alpha coefficient is usually used to assess reliability, but this approach assumes a unidimensional scale. The purpose of this article is to address the dimensionality and internal consistency reliability of the SCHFI. This was a secondary analysis of data from 629 adults with heart failure enrolled in three separate studies conducted in the northeastern and northwestern United States. Following testing for scale dimensionality using confirmatory factor analysis, reliability was tested using coefficient alpha and alternative options. Confirmatory factor analysis demonstrated that: (a) the Self-Care Maintenance Scale has a multidimensional four-factor structure; (b) the Self-Care Management Scale has a two-factor structure, but the primary factors loaded on a common higher-order factor; and (c) the Self-Care Confidence Scale is unidimensional. Reliability estimates for the three scales, obtained with methods compatible with each scale's dimensionality, were adequate or high. The results of the analysis demonstrate that issues of dimensionality and reliability cannot be separated. Appropriate estimates of reliability that are consistent with the dimensionality of the scale must be used. In the case of the SCHFI, coefficient alpha should not be used to assess reliability of the self-care maintenance and the self-care management scales, due to their multidimensionality. When performing psychometric evaluations, we recommend testing dimensionality before assessing reliability, as well using multiple indices of reliability, such as model-based internal consistency, composite reliability, and omega and maximal reliability coefficients.

### Keywords

Adult, Factor Analysis, Statistical, Female, Heart Failure, Humans, Male, Middle Aged, Psychometrics, Quality of Life, Reproducibility of Results, Self Care, United States

### Disciplines

Cardiology | Cardiovascular Diseases | Circulatory and Respiratory Physiology | Health Services Research | Medical Humanities | Medicine and Health Sciences | Nursing | Preventive Medicine



Published in final edited form as:

*Res Nurs Health*. 2014 December ; 37(6): 524–537. doi:10.1002/nur.21623.

## Dimensionality and Reliability of the Self-Care of Heart Failure Index Scales: Further Evidence from Confirmatory Factor Analysis

**Claudio Barbaranelli,**

Professor, Department of Psychology, Sapienza University, Rome, Italy

**Christopher S. Lee,**

Associate Professor, School of Nursing, Oregon Health & Science University, Portland, OR, USA

**Ercole Vellone,** and

Research Fellow, Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy

**Barbara Riegel**

Professor and Edith Clemmer Steinbright Chair of Gerontology, School of Nursing, University of Pennsylvania, 418 Curie Boulevard, Philadelphia, PA USA 19104-4217

### Abstract

The Self-Care of Heart Failure Index (SCHFI) is used widely, but issues with reliability have been evident. Cronbach alpha coefficient is usually used to assess reliability, but this approach assumes a unidimensional scale. The purpose of this article is to address the dimensionality and internal consistency reliability of the SCHFI. This was a secondary analysis of data from 629 adults with heart failure enrolled in three separate studies conducted in the northeastern and northwestern United States. Following testing for scale dimensionality using confirmatory factor analysis, reliability was tested using coefficient alpha and alternative options. Confirmatory factor analysis demonstrated that: a) the self-care maintenance scale has a multidimensional 4-factor structure; b) the self-care management scale has a 2-factor structure, but the primary factors loaded on a common higher-order factor; and c) the self-care confidence scale is unidimensional. Reliability estimates for the three scales, obtained with methods compatible with each scale's dimensionality, were adequate or high. The results of the analysis demonstrate that issues of dimensionality and reliability cannot be separated. Appropriate estimates of reliability that are consistent with the dimensionality of the scale must be used. In the case of the SCHFI, coefficient alpha should not be used to assess reliability of the self-care maintenance and the self-care management scales, due to their multidimensionality. We recommend testing dimensionality before assessing reliability, as well using multiple indices of reliability, such as model-based internal consistency, composite reliability, and omega and maximal reliability coefficients.

## Keywords

heart failure; psychometrics; reliability; factor analysis; self-care; measurement; instrument validation

Heart failure (HF) is one of the most common chronic illnesses among adults worldwide (Bui, Horwich, & Fonarow, 2011; Huffman & Prabhakaran, 2010; McDonagh, 2013). Population longevity, better survival rates after acute myocardial infarction (Ford et al., 2007), and the increasing prevalence of hypertension (Ong, Cheung, Man, Lau, & Lam, 2007) have ballooned the incidence of HF in recent years. In the United States, among those older than 80 years of age, almost 12% of men and women have HF (Go et al., 2014) and the prevalence is projected to increase by 25% between 2013 and 2030 (Heidenreich et al., 2013). HF is associated with a high symptom burden, poor quality of life (Banerjee et al., 2013), high mortality, and frequent hospital admissions that contribute enormously to health care costs (Heidenreich et al., 2013). Self-care has been shown to improve symptoms, quality of life, hospitalization rates, cost, and mortality (Jones et al., 2012; Lee, Moser, Lennie, & Riegel, 2011; Lee, Moser, et al., 2011a; Rathman, Lee, Sarkar, & Small, 2011).

Much of our understanding of the contributions of self-care to HF outcomes followed publication of the Self-Care of HF Index (SCHFI; Riegel et al., 2004), a self-report, disease-specific measure of HF self-care based on the situation-specific theory of HF self-care (Riegel & Dickson, 2008). The SCHFI, last updated in 2009 (Riegel, Lee, Dickson, & Carlson, 2009), is freely available in the public domain (<http://www.self-careofheartfailureindex.com/>) and is used worldwide (Jaarsma et al., 2013). However, a consistent finding among investigators using the SCHFI has been marginal or poor reliability of the self-care maintenance and self-care management scales.

Reliability is a reflection of measurement error, defined as “chance effects that are momentary and have nothing to do with the construct being measured” (Raykov & Marcoulides, 2011, p. 116). Internal consistency reliability of the SCHFI has been tested mainly with Cronbach alpha coefficient, which may not be the best approach for an instrument with multiple dimensions. Raykov (1997) demonstrated that alpha is a lower-bound estimate for reliability (or even fairly close to reliability) *only* when 1) the indicators composing the scale are congeneric (i.e., there is a single factor explaining their correlations), 2) their unique variances are not correlated, and 3) their loadings on the single factor are “uniformly high” ( $> |0.6|$ ), that is when items are at least *tau* equivalent (Novick & Lewis, 1967). When these three conditions are not met, alpha can give a biased (upward or downward) estimate of reliability (Maxwell, 1968; Raykov, 1997, 1998, 2001a, 2001b). However, because population reliability of the Index is unknown, there is no way of knowing whether these conditions are met, so alpha can be unpredictable, and reliance on it can be seriously misleading.

Thus, the purpose of this study was to assess dimensionality of the SCHFI and then assess reliability of each scale using the best alternative approaches. In doing so, we produced evidence of the critical importance of considering alpha as a measure of test homogeneity,

rather than measurement error or reliability, in the absence of a thorough assessment of dimensionality.

## Measurement of Self-Care

The situation-specific theory of HF self-care defines self-care as a naturalistic decision making process composed of two dimensions: self-care maintenance and self-care management (Riegel & Dickson, 2008). Self-care maintenance reflects behaviors in which patients engage to maintain physiological stability, including monitoring of signs and symptoms (e.g., checking ankle edema) and adhering to prescribed treatments (e.g., being physically active). Self-care management is a process of recognizing symptoms (e.g., shortness of breath), implementing treatments aimed at addressing the symptoms (e.g., limiting salt intake) and evaluating the effectiveness of the treatment implemented. This process was recently confirmed in an Italian sample (Vellone, Riegel, D'Agostino, et al., 2013c). Self-care confidence (i.e., self-efficacy related to the specific tasks of self-care) is thought to be an important influence on the effectiveness of HF self-care, and there is evidence to support this proposition (Cene et al., 2013; Lee, Moser, et al., 2011b; Lee, Suwanno, & Riegel, 2009).

The SCHFI version 6.2 directly reflects this naturalistic decision-making process and thus is useful to investigators seeking to describe self-care and test the effectiveness of interventions. However, reliability is a continuing issue with the SCHFI. In the earlier version (SCHFI V.4), when all three scales were added to yield a single self-care score, Cronbach alpha was .76 for the full scale, and .56, .70 and .82 for the self-care maintenance, management and confidence scales respectively (Riegel et al., 2004). Similar results have been reported by others. Yu and colleagues found that Cronbach alpha of the aggregated scales was .73; separate scale coefficients were not reported (Yu et al., 2011). Kato and colleagues tested SCHFI v.6.1 and reported Cronbach alpha coefficients of .68, .58, and .87 for the self-care maintenance, management, and confidence scales respectively (Kato et al., 2013). When the SCHFI v.6.2 update was published, Cronbach alpha was reported to be .55 for self-care maintenance, .60 for self-care management, and .83 for self-care confidence (Riegel et al., 2009).

One reason why Cronbach alpha coefficient is low may be that the SCHFI is not unidimensional. The alpha coefficient is an index of the degree to which an instrument's items are correlated on average (Sijtsma, 2009). The SCHFI, however, has three scales. For example, items in the SCHFI self-care maintenance scale reflect the many behaviors noted in clinical guidelines to be important to maintaining physiologic stability (e.g., following a low-salt diet, exercising), but the items reflect more than one dimension of self-care maintenance. Accordingly, items are homogeneous within the specific facets constituting self-care maintenance, but they may not be homogenous when considering the construct as a whole.

## Use and Misuse of Cronbach Alpha Coefficients

Internal consistency has been improperly used as a synonym of homogeneity, that is, of unidimensionality of the scale, and a high alpha coefficient has been used as proof that a

single factor explains correlations among scale items (Raykov, 2012; Raykov & Marcoulides, 2011). However, while homogeneity refers to the fact that only one latent dimension explains the correlations among a set of items composing a scale, internal consistency refers to the fact that items are simply related, but sheds no light on their latent structure (Schmitt, 1996).

Alpha as a measure of internal consistency is basically an index of how much the items are correlated on average, assuming that all item covariances are positive. Moreover, the alpha value is strictly dependent on the sum of the inter-item covariances (Sijtsma, 2009), as well as on the number of items composing the test (Nunnally & Bernstein, 1994). Alpha does not tell anything about how many latent dimensions are responsible for item covariances. Evidence of unidimensionality comes from exploratory and confirmatory factor analysis (EFA and CFA), with goodness of fit indices that explicitly test the null hypothesis that all covariances among items are explained by a single latent dimension and/or with goodness of fit indices that measure the magnitude of residual covariance once a single latent dimension is partialled out (see also Hattie [1985] and McDonald [1981] on this topic). Thus, only a careful examination of the items' latent structure (e.g., dimensionality) can guide the definition of a pertinent summed score. The risk of not doing so is ending up with summed scores that reflect mixed sources of variance. This eventually jeopardizes the validity of the scale or the extent to which the scale measures the construct it is intended to measure.

In summary, scale dimensionality must be assessed before choosing a method for estimating reliability. If a scale, such as the SCHFI, is multidimensional, especially if the different dimensions composing the scale are measured with a relatively small number of items, other approaches than alpha to assessing reliability are more appropriate. These options are discussed below.

## Linking Reliability Estimates to Factor Analysis Results

In this section we will consider indices of scale reliability that are alternatives to Cronbach alpha. These indices use results from previous confirmatory or exploratory factor analysis in their computations.

### Model-based internal consistency coefficients: Congeneric scales

Fornell and Larcker (1981) introduced two different indices to assess scale reliability. When items are congeneric (i.e., a model with one homogenous factor that fits the data) and measurement errors are not correlated, the reliability of a single item can be estimated from factor analysis results, as shown in equation 1:

$$\rho_y = \frac{\lambda_y^2}{\lambda_y^2 + \text{Var}(\varepsilon_y)} \quad (1)$$

In equation 1,  $\lambda$  is the factor loading of the indicator, and  $\text{Var}(\varepsilon_y)$  is its residual variance, assuming that  $\varphi$ , the variance of the factor is fixed at 1 for identification purposes and that the factor and the measurement error are independent. Under the same assumptions, the

composite reliability index can be computed for a set of  $p$  congeneric indicators of a construct  $\eta$ , as shown in Equation 2 (see Raykov & Marcoulides [2011] for a generalization when measurement errors are correlated):

$$\rho_{\eta} = \frac{\left(\sum_{i=1}^p \lambda_i\right)^2}{\left(\sum_{i=1}^p \lambda_i\right) + \sum_{i=1}^p \text{Var}(\varepsilon_i)} \quad (2)$$

These coefficients are consistent estimators of scale reliability (Raykov, 2012) and can be easily computed using standard output for exploratory or confirmatory factor analysis from commercial or freeware software. This coefficient has a close resemblance to the so-called omega coefficient (McDonald, 1999; Revelle & Zinbarg, 2009). Omega can be used as a measure of reliability only when one latent variable accounts for item correlations (McDonald, 1999).

### Multidimensional scales

Composite reliability coefficient can be extended to consider the more general case of a multidimensional scale, that is, a scale where more than one latent variable explains correlations among observed variables in a data set (Camilli, Wang, & Fesq, 1995) such as the SCHFI. The formula for the global reliability index for multidimensional scales proposed by Raykov and Marcoulides (2011) is shown in equation 3:

$$\rho_U = \frac{\mathbf{1}' \mathbf{A} \Phi \mathbf{A}' \mathbf{1}}{(\mathbf{1}' \mathbf{A} \Phi \mathbf{A}' \mathbf{1} + \mathbf{1}' \Theta \mathbf{1})} \quad (3)$$

Here,  $\mathbf{A}$  is the matrix of factor loadings onto common factors and  $\mathbf{A}'$  its transpose,  $\Phi$  is the covariance matrix of the common factors,  $\Theta$  is the covariance matrix of measurement errors,  $\mathbf{1}$  is a unity vector and  $\mathbf{1}'$  its transpose. This formula assumes that it is rarely possible to distinguish among specific factor and measurement errors (Bollen, 1989; McDonald, 1985; McDonald, 1999), and that these two components sum into the uniqueness term (Bollen, 1989).

Another model-based internal consistency index that can be used when a scale has more than one factor was proposed by Bentler (2009) and shown in equation 4:

$$\hat{\rho}_{xx} = 1 - \frac{\mathbf{1}' \hat{\Psi} \mathbf{1}}{\mathbf{1}' \hat{\Sigma} \mathbf{1}}, \quad (4)$$

Here  $\hat{\Sigma}$  is the fitted covariance matrix obtained from model parameter estimates, and  $\hat{\Psi}$  is the error covariance matrix.

### Maximal reliability (MR)

Items in a scale can be appropriately weighted so that the weighted combination of the items has a higher reliability in comparison to that obtained when the items are summed using a uniform weight of 1. This situation seems to be particularly relevant when the items show remarkable differences in their factor loading (i.e., when they are simply congeneric measures of the construct, with a different “true” component as well as different precision). Optimal weights  $w_j$  are then obtainable for each item from the ratio of the loading and the unique variance, so that  $w_j = b_j/\theta_j$ . The maximal reliability index can be obtained as shown in equation 5 (Raykov, 2012):

$$\rho_{U,m} = \frac{\sum_{j=1}^r (\lambda_j^2 / \theta_j)}{1 + \sum_{j=1}^r (\lambda_j^2 / \theta_j)} \quad (5)$$

### Factor score determinacy coefficient

The similarities between classical test theory and common factor analysis are well-known (e.g., Kline, 2005). In this domain, the factor score determinacy coefficient is used to evaluate the internal consistency of the factor solution (Tabachnick & Fidell, 2013). As noted by Brown (2003), this coefficient represents an important result of factor analysis. In particular, a high degree of determinacy indicates that “the factor score estimates could serve as suitable substitutes for the factor itself” (Brown, 2003, p. 1418).

Factor score determinacy represents the correlation between the estimated and true factor scores. It ranges from 0 to 1 and describes how well the factor is measured, with 1 being the best value (Muthén & Muthén, 1998–2012; see also Grice, 2001). The factor score determinacy coefficient ( $\rho^2$ ) is given by  $\mathbf{f}'\mathbf{R}^{-1}\mathbf{f}$  (McDonald & Mulaik, 1979, equation 6, p. 299) where  $\mathbf{f}$  are common factor loadings and  $\mathbf{R}$  is the correlation matrix among indicators (items) of the factor. The larger the coefficient (e.g., .70, Tabachnick & Fidell, 2013), the more stable the factors, in the sense that the observed variables account for substantial variance in the factor scores, while low values mean the factors are poorly defined by the observed variables.

### Alternative Approaches to Reliability of the SCHFI

As noted above, it is critical that dimensionality testing precede reliability testing. Evidence of dimensionality comes from exploratory and confirmatory factor analysis. As the SCHFI is based on theory, confirmatory factor analysis (CFA) is typically preferred.

In the first psychometric testing of the SCHFI v.6.2 (Riegel et al., 2009) self-care maintenance, self-care management and self-care confidence scales were tested in a single CFA. Results showed poor fit:  $\chi^2 = 356.92$ , comparative fit index (CFI) = .73 and root mean square error of approximation (RMSEA) = .07. In a recent study from Italy (Vellone et al., 2013a), a combination of EFA and CFA was used to test the dimensionality of each scale of the SCHFI v.6.2. This analysis revealed a new factorial structure within each scale that fits



the data with better, more supportive fit indices (CFI = .92, RMSEA = .05 for the self-care maintenance scale; CFI = .95, RMSEA = .07 for the self-care management scale; CFI = .99, RMSEA = .02 for the self-care confidence scale). In the Italian study, reliability was tested with factor score determinacy coefficients that were adequate (Tabachnick & Fidell, 2013), ranging from .74 to .90. In these analyses, when the data were normally distributed we used maximum likelihood estimation, and when not normally distributed we used the Satorra-Bentler correction.

Based on this evidence of multidimensionality and promising evidence of reliability using approaches other than Cronbach alpha, in the current study we explored the factorial structure of the SCHFI v.6.2 in US samples and used those results to identify the best method for users of the SCHFI to assess reliability.

## Methods

### Instrument

This was a secondary analysis of data from 629 adults with heart failure enrolled in three separate studies conducted in the northeastern and northwestern United States. In each study, self-care was measured using the Self-Care of HF Index (SCHFI) v.6.2, a 22-item instrument with three scales that measure the two components of HF self-care: maintenance and management, and confidence - a major influence on self-care (Riegel et al., 2004; Riegel et al., 2009).

The self-care maintenance scale has 10 items that measure symptom monitoring and adherence to behaviors that are advocated in clinical guidelines (Yancy et al., 2013) to maintain patients in a stable state (e.g., monitoring weight, eating a low salt diet, exercising). The 6 items of the self-care management scale measure patients' abilities to recognize symptoms when they occur, treatment implementation in response to symptoms (e.g., reduce fluid intake, take an extra water pill), and treatment evaluation. A self-care management scale score is calculated only in patients who have been recently symptomatic. The self-care confidence scale uses 6 items to evaluate patients' perceived abilities to engage in each phase of the self-care process (e.g., preventing symptom onset, recognizing symptom changes).

Each scale uses a 4-point self-report response format (*never or rarely, sometimes, frequently, always or daily*). The Institutional Review Boards of each site approved the studies before data collection began, and all participants gave informed consent.

### Participants

**Sample 1**—The first sample consisted of 280 adults with chronic HF enrolled from three outpatient settings in Philadelphia, Pennsylvania, and Newark, Delaware. The purpose of this longitudinal descriptive study was to test the relationship between excessive daytime sleepiness and HF self-care. Only baseline data were used in the current analysis. Participants had a diagnosis of HF, confirmed by echocardiography and clinical evidence, were 18 years of age or older, and were medically stable, without an acute coronary event in the past three months. Patients were excluded if they had evidence of dementia, major

depression, or drug or alcohol abuse within the past year. Because this was a study of daytime sleepiness, patients who worked the night shift also were excluded. Data collection took place during home visits by trained research associates, as described in detail elsewhere (Riegel, Moelter, et al., 2011). These patients were predominately older adults, male, and Caucasian. Their demographic and clinical characteristics are shown in Table 1.

**Sample 2**—The second sample consisted of 202 adults with moderate to advanced chronic HF enrolled through a single outpatient HF clinic in Portland, Oregon. The purpose of this prospective cohort study was to describe gender differences in HF symptoms, self-care behaviors, and event-free survival (Lee, Gelow, Denfeld, et al., 2013). All participants had a confirmed diagnosis of HF by echocardiographic and physical examination evidence, were 21 years of age or older and responsible for making their own decisions about their care, and had New York Heart Association (NYHA) class II-IV HF (current symptoms). Patients were excluded if they had a diagnosis of moderate or greater cognitive impairment (i.e., dementia), major and uncorrected visual or hearing deficits, or had received a heart transplant or mechanical circulatory support to manage their HF. Questionnaires including the SCHFI were completed in the clinic or at home and returned by mail; only baseline data were used for this analysis. In general, the sample was mainly middle-aged adults, predominantly Caucasian, and with an equal proportion of males and females (Table 1).

**Sample 3**—The third sample consisted of 146 adults with moderate to advanced chronic HF enrolled through the same single outpatient HF clinic in Portland, Oregon as sample 2. The purpose of this prospective cohort study was to characterize naturally-occurring patterns of change in HF self-care management and their association with quality of life (Lee, Gelow, Mudd, et al., 2013). All participants had a confirmed diagnosis of HF by echocardiographic evidence and history and physical examination, were 21 years of age or older, and had NYHA class II-IV HF. Patients were excluded if they had a diagnosis of dementia, major and uncorrected visual or hearing deficits, or had received a heart transplant or mechanical circulatory support to manage their HF. Questionnaire data were completed in the clinic or at home and submitted by mail; only baseline data were used for this analysis. Many patients participated in both studies in Portland; thus, the total number of unique participants from Studies 2 and 3 for this analysis was 273. The characteristics of these patients are detailed in Table 1, and they were predominantly male, Caucasian, and in middle adulthood.

## Data Analysis

As the inclusion criteria were essentially the same, data from the three samples were combined for analysis. Analyses using the combined sample were then conducted separately on each of the three scales comprising the SCHFI. Consistent with our arguments in the previous section, the following reliability coefficients were considered: a) standardized Cronbach alpha; b) composite reliability (Fornell & Larcker, 1981) or omega (McDonald, 1999); c) maximal reliability (Raykov, 2012); and d) factor score determinacy (McDonald & Mulaik, 1979).

Reliability coefficient  $\alpha$  (which does not use any parameter from the factor model) was obtained from the freeware software FACTOR (Lorenzo-Seva & Ferrando, 2006). All other coefficients were computed from the Mplus output (see below).

As dimensionality testing must precede reliability testing, we began the analysis with CFA and then assessed reliability. Confirmatory factor analyses were conducted using Mplus 7.1 (Muthén & Muthén, 1998–2012). Because the SCHFI item response format presents only four ordered categories, and due to non-normality in the data, the following methods were used for parameter estimation in CFA: a) robust maximum likelihood (ML) estimator, the Satorra-Bentler correction, implemented in Mplus via the estimator MLM); and b) WLS-MV estimator (Muthén & Muthén, 1998–2012).

While ML-based estimators are suitable for continuous (or almost continuous) indicators, they may introduce bias in parameter estimation when used with categorical data. Thus, we expected that internal coherence estimates derived from ML would be underestimated, although this bias may be moderated if data do not present floor or ceiling effects (B. O. Muthén, personal communication, March 24, 2014).

WLS-MV is a weighted least square method highly recommended to applied researchers for the analysis of ordinal or dichotomous variables (Flora & Curran, 2004) and appears to be the ideal choice for data such as those of the SCHFI. However, Raykov and Marcoulides (2011) warned to use internal coherence coefficients such as those introduced above when items have fewer than 5 possible ordinal responses. In fact, WLS-MV method, like other methods for factor-analyzing categorical ordinal variables, considers dichotomous and observed categorical variables as discretization of an underlying continuous variable. In this approach the focus is not on  $X$ , the scale obtained by summing the observed categorical variables, but on  $Z$ , the sum of the underlying normal variables.

As noted by Bentler (2009), internal consistency reliability of  $Z$  can be derived once correlations among underlying normal variables are estimated using polychoric or tetrachoric correlations, and these correlations are used as an input for factor analysis. It must be noted, however, that when internal coherence estimates are based on parameters derived from polychoric correlations, these estimates correspond to the  $Z$  continuous variable that underlies the  $X$  observed discrete responses, not to the latter. In practice, these hypothetical latent variables are never available, so the internal consistency obtained from polychoric coefficients does not reflect the sum of observed items. Finally, in the same way that the polychoric-based factor loadings are generally higher than the product-moment-based factor loadings (due to attenuation to coarse grouping in categorical variables), we expected that these internal coherence estimates would be higher than those corresponding to the observed discrete scores. Therefore, we believed that to have a complete picture of internal coherence of the SCHFI scales, estimates deriving from both robust ML and WLS-MV should be presented.

Based on Hoyle's (1995) recommendations, and according to a multifaceted approach to the assessment of model fit (Tanaka, 1993), we considered a number of goodness of fit indices. Table 2 presents a summary of appropriate fit cut-points (Bentler, 1990; Jöreskog &

Sörbom, 1984; Hu and Bentler, 1999; Yu, 2002). Additionally, traditional chi-square statistics are reported. However, due to the sample size and the sensitivity of the chi-square likelihood ratio test to sample size, chi-square test results were not used in interpreting model fit.

One may question whether the results regarding the dimensionality of SCHFI scales are generalizable across the different datasets in this study. Accordingly a series of multiple-group (MG) confirmatory factor analyses was devised. First, data from samples 2 and 3 (from the two studies conducted in the Pacific Northwest) were merged to a single sample because the number of unique and complete responses from each group was too small for separate analysis. Many of these subjects participated in both studies, and the Self-Care Management scale is completed only by symptomatic patients, leaving insufficient cases in each group. M-Box tests of homogeneity of covariances (Tabachnick & Fidell, 2013) between samples 2 and 3 were non-significant, so covariances in samples 2 and 3 were considered as homogenous. Then, MG-CFAs were performed for the two resulting groups: Sample 1 vs. Samples 2 and 3 combined. In these analyses, we constrained factor loadings to be equal, in order to test for so-called metric invariance (see Meredith, 1993). Because we did not address differences in latent or in observed means, we believed this was an appropriate level of invariance for this analysis.

## Results

### Correlations and Normality of Distribution within Scales

Average correlation (Table 3) for the self-care maintenance items was  $r=.15$ . While there were some higher correlations, in the order of .62 and .49, many other correlations were much lower, illustrating a complex pattern unlikely to be consistent with a single-dimension factor structure. Notably, item 8 (the one reversed item) was poorly correlated with the other items.

Average correlation for the self-care management items (Table 4) was .21. Again, some correlations were fairly adequate, but others were much lower, providing evidence of a complex pattern. Item 14 (extra diuretic) was poorly correlated with all other items.

Average correlation for the self-care confidence items (Table 5) was .47. All correlations in the matrix were higher than .30, reaching high values such as .60, a pattern that is often compatible with a one-factor structure.

Regarding normality of distribution (Tables 3, 4, and 5), the self-care maintenance scale items presented low to moderate non-normality, with the sole exception of item 8, which had high negative skewness and kurtosis. Self-care management scale and self-care confidence scale items had low to moderate non-normality.

### Self-care Maintenance Scale

**Dimensionality**—Riegel and colleagues (2009) posited a single factor underlying the 10 items composing the self-care maintenance scale. Thus, we first specified a one-factor

model. This model proved to be largely unsatisfactory, with poor fit indices shown in Tables 6 and 7, row 1. Two different multidimensional alternative models were then specified.

First a two-factor model derived from the Italian analysis of the SCHFI described above (Vellone et al., 2013a) was tested. Here, items 1, 2, 4, 5, 6 and 9 were tested as indicators of an Autonomous Maintenance factor and items 3, 5, 8 and 10 were tested as indicators of a Provider-Directed Maintenance factor. We also tested a four-factor model derived from Vellone, Riegel and colleagues (2013b), where items 1, 2 were measures of a Symptom Monitoring factor, items 4, 7 were measures of a Physical Activity factor, items 6, 9 were measures of a Sodium Intake factor, and items, 3, 5, 8, 10 were measures of a Medical Treatment Adherence factor. The two-factor model resulted in poor fit, while the four-factor model resulted in an excellent fit (see Tables 6 and 7, rows 2 and 3).

However, item 8 (“Forget to take one of your medicines”) was extremely problematic, showing low and non-significant loading on the posited factor. As previously noted, this item also had very low zero-order correlations (see Table 3) with all other items. Notably, this item is the only negatively worded item in the scale (its scores were reversed before the analysis in order to avoid negative signs in loadings). We believed that item 8 could be considered an outlying variable (Comrey & Lee, 1992), so it was excluded from the next analyses.

The four-factor model was then re-specified without item 8, and this model resulted in an excellent fit (see Tables 6 and 7, row 4). Figure 1 gives a graphic representation of this four-factor model, including estimates from the Mplus output’s completely standardized solutions. Factor loadings were generally medium to high, attesting to a substantial proportion of common variance among the items. All factor correlations but one were positive and significant, attesting to a significant and coherent association among the different facets of self-care maintenance.

**Internal coherence**—The self-care maintenance scale was intended to yield a single score and not four different scores related to the different aspects of the construct. It is possible, however, to achieve this aim despite the substantial multidimensionality seen in the scale. When the alpha coefficient was computed with the nine remaining items in the scale, a poor coefficient of .65 was obtained. However, alpha assumes that the items are satisfying a unidimensional structure. Knowing that there are four dimensions represented in this scale, more appropriate reliability coefficients that takes into account the multidimensionality of the scale are the global reliability index for multidimensional scales (Raykov, 2012) and the model-based internal consistency coefficient (Bentler, 2009). These coefficients were .75 and .76, respectively, when derived from MLM results, and .82 and .83 when derived from WLS-MV results. Although the dimensionality of this scale is complex, Bentler (2006) noted, “every multidimensional coefficient implies a particular composite with maximal unidimensional reliability” (p. 343). Thus, the final reliability estimates derived with appropriate methods “can be interpreted to represent a unidimensional composite” (Bentler, 2006, p. 341).

## Self-care Management Scale

**Dimensionality**—Riegel and colleagues (2009) posited a single factor underlying the 6 items composing the self-care management scale. Accordingly we first specified a one-factor model. This model proved to be largely unsatisfactory, with poor fit indices (see Tables 6 and 7, row 5).

An alternative model was then specified, derived from the Italian study discussed above (Vellone et al., 2013b), in which items 11, 12, 13, and 16 were measures of an Autonomous Management factor, and items 14 and 15 were measures of a Provider-Directed Management factor. This model also proved to be largely unsatisfactory (see Tables 6 and 7, row 6). A close scrutiny of the correlation matrix revealed that item 14 (“Take an extra water pill”) was not adequately correlated with any of the other items in the scale (average correlation = .16). The poor correlation was not entirely unexpected, as patients having trouble recognizing shortness of breath or ankle swelling would be unlikely to take an extra water pill as a remedy (Cocchieri et al., 2014; Prasun, Kocheril, Klass, Dunlap, & Piano, 2005). Further, self-administration of diuretics is not uniformly advocated by providers, and not all patients are willing or able to self-medicate. Thus, the response to this item is dependent on the instructions patients receive from their providers, as well as individual motivation. Following this reasoning, we eliminated item 14 from subsequent analyses and tested the original one-factor model with the remaining 5 items. However, this solution did not improve sufficiently (see Tables 6 and 7, row 7).

In the one-factor solution, items 12 (reduce the salt in your diet), 13 (reduce your fluid intake), and 15 (call your doctor or nurse for guidance) had high loadings on the specified factor (treatment implementation), while items 11 (symptom recognition) and 16 (treatment evaluation) were poorly associated with it. Thus, while items 12, 13 and 15 reflect treatment implementation as in the original theory (Riegel & Dickson, 2008), items 11 and 16 seem to reflect uncertainty in the evaluation of symptoms and the management thereof. Accordingly, we specified a correlated two-factor model. When the model was tested with two factors (Evaluation, items 11 and 16, and Implementation, items 12, 13, and 15), this model now showed an excellent fit to the data (see Tables 6 and 7, row 8).

This final model was statistically equivalent with a second-order model in which the two primary factors of Evaluation and Implementation loaded on a second-order factor of Self-Care Management, which has always been conceptualized as entailing elements of both evaluation and implementation. Figure 2 gives a graphic representation of the hierarchical model including estimates from the Mplus output’s completely standardized solution. In the 2-correlated-factors model shown in Figure 2, the five items are shown as loading on two primary factors, while the second-order factor is replaced by a correlation ( $r = .33$  on MLM and  $.34$  on WLS-MV) between the two factors. As can be seen in Figure 2, factor loadings were generally medium to high, attesting to a substantial proportion of common variance among the items.

**Internal coherence**—As in the case of the self-care maintenance scale, the self-care management scale was intended to yield a single score and not different scores related to the different aspects of the construct. When the alpha coefficient was computed for the five

remaining items in the scale, a poor coefficient of .59 was obtained. As alpha assumes that the items are satisfying a unidimensional structure, and this scale is not unidimensional, we again computed the global reliability index for multidimensional scales (Raykov, 2012) and the model-based internal consistency coefficient (Bentler, 2009). These coefficients were .68 and .66, respectively, when derived from MLM results, and .77 and .76 when derived from WLS-MV results. Although these coefficients are all better than the alpha coefficient, they reached the .70 level of an adequate reliability only when based on WLS-MV estimates (Nunnally & Bernstein, 1994). As noted above, while internal coherence coefficients based on MLM may tend to be downward-biased (due to the low number of ordered categories of item responses), coefficients derived from WLS-MV estimates measure the internal coherence of the latent variables underlying the coarse categorical responses, and may tend to be upward-biased. Therefore, the reliability of the composite obtained by summing the 5 observed scores of self-care management scale must be interpreted with caution.

### Self-care Confidence Scale

**Dimensionality**—Riegel and colleagues (2009) posited a single factor underlying the 6 items composing this scale. Accordingly, we first specified a one-factor model. This model showed an excellent fit (see Tables 6 and 7, row 9). Figure 3 presents the final fitted model: all factor loadings were high, significant and positive.

**Internal coherence**—Table 8 presents reliability estimates for this final model. All the coefficients, regardless of the estimator used, confirm the high reliability of the scale obtained by summing the 6 items related to self-care confidence.

### Generalizability of the Results

The model parameters are substantially consistent between the two samples included in this analysis, with good or very good fit indices, both in the MLM and WLS-MV models (see Table 9). In all these solutions, none of the modification indices associated with constrained parameters resulted in significance at the  $p = .05$  level or below, with the sole exception of item 8 in MLM analysis of the Self-Care Maintenance scale; when the constraint was relaxed, item 8 showed a factor loading of .87 in Sample 1 and .54 in Samples 2+3.

### Discussion

The purpose of this study was to assess the psychometric properties of the SCHFI. To do so, we first tested the dimensionality of each scale using CFA and then chose measures of internal consistency based on scale dimensionality. This analysis revealed that only the self-care confidence scale is unidimensional, while both self-care maintenance and self-care management scales are multidimensional. After eliminating one item from each of these scales, excellent model fit was obtained for each scale. Internal consistency coefficients ranged from .75 to .83 for the self-care maintenance scale and from .66 to .77 for the self-care management scale, using indices that consider the multidimensionality of the scales. Internal consistency of the self-care confidence scale ranged from .84 to .90 when tested using a variety of indices.

The finding that the self-care maintenance and management scales are not unidimensional illustrates that reliability of the SCHFI should not be assessed using coefficient alpha. When the reliability of the self-care maintenance scale was measured with a coefficient that reflects its multidimensional nature, the results were adequate, attesting to the possibility of using a fairly reliable total summed score as a measure of the general construct. Reliability would be higher with subscale scores, but two items are insufficient to produce a robust subscale.

The self-care management scale also is not unidimensional. When reliability was measured with a coefficient that reflects its multidimensional nature, the results were better but not completely adequate, especially when the internal consistency index was computed using MLM-based estimates. Given these results, we recommend that a total summed score be used with caution. Again, reliability would certainly be higher with subscale scores, but new items would be needed.

At this point we advocate that scoring of the self-care management scale take into account whether patients have been instructed to self-administer a diuretic dose for a weight gain. In the most recent update, we noted that item 14 (“Take an extra water pill”) can be skipped and the scoring adjusted using the same general formula (raw score sum minus lowest possible raw score, then divided by the possible range of scores, and finally multiplied by 100). In this update, we go a step further to specify that patients who report that they are “not likely” to take an extra water pill should be asked what they were told to do by their provider. If the provider has not advocated taking an extra water pill, this item should be skipped in the scoring. The scoring formula would need to be adjusted as noted above. Efforts are underway to update the SCHFI in a more thorough fashion in the near future. As we gain insight into differentiating self-care from provider-guided care (Lee, Gelow, Mudd, et al., 2013), and get better at identifying experts in self-care (Riegel, Lee, et al., 2011) who may be well-positioned to self-titrate diuretics in response to symptoms, this item on taking an extra water pill may gain increasing importance.

In practical applications, we often rely on the illusion that factor analysis (exploratory or confirmatory) would remove from the observed variable any sources of variability extraneous to that due to common factors. However, although the factor model accounts for the common variance among the indicators of a construct, it does not remove other sources of variability from the indicators. Therefore, if item scores are simply added together to obtain the total score on a scale, this scale score will reflect both what is due to the common factor and what is due to the residual components (DeShon, 2004).

To limit the impact of such unwanted components, using factor scores to compute total scores may be a possible solution (see Tabachnick & Fidell, 2013). In computing such scores, observed variables would be weighted by a coefficient that is proportional to the factor loading obtained in the factor solution. While factor score coefficients can be easily derived from software output for CFA or for EFA, the computation of weighted scores is more complex and not currently easily implemented. Moreover, factor scores are dependent on the data set in which they were developed, making them difficult to generalize unless they are calibrated on large samples. Thus, at this point, our recommendation is that total



scores continue to be used; we hope to devise and make public a more refined scoring method in the near future.

We acknowledge several limitations to the approach used in this study. We adopted a top-down approach in the definition of the factor models based on prior testing through CFA. In doing so, we considered as a starting point the many studies of the dimensionality of the SCHFI in prior years. Although the present results generally supported the previously-tested models, we still eliminated two items with poor correlations. Doing so introduced a post-hoc modification in our top-down model-testing strategy. One may question whether a fully exploratory approach based on EFA could have provided useful information about the number of dimensions. One approach might have been to split the sample and perform an EFA on one half and CFA on the other, using a cross-validation approach. However, this approach would have resulted in a dramatic reduction in sample size, especially in the analysis of the self-care management scale, for which we had only a sample of 322 subjects. Investigators with more subjects could test such an approach in future studies.

Future investigators also should clarify whether using global scores for constructs that are multidimensional is preferable to using specific scores for the facets or components of these multidimensional constructs. While we demonstrated the reliability of these global scores in our sample, reliability must be demonstrated in other samples to extend the generalizability of the results. Moreover, the current SCHFI intentionally includes only a limited number of items, which constrains its ability to produce robust subscale scores. The trade-off between simplicity and the potential for enlargement with subscale scores is an issue requiring serious consideration. In conclusion, we reaffirm the wisdom of Bentler (2009) that coefficients of internal coherence are only ancillary measures, and the more important objective of the analysis of psychometric characteristics of a scale is to understand its underlying structure.

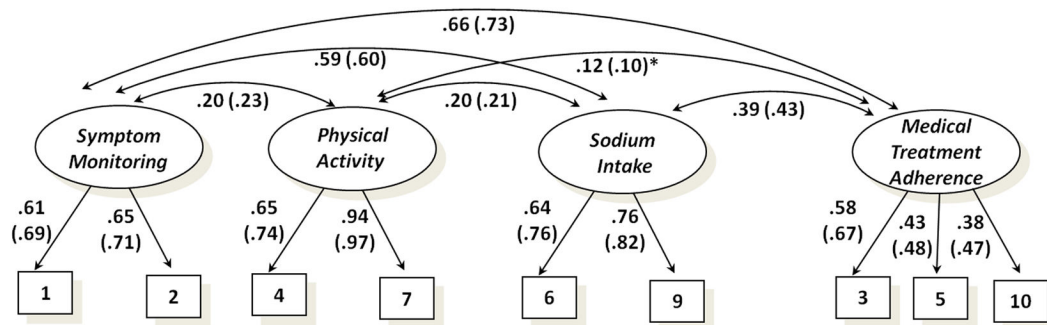
## References

- Banerjee T, Lee KS, Browning SR, Hopenhayn C, Westneat S, Biddle MJ, Moser DK. Limited association between perceived control and health-related quality of life in patients with heart failure. *Journal of Cardiovascular Nursing*. 2013;10.1097/JCN.0b013e31828b2b23
- Bentler PM. Comparative fit indexes in structural models. *Psychological Bulletin*. 1990; 107:238–246. [PubMed: 2320703]
- Bentler, PM. EQS 6 structural equations program manual. Encino, CA: Multivariate Software, Inc; 2006.
- Bentler PM. Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*. 2009; 74:137–143.10.1007/s11336-008-9100-1 [PubMed: 20161430]
- Bollen, KA. Structural equations with latent variables. New York: Wiley; 1989.
- Brown TA. Confirmatory factor analysis of the Penn State Worry Questionnaire: Multiple factors or method effects? *Behaviour Research and Therapy*. 2003; 41:1411–1426. [PubMed: 14583411]
- Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Newbury Park, CA: Sage; 1993. p. 136-162.
- Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. *Nature Reviews Cardiology*. 2011; 8:30–41.10.1038/nrcardio.2010.165
- Camilli G, Wang M, Fesq J. The effects of dimensionality on equating the law school admission test. *Journal of Educational Measurement*. 1995; 32:79–96.

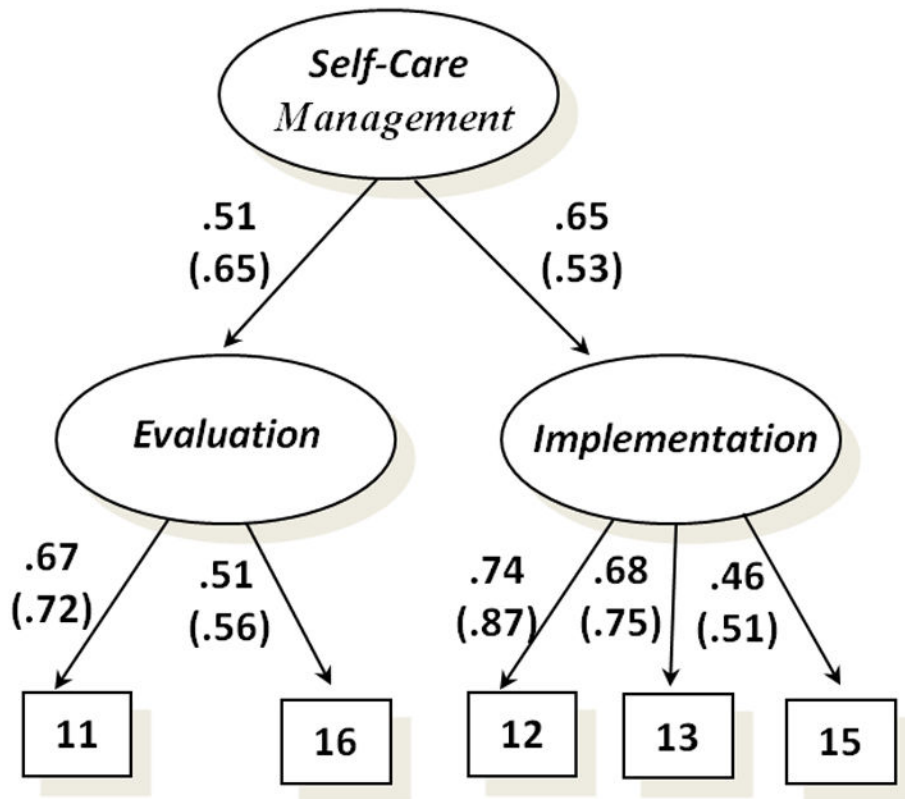
- Cene CW, Haymore LB, Dolan-Soto D, Lin FC, Pignone M, Dewalt DA, Corbie-Smith G. Self-care confidence mediates the relationship between perceived social support and self-care maintenance in adults with heart failure. *Journal of Cardiac Failure*. 2013; 19:202–210.10.1016/j.cardfail.2013.01.009 [PubMed: 23482082]
- Cochieri A, Riegel B, D'Agostino F, Rocco G, Fida R, Alvaro R, Vellone E. Describing self-care in Italian adults with heart failure and identifying determinants of poor self-care. *European Journal of Cardiovascular Nursing*. 2014 Jan 9. epub ahead of print. 10.1177/1474515113518443
- Comrey, AL.; Lee, HB. *A first course in factor analysis*. 2. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.
- DeShon RP. Measures are not invariant across groups with error variance homogeneity. *Psychology Science*. 2004; 46:137–149.
- Flora DB, Curran PJ. An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*. 2004; 9:466–491. [PubMed: 15598100]
- Ford ES, Ajani UA, Croft JB, Critchley JA, Labarthe DR, Kottke TE, Capewell S. Explaining the decrease in U.S. deaths from coronary disease, 1980–2000. *New England Journal of Medicine*. 2007; 356:2388–2398.10.1056/NEJMsa053935 [PubMed: 17554120]
- Fornell C, Larcker D. Evaluating structural equation models with unobservable variable and measurement error. *Journal of Marketing Research*. 1981; 18:39–50.
- Go AS, Mozaffarian D, Roger VL, Benjamin EJ, Berry JD, Blaha MJ. Stroke Statistics Subcommittee. Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation*. 2014; 129:e28–e292.10.1161/01.cir.0000441139.02102.80 [PubMed: 24352519]
- Grice JW. Computing and evaluating factor scores. *Psychological Methods*. 2001; 6:430–450. [PubMed: 11778682]
- Hattie JA. Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*. 1985; 9:139–164.
- Heidenreich PA, Albert NM, Allen LA, Blumke DA, Butler J, Fonarow GC, Pr. Forecasting the impact of heart failure in the United States: A policy statement from the American Heart Association. *Circulation: Heart Failure*. 2013; 6:606–619.10.1161/HHF.0b013e318291329a [PubMed: 23616602]
- Hoyle, RH. *Structural equation modeling: Concepts, issues, and application*. Thousand Oaks, CA: Sage; 1995.
- Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*. 1999; 6(1):1–55.
- Huffman MD, Prabhakaran D. Heart failure: Epidemiology and prevention in India. *The National Medical Journal of India*. 2010; 23(5):283–288. [PubMed: 21250584]
- Jaarsma T, Strömberg A, Ben Gal T, Cameron J, Driscoll A, Duengen HD, Riegel B. Comparison of self-care behaviours of heart failure patients in 15 countries worldwide. *Patient Education and Counseling*. 2013; 92:114–120.10.1016/j.pec.2013.02.017 [PubMed: 23579040]
- Jones CD, Holmes GM, Dewalt DA, Erman B, Broucksou K, Hawk V, Pignone M. Is adherence to weight monitoring or weight-based diuretic self-adjustment associated with fewer heart failure-related emergency department visits or hospitalizations? *Journal of Cardiac Failure*. 2012; 18:576–584.10.1016/j.cardfail.2012.05.004 [PubMed: 22748492]
- Joreskog, KG.; Sorbom, D. *Lisrel VI. Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods*. Mooresville, Indiana: Scientific Software; 1984.
- Kato N, Kinugawa K, Nakayama E, Tsuji T, Kumagai Y, Miura C, Komuro I. Psychometric testing of the Japanese Version of the Self-Care of Heart Failure Index. *Journal of Cardiac Failure*. 2013; 19(8S)
- Kline, RB. *Principles and practice of structural equation modeling*. New York: Guilford Press; 2005.
- Lee CS, Gelow JM, Denfeld QE, Mudd JO, Burgess D, Green JK, Jurgens CY. Physical and psychological symptom profiling and event-free survival in adults with moderate to advanced heart failure. *Journal of Cardiovascular Nursing*. 2013; 29:315–323.10.1097/JCN.0b013e318285968a [PubMed: 23416942]

- Lee CS, Gelow JM, Mudd JO, Green JK, Hiatt SO, Chien C, Riegel B. Profiles of self-care management versus consulting behaviors in adults with heart failure. *European Journal of Cardiovascular Nursing*. 2013 Dec 31. epub ahead of print. 10.1177/1474515113519188
- Lee CS, Moser DK, Lennie TA, Riegel B. Event-free survival in adults with heart failure who engage in self-care management. *Heart and Lung*. 2011; 40:12–20.10.1016/j.hrtlng.2009.12.003 [PubMed: 20561885]
- Lee CS, Moser DK, Lennie TA, Tkacs NC, Margulies KB, Riegel B. Biomarkers of myocardial stress and systemic inflammation in patients who engage in heart failure self-care management. *Journal of Cardiovascular Nursing*. 2011a; 26:321–328.10.1097/JCN.0b013e31820344be [PubMed: 21263344]
- Lee CS, Moser DK, Lennie TA, Tkacs NC, Margulies KB, Riegel B. Biomarkers of myocardial stress and systemic inflammation in patients who engage in heart failure self-care management. *Journal of Cardiovascular Nursing*. 2011b; 26:321–328.10.1097/JCN.0b013e31820344be [PubMed: 21263344]
- Lee CS, Suwanno J, Riegel B. The relationship between self-care and health status domains in Thai patients with heart failure. *European Journal of Cardiovascular Nursing*. 2009; 8:259–266.10.1016/j.ejcnurse.2009.04.002 [PubMed: 19411188]
- Lorenzo-Seva U, Ferrando PJ. FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods*. 2006; 38(1):803–811.
- Maxwell AE. The effect of correlated errors on estimates of reliability coefficients. *Educational and Psychological Measurement*. 1968; 28:803–811.
- McDonagh GK. Heart failure epidemiology: European perspective. *Current Cardiology Reviews*. 2013; 9:123–127. [PubMed: 23597298]
- McDonald RP. The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*. 1981; 34:100–117.
- McDonald, RP. *Test theory. A unified treatment*. Mahwah, NJ: Erlbaum; 1999.
- McDonald RP, Mulaik SA. Determinacy of common factors. *Psychological Bulletin*. 1979; 86:297–306.
- McDonald, RP. *Factor analysis and related methods*. Hillsdale, NJ: Erlbaum; 1985.
- Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika*. 1993; 58:525–543.
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. 7. Los Angeles, CA: Muthén & Muthén; 1998–2012.
- Novick MR, Lewis C. Coefficient alpha and the reliability of composite measurement. *Psychometrika*. 1967; 32:1–13. [PubMed: 5232569]
- Nunnally, JC.; Bernstein, IH. *Psychometric theory*. New York: McGraw Hill, Inc; 1994.
- Ong KL, Cheung BM, Man YB, Lau CP, Lam KS. Prevalence, awareness, treatment, and control of hypertension among United States adults 1999–2004. *Hypertension*. 2007; 49(1):69–75.10.1161/01.HYP.0000252676.46043.18 [PubMed: 17159087]
- Prasun MA, Kocheril AG, Klass PH, Dunlap SH, Piano MR. The effects of a sliding scale diuretic titration protocol in patients with heart failure. *Journal of Cardiovascular Nursing*. 2005; 20:62–70. [PubMed: 15632815]
- Rathman LD, Lee CS, Sarkar S, Small RS. A critical link between heart failure self-care and intrathoracic impedance. *Journal of Cardiovascular Nursing*. 2011; 26:E20–26.10.1097/JCN.0b013e3181ee28c8 [PubMed: 21076309]
- Raykov T. Scale reliability, Cronbach's coefficient alpha, and violations of essential tau-equivalence for fixed congeneric components. *Multivariate Behavioral Research*. 1997; 32:329–354.
- Raykov T. Cronbach's alpha and reliability of composite with interrelated nonhomogenous items. *Applied Psychological Measurement*. 1998; 22:375–385.
- Raykov T. Bias of coefficient alpha for congeneric measures with correlated errors. *Applied Psychological Measurement*. 2001a; 25:69–76.
- Raykov T. Estimation of congeneric scale reliability via covariance structure models with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*. 2001b; 54:315–323. [PubMed: 11817096]

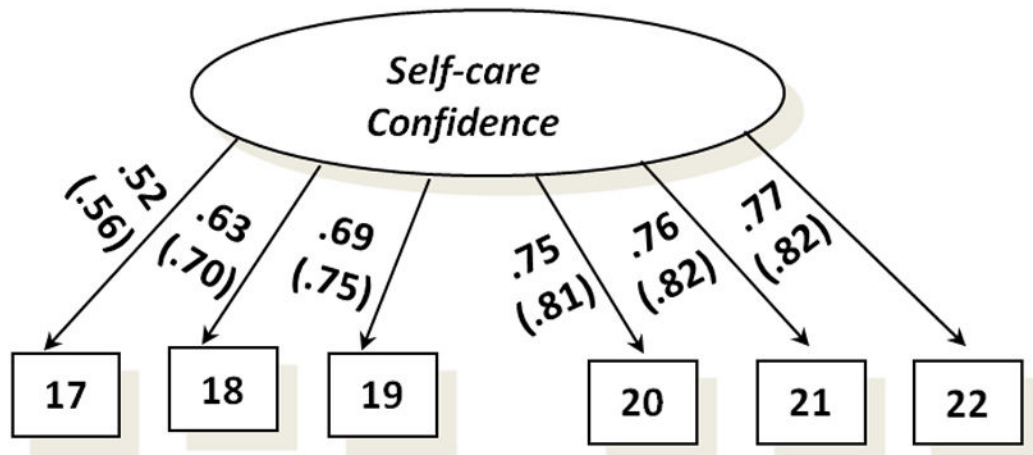
- Raykov, T. Scale construction and development using structural equation modeling. In: Hoyle, RH., editor. *Handbook of Structural Equation Modeling*. New York: The Guilford Press; 2012. p. 472-492.
- Raykov, T.; Marcoulides, GA. *Introduction to psychometric theory*. New York: Routledge; 2011.
- Revelle W, Zinbarg RE. Coefficients alpha, beta, omega, and the GLB: Comments on Sijtsma. *Psychometrika*. 2009; 74:145–154.
- Riegel B, Carlson B, Moser DK, Sebern M, Hicks FD, Roland V. Psychometric testing of the Self-Care of Heart Failure Index. *Journal of Cardiac Failure*. 2004; 10:350–360.10.1016/j.cardfail.2003.12.001 [PubMed: 15309704]
- Riegel B, Dickson VV. A situation-specific theory of heart failure self-care. *Journal of Cardiovascular Nursing*. 2008; 23:190–196.10.1097/01.JCN.0000305091.35259.85 [PubMed: 18437059]
- Riegel B, Lee CS, Albert N, Lennie T, Chung M, Song EK, Moser DK. From novice to expert: Confidence and activity status determine heart failure self-care performance. *Nursing Research*. 2011; 60:132–138.10.1097/NNR.0b013e31820978ec [PubMed: 21317825]
- Riegel B, Lee CS, Dickson VV, Carlson B. An update on the Self-Care of Heart Failure Index. *Journal of Cardiovascular Nursing*. 2009; 24:485–497.10.1097/JCN.0b013e3181b4baa0 [PubMed: 19786884]
- Riegel B, Moelter ST, Ratcliffe SJ, Pressler SJ, De Geest S, Potashnik S, Goldberg LR. Excessive daytime sleepiness is associated with poor medication adherence in adults with heart failure. *Journal of Cardiac Failure*. 2011; 17:340–348.10.1016/j.cardfail.2010.11.002 [PubMed: 21440873]
- Schmitt N. Uses and abuses of coefficient alpha. *Psychological Assessment*. 1996; 8:350–353.
- Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*. 2009; 74:107–120. [PubMed: 20037639]
- Tabachnick, BG.; Fidell, LS. *Using multivariate statistics*. 6. Boston: Allyn & Bacon; 2013.
- Tanaka, JS. Multifaceted conceptions of fit. In: Kennet, AB.; Long, JS., editors. *Structural Equation Models*. Newbury Park, CA: Sage; 1993. p. 10-39.
- Vellone E, Riegel B, Cocchieri A, Barbaranelli C, D'Agostino F, Antonetti G, Alvaro R. Psychometric properties of the Self-Care of Heart Failure Index Version 6.2. *Research in Nursing & Health*. 2013a; 36:500–511.10.1002/nur.21554 [PubMed: 23832431]
- Vellone E, Riegel B, Cocchieri A, Barbaranelli C, D'Agostino F, Glaser D, Alvaro R. Validity and reliability of the Caregiver Contribution to Self-care of Heart Failure Index. *Journal of Cardiovascular Nursing*. 2013b; 28:245–255.10.1097/JCN.0b013e318256385e [PubMed: 22760172]
- Vellone E, Riegel B, D'Agostino F, Fida R, Rocco G, Cocchieri A, Alvaro R. Structural equation model testing the situation-specific theory of heart failure self-care. *Journal of Advanced Nursing*. 2013c; 69:2481–2492.10.1111/jan.12126 [PubMed: 23521633]
- Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Wilkoff BL. 2013 ACCF/AHA Guideline for the Management of Heart Failure: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2013; 128:e240–e327.10.1161/CIR.0b013e31829e8776 [PubMed: 23741058]
- Yu, CY. Unpublished doctoral dissertation. University of California; Los Angeles: 2002. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.
- Yu DS, Lee DT, Thompson DR, Jaarsma T, Woo J, Leung EM. Psychometric properties of the Chinese version of the European Heart Failure Self-care Behaviour Scale. *International Journal of Nursing Studies*. 2011; 48(4):458–467.10.1016/j.ijnurstu.2010.08.011 [PubMed: 20970800]



**Figure 1.** Confirmatory factor analysis model for the Self-Care Maintenance Scale ( $N=549$ ). All factor correlations are significant ( $p<.05$  or lower) except those with asterisks. The standardized coefficients derived from MLM estimators are followed by coefficients (in parentheses) derived from WLS-MV estimators.



**Figure 2.** Confirmatory factor analysis model for Self-Care Management Scale ( $N=322$ ). The standardized coefficients derived from MLM estimators are followed by coefficients (in parentheses) derived from WLS-MV estimators.



**Figure 3.**

Confirmatory factor analysis model for the Self-Care Confidence Scale ( $N=554$ ). The standardized coefficients derived from MLM estimators are followed by coefficients (in parentheses) derived from WLS-MV estimators.

**Table 1**

Demographic and Clinical Characteristics of the Three Samples

Variable	Sample #1 (n=280)		Sample #2 (n=202)		Sample #3 (n=146)	
	M	SD	M	SD	M	SD
Age (years)	62.0	12.5	56.9	13.3	57.4	12.8
Body mass index (BMI)	31.0	8.0	30.7	7.4	32.0	7.1
Systolic blood pressure	116.1	18.4	109.3	16.8	109.3	13.9
Diastolic blood pressure	69.0	11.2	67.2	10.2	68.7	9.5
Months with HF	73.4	71.1	67.8	71.0	72.4	62.4
Ejection fraction	35.4	17.0	28.6	12.4	28.4	11.6
Number of comorbid illnesses	3.1	2.1	2.3	1.5	2.2	1.1
	n	%	n	%	n	%
Male	180	64.3	101	50.0	102	69.9
Race/Ethnicity						
White	175	62.5	173	85.6	126	86.3
Black	96	34.3	2	1.0	5	3.4
Other	9	3.2	27	13.4	11	10.3
Education						
Less than high school	27	9.7	14	6.9	12	8.2
High school	101	36.2	51	25.3	33	22.6
At least some college	78	28.0	137	67.8	101	69.2
Household income insufficient to make ends meet	45	16.1	76	37.8	61	41.8
Perceived overall health						
Excellent/very good	34	12.1	7	3.5	8	5.5
Good	94	33.6	49	24.2	40	27.4
Fair/poor	152	54.3	146	72.3	98	67.1
Atrial fibrillation	94	33.6	78	38.6	63	42.2
Diabetic	108	38.6	48	23.8	37	25.3
Cerebrovascular disease	42	15.0	25	12.4	13	8.9
Renal disease	74	26.4	53	26.2	40	27.9



Variable	Sample #1 (n=280)		Sample #2 (n=202)		Sample #3 (n=146)	
	M	SD	M	SD	M	SD
Heart failure type						
Systolic/mixed	226	80.7	187	92.6	142	97.3
Diastolic	53	18.9	15	7.4	4	2.7
NYHA functional class						
Class I & II	66	23.6	81	40.1	60	41.1
Class III	164	58.6	113	55.9	83	56.8
Class IV	50	17.9	8	4.0	3	2.1
Angiotensin-Converting	162	57.9	107	53.0	81	55.5
Enzyme (ACE) inhibitor						
Angiotensin II Receptor	83	29.6	56	27.7	45	30.8
Blocker (ARB)						
Diuretic	226	80.7	173	85.6	127	87.0
Beta-Blocker	259	92.5	183	90.6	134	91.8

**Table 2**

## Summary of Appropriate Fit Statistics and Acceptable Cut-Points

Statistic	Indicator of Acceptable Fit	Source
Comparative Fit Index (CFI) or Tucker and Lewis Index (TLI)	.90-.95 indicates acceptable fit > .95 indicates good fit	Hu & Bentler, 1999
Root Mean Square Error of Approximation (RMSEA) estimates lack of model fit and compensates for model complexity	.05 indicates a well-fitting model, .05-.08 indicates a moderate fit, .10 indicates poor fit	Browne & Cudeck, 1993
RMSEA is given along with 90% confidence interval limits	.05 to .08 indicates good fit	
RMSEA is given along with the test of close fit examining the probability that the approximation error is low	<i>p</i> -values > .05 indicate good fit	
Standardized Root Mean Square Residual (SRMR); Jöreskog & Sörbom, 1984) is based on residual covariance matrix, assesses discrepancy between observed and predicted covariances	.08 indicates good fit	Hu & Bentler, 1999
The Weighted Root Mean Square Residual (WRMR) is used with WLS-MV estimators	1.0 indicates a good fit	Yu, 2002
Traditional chi-square statistics should be reported as well		

**Table 3**  
 Self-Care Maintenance Scale: Correlations and Descriptors of Distribution and Normality (N = 549)

	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	M	SD	S	K
Item 1	--									3.05	1.22	-0.65	-0.94
Item 2	.40	--								3.02	1.39	-0.69	-0.91
Item 3	.18	.32	--							3.36	0.70	-1.33	0.76
Item 4	.19	.01	.03	--						2.70	0.97	-0.11	-1.10
Item 5	.18	.17	.24	-.06	--					3.43	0.39	-1.23	0.67
Item 6	.26	.20	.16	.14	.13	--				3.20	0.77	-0.87	-0.34
Item 7	.18	.06	.07	.62	-.01	.13	--			2.10	1.17	0.48	-0.99
Item 8	.02	.04	.05	.15	.00	.06	.13	--		3.63	0.17	-2.05	4.89
Item 9	.27	.30	.15	.06	.14	.49	.14	.02	--	2.34	1.90	0.20	-1.45
Item 10	.15	.12	.22	.05	.19	.13	.11	.05	.09	3.30	2.01	-1.25	-0.27

Note. S = skewness; K = kurtosis. All correlations are statistically significant ( $p < .05$ ) except those in italics.

**Table 4**  
Self-Care Management Scale: Correlations and Descriptors of Distribution and Normality (n = 322)

	Item 11	Item 12	Item 13	Item 14	Item 15	M	SD	S	K
Item 11	--					2.67	3.23	-0.75	-0.53
Item 12	.15	--				3.39	0.81	-1.43	0.89
Item 13	.20	.50	--			2.88	1.57	-0.49	-1.17
Item 14	.19	.08	.25	--		2.78	3.36	0.26	0.94
Item 15	.05	.35	.31	.13	--	2.72	1.94	-0.23	-1.47
Item 16	.34	.16	.06	.20	.11	2.45	3.13	-0.49	-0.85

Note. S = skewness; K = kurtosis. All correlations are statistically significant ( $p < .05$ ) except those in italics.

**Table 5**  
 Self-Care Confidence Scale: Correlations and Descriptors of Distribution and Normality (n = 554)

	Item 17	Item 18	Item 19	Item 20	Item 21	M	SD	S	K
Item 17	--					2.32	0.65	0.22	-0.70
Item 18	.30	--				3.22	0.34	-0.58	-0.47
Item 19	.36	.46	--			3.16	0.37	-0.63	-0.13
Item 20	.35	.51	.52	--		3.11	0.42	-0.53	-0.41
Item 21	.45	.44	.53	.53	--	3.08	0.43	-0.55	-0.30
Item 22	.39	.46	.49	.60	.60	2.94	0.49	-0.36	-0.57

Note. S = skewness; K = kurtosis. All correlations are statistically significant ( $p < .05$ ).

**Table 6**  
Fit Indices from CFA Models (Analysis of Covariance Matrix Using MLM Robust Estimator)

Model	$\chi^2$	DF	$p(\chi^2)$	TLI	CFI	SRMR	RMSEA	$p(\text{RMSEA} < .05)$	RMSEA 90% CI
Self-care maintenance ( $n = 549$ )									
1. One-Factor Model	380.00	35	.001	.41	.54	.090	.130	(.001)	[.12, .15]
2. Two-Factor Model	354.00	34	.001	.43	.57	.090	.130	(.001)	[.12, .14]
3. Four-Factor Model	65.34	29	.001	.93	.95	.040	.048	(.570)	[.032, .063]
4. Four-Factor Model without item #8	49.19	21	.001	.93	.96	.035	.049	(.490)	[.032, .068]
Self-care management ( $n = 322$ )									
5. One-Factor Model	55.13	9	.001	.62	.77	.074	.130	(.001)	[.095, .16]
6. Two-Factor Model	57.89	8	.001	.54	.75	.074	.140	(.001)	[.11, .17]
7. One-Factor Model without item #14	39.00	5	.001	.63	.81	.070	.140	(.001)	[.10, .19]
8. Two-Factor Model without item #14	8.32	4	.080	.94	.98	.023	.058	(.340)	[.00, .11]
Self-care confidence ( $n = 554$ )									
9. One-Factor Model	23.35	9	.010	.99	.99	.023	.050	(.370)	[.03, .08]

Note. DF = degree of freedom; TLI = Tucker Lewis Index; CFI = comparative fit index; SRMR = Standardized Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; CI = confidence interval.

**Table 7**  
Fit Indices from CFA Models (Analysis of Polychoric Correlations Using WLS-MV Estimator)

Model	$\chi^2$	DF	$p(\chi^2)$	TLI	CFI	WRMR	RMSEA	$p(\text{RMSEA} < .05)$	RMSEA 90% CI
Self-care maintenance ( $n = 549$ )									
1. One-Factor Model	589.00	35	.001	.55	.65	2.770	.170	(.001)	[.16, .18]
2. Two-Factor Model	555.00	34	.001	.57	.68	2.680	.170	(.001)	[.15, .18]
3. Four-Factor Model	83.58	29	.001	.95	.97	.917	.059	(.160)	[.044, .073]
4. Four Factor Model without item #8	59.13	21	.001	.98	.96	.795	.058	(.220)	[.040, .075]
Self-care management ( $n = 322$ )									
5. One-Factor Model	96.96	9	.001	.67	.80	1.392	.170	(.001)	[.14, .21]
6. Two-Factor Model	97.27	8	.001	.62	.80	1.389	.190	(.001)	[.15, .22]
7. One-Factor Model without item #14	79.00	5	.001	.60	.80	1.410	.210	(.001)	[.17, .26]
8. Two-Factor Model without item #14	7.04	4	.130	.99	.98	.391	.049	(.440)	[.00, .106]
Self-care confidence ( $n = 554$ )									
9. One-Factor Model	36.18	9	.001	.99	.99	.715	.070	(.052)	[.05, .10]

Note. DF = degree of freedom; TLI = Tucker Lewis Index; CFI = comparative fit index; WRMR = Weighted Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; CI = confidence interval.

**Table 8**

## Self-Care Confidence Scale: Reliability Estimates for Single Factor in Confirmatory Factor Analyses

<b>Index</b>	<b>Coefficients Derived from MLM Estimators</b>	<b>Coefficients Derived from WLS-MV Estimators</b>
Standardized Cronbach alpha ( $\alpha$ )	.84	.88
Composite Reliability (CR)	.84	.88
McDonald's Omega ( $\omega_i$ )	.84	.88
Maximal Reliability (MR)	.86	.90
Factor Score Determinacy (FSD)	.86	<i>a</i>

*Note.* The FSD coefficient has been derived from results of the CFA solution using the coefficient described in McDonald & Mulaik (1979).

<sup>a</sup>FSD is not presented in the solution derived from WLS-MV because this index does not make sense for categorical factor indicators.



Table 9

Fit Indices from Multiple-Groups Confirmatory Factor Analysis Models

	$\chi^2$	DF	$p(\chi^2)$	TLI	CFI	SRMR or WRMR	RMSEA	$p(\text{RMSEA} < .05)$
Self-care Maintenance Scale ( $n = 549$ )								
ML	100.62	51	.001	.91	.94	.055	.059	(.17)
WLS	91.59	51	.001	.97	.98	1.03	.054	(.35)
Self-care Management Scale ( $n = 322$ )								
ML	15.24	11	.170	.97	.99	.038	.042	(.55)
WLS	15.96	11	.140	.98	.99	.652	.046	(.51)
Self-care Confidence Scale ( $n = 554$ )								
ML	35.00	20	.020	.98	.99	.055	.052	(.42)
WLS	48.87	21	.001	.99	.99	.971	.07	(.10)

Note. DF = degree of freedom; TLI = Tucker Lewis Index; CFI = comparative fit index; WRMR = Weighted Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation.