



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *Journal of Cardiovascular Nursing*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Lin, C-Y., Pakpour, A H., Broström, A., Fridlund, B., Årestedt, K. et al. (2018)
Psychometric Properties of the 9-item European Heart Failure Self-care Behavior Scale
Using Confirmatory Factor Analysis and Rasch Analysis Among Iranian Patients.
Journal of Cardiovascular Nursing, 33(3): 281-288
<https://doi.org/10.1097/JCN.0000000000000444>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-39222>

**Psychometric Properties of the 9-item European Heart Failure Self-Care Behavior Scale
Using Confirmatory Factor analysis and Rasch Analysis Among Iranian Patients**

Chung-Ying Lin, PhD

Assistant Professor, Department of Rehabilitation Sciences, Faculty of Health and Social Sciences, Hong Kong Polytechnic University, Hung Hom, Hong Kong

*Amir H. Pakpour, PhD

Associate Professor, Social Determinants of Health Research Center, Qazvin University of Medical Sciences, Shahid Bahonar Blvd, Qazvin 3419759811, Iran; Department of Nursing, School of Health and Welfare, Jönköping University, Jönköping, Sweden

Anders Broström, PhD

Professor, Department of Nursing, School of Health and Welfare, Jönköping University, Jönköping, Sweden

Bengt Fridlund, PhD

Professor, School of Health and Welfare, Jönköping University, Jönköping, Sweden

Kristofer Årestedt, PhD

Professor, Faculty of Health and Life Sciences, Linnaeus University, Kalmar; Department of Medical and Health Sciences, Linköping University, Sweden

Anna Strömberg, PhD

Professor, Department of Medical and Health Sciences, Linköping University, Sweden

Tiny Jaarsma, PhD

Professor, Department of Social and Welfare Studies, Linköping University, Sweden

Jan Mårtensson, PhD³

Professor, Department of Nursing, School of Health and Welfare, Jönköping University,

Jönköping, Sweden

*Corresponding author:

Dr. Amir H. Pakpour

Present/permanent address: Social Determinants of Health Research Center, Qazvin

University of Medical Sciences, Shahid Bahonar Blvd, Qazvin 3419759811, Iran.

E-mail: pakpour_amir@yahoo.com (A.H. Pakpour)

Phone: +98-28-33239259

Fax: +98-28-33239259

The authors have no funding or conflicts of interest to disclose

Number of words in text: 2971

Number of tables and figures: Table=3; Figure=0

Abstract

Background: The nine-item European Heart Failure Self-Care Behavior scale (EHFScB-9) is a self-reported questionnaire commonly used to capture the self-care behavior of people with heart failure (HF).

Objective: To investigate the EHFScB-9's factorial structure and categorical functioning of the response scale and differential item functioning (DIF) across sub-populations in Iran.

Methods: Patients with HF (n=380; 60.5% male; mean [SD] age=61.7 [9.1]) participated in this study. The median (interquartile range) of the duration of their HF was 6.0 (2.4, 8.8) months. The majority of the participants were in New York Heart Association classification II (NYHA II, 61.8%); few of them had left ventricular ejection fraction assessment (11.3%). All participants completed the EHFScB-9. Confirmatory factor analysis was used to test the factorial structure of the EHFScB-9; Rasch analysis was used to analyze categorical functioning and DIF items across two characteristics (gender and NYHA).

Results: The two-factor structure (*adherence to regimen* and *consulting behavior*) of the EHFScB-9 was confirmed and the unidimensionality of each factor was found. Categorical functioning was supported for all items. No items displayed substantial DIF across gender (DIF contrast=-0.25-0.31). Except for Item 3—(*Contact doctor or nurse if legs/feet are swollen*; DIF contrast=-0.69)—no items displayed substantial DIF across NYHA classes (DIF contrast=-0.40 to 0.47).

Conclusions: Despite the DIF displayed in one item across the NYHA classes, the EHFScB-9 demonstrated sound psychometric properties in patients with HF.

Keywords: confirmatory factor analysis; heart failure; Rasch; self-care behavior scale

Introduction

Heart failure (HF)—a complex syndrome with high rates of mortality and morbidity^{1,2}—remains a major health problem globally.¹⁻⁴ Recent data from the American Heart Association show a significantly increased number: over 6.5 million people in the US and over 15 million in Europe suffer from HF.¹ The prevalence of HF in developed countries is ~1 to 2%, and much higher (>10%) if only focusing on older populations over 70 years of age (> 10%).^{5,6} The prevalence of HF in developing country is high as well; for example, the prevalence in Iran in the near future is projected to be ~3.5%.⁷

To decrease the social burden and caregiver burden from HF, the European Society of Cardiology guidelines for acute and chronic HF propose that HF patients should receive a holistic and multidisciplinary approach, including psychoeducation and promotion of self-care behavior.^{5,6} Self-care behavior for HF is defined as “a naturalistic decision-making process involving the choice of behaviors that maintain physiologic stability (maintenance) and the response to symptoms when they occur (management).”^{8,9} Patients with HF can improve their health if they engage in such effective behaviors.^{10,11} Therefore, given that the importance of self-care behavior for patients with HF is highlighted, an effective and feasible instrument measuring such behaviors is warranted.

The most commonly used instruments developed for self-care behavior for patients with HF are the Self-Care of Heart Failure Index (SCHFI)¹² and the European Heart Failure Self-Care Behavior scale (EHFScB)¹³. However, SCHFI and EHFScB have different conceptualization of HF self-care: SCHFI focuses on three dimensions of self-care (maintenance, management, and confidence); EHFScB was developed first using the three dimensions of complying with regimen, asking for help, and adapting activities. We investigated the use of EHFScB because it has the following strengths: (1) short and easy to understand, (2) developed through a strong procedure by many HF experts in HF,¹³ (3) useful

to evaluate the effectiveness of psychoeducation on self-care behavior because the EHFSdB covers the patient's ability to participate in effective self-care, and (4) a theory-driven approach to design a two-factor structure (i.e., adherence to regimen and consulting behavior) instead of creating the factorial structure based on data properties. The EHFSdB originally contained 12 items but has been shortened to 9 items (EHFSdB-9) because of improvements in psychometric properties and reduced respondent burden.¹⁴

Although the EHFSdB-9 is a promising instrument to assess HF self-care behavior, its psychometric properties are underdeveloped. Several studies have reported on the 12-item version, but only five studies have been identified that tested the psychometric properties of the EHFSdB-9.¹⁵ In addition, studies found different factorial structures for the EHFSdB-9 using confirmatory factor analysis (CFA).¹⁵ There are four proposed factorial structures for the EHFSdB-9: an overall underlying domain;¹⁶ two domains (*adherence to regimen* and *consulting behavior*);¹⁴ three domains (*adhering to recommendations, fluid and sodium management, and physical activity and recognition of deteriorating symptoms*);¹⁷ and three domains (*consulting behavior, autonomy-based adherence, and provider-based adherence*).¹⁸ Inconsistent findings regarding the factorial structure of the scale give no definite answers about the underlying structure of the scale; therefore, additional analyses comparing different factorial structures are needed. Østergaard et al.¹⁹ identified the problems of different factorial structures, and used several CFAs to compare the four proposed factorial structures. They found that the two-factor structure (*adherence to regimen* and *consulting behavior*) outperformed other factorial structures; however, further corroborations on their findings are needed. Specifically, they did not use diagonally weighted least squares (DWLS), an estimator for Likert-type scale,^{20,21} to take care of their models; they treated the indicator variables (i.e., items) as continuous, and the estimated parameters (e.g., loading and standard error) may be biased.

Another underdeveloped issue for the psychometric properties of the EHFS_cB-9 regards the theory used for psychometric testing. Studies have mainly adopted the classical test theory to evaluate the psychometric properties of the EHFS_cB-9.^{14,16-19} Although the internal consistency of the EHFS_cB-9 was satisfactory across several European countries using Cronbach's α ,^{14,19} to the best of our knowledge, no one has used Rasch models. Rasch analysis uses a mathematical formula to estimate the probability of an individual responding with a certain answer (e.g., scores 1 to 5 in the Likert-type scale) on an item. Using the formula, both item and person properties can be estimated; that is, we can estimate the item difficulty (whether the HF self-care is easy to be complied with) and the person's ability (whether the person has the capability to do the HF self-care). We acknowledge that using classical test theory has the benefit of being easily understood;²² however, without the results analyzed using Rasch models, some psychometric characteristics of the EHFS_cB-9 cannot be identified. For example, it is unclear how each item is embedded in the EHFS_cB-9 (i.e., how much the concept of the item is out of the EHFS_cB-9, or how much of it is redundant to the information provided by other items); whether the categorical functioning of the response scale is in order; and whether different sub-populations of patients with HF interpret the EHFS_cB-9 items similarly.

Testing whether different subgroups interpret the EHFS_cB-9 items differently is a critical issue because the score differences of a measure can be the true group differences or the various understandings toward the same measure contents.²³ If we want to ensure that the score difference in the EHFS_cB-9 is from the true group difference, we should exclude the possibility of various understandings. Differential item functioning (DIF) in the Rasch analysis is a recommended method to tackle this issue.^{24,25}

This study aimed to: (1) compare competing CFA models previously described in the literature, and (2) evaluate the instrument using the Rasch model.

Methods

Instrument

The EHFS_cB-9 contains nine items measuring how well a patient with HF manages self-care. The item contents regarding how to self-care were designed specifically on HF symptoms, and all the items are rated on a five-point Likert scale ranging from 1 (*completely agree*) to 5 (*completely disagree*). A lower score represents better self-care of a patient with HF.¹⁴

Design

The ethics committee of Qazvin University of Medical Sciences approved the study. All participants gave written informed consent for inclusion in the study. The psychometric evaluation study was conducted in two steps: (1) translation into Persian and pilot-testing the EHFS_cB-9^{26,27} and (2) psychometric evaluation. Detailed information of the first step is described in Appendix A.

Participants and procedures for psychometric evaluation

Patients with HF were recruited from seven university hospitals of Tehran, Qazvin, Tabriz, Mashhad, Zanjan, Sari, and Ilam. All hospitals were referral university hospitals which have the highest admission rates of heart diseases. Inclusion criteria: patients at least 18 years of age, had New York Heart Association classification (NYHA) II–IV, and able to read and write Persian. Exclusion criteria: have cognitive impairment (Mini-Mental State Examination score <25), received heart transplant, or inability to understand Persian. Clinical variables were collected from patients' medical records.

Statistical analysis

We tested the internal consistency of the EHFS_cB-9 using polychoric correlation (ρ):²⁸ the ordinal alpha ($\alpha > 0.7$ suggests acceptable) and corrected item-total polychoric correlations ($\rho > 0.4$ suggests acceptable).

We used CFA to determine the factorial structure of the EHFS_cB-9. Specifically, we

tested four different competing factorial models (Table 2) suggested in previous research: one-factor,¹⁶ two-factor (*adherence to regimen* and *consulting behavior*),¹⁴ three-factor (*adherence to recommendations; fluid and sodium management; physical activity and recognition of deteriorating symptoms*) proposed by Lambrinou et al.,¹⁷ and another three-factor model (*consulting behavior; autonomy-based adherence; provider-based adherence*) proposed by Vellone et al.¹⁸ We adopted the estimator of DWLS, which is based on polychoric correlations and recommended for Likert-type scales.^{20,21} To identify supported models, we used comparative fit index (CFI) and Tucker-Lewis index (TLI) > 0.95, root mean square of error approximation (RMSEA) < 0.05, and weighted root mean square residual (WRMR) < 0.90.²⁹⁻³¹ Moreover, we detected whether any model contained offending estimates (e.g., correlations coefficient > 1) and considered the model with offending estimates as an inappropriate model (i.e., Heywood cases).

After the factorial structure was confirmed by CFA, we used the Rasch partial-credit model to examine the unidimensionality of each EHFScB-9 factor, item difficulty, item fit statistics, separation reliability, categorical functioning, and DIF for gender and NYHA. We used principal component analysis (PCA) on the standardized residuals retrieved from Rasch models to test the unidimensionality and the first component's eigenvalue of < 2 to indicate unidimensionality.²² Item difficulty was presented using the log-odd unit, namely, *logit*, where a higher logit represents a more difficult item. Item fit statistics included information-weighted fit statistic (infit) mean square (MnSq) and outlier-sensitive fit statistic (outfit) MnSq (the ratio of the observed response to the predicted response with an ideal value at 1.0); however, some variation from expectation (i.e., 40%) is allowed because Rasch analysis is a probability model. Infit and outfit MnSq between 0.6 and 1.4 of each item suggests an acceptable fit.³² Separation reliability included person and item separation reliabilities; a value > 0.7 is acceptable.³³ Categorical functioning was used to understand whether

successive response categories for each item were located in their expected order; for example, the difficulty of the response “Completely agree” should be lower than that of the response “Agree” in all items. Both average measure (the estimated average ability on a particular category) and step measure (the thresholds between categories) should monotonically increase with categories.³⁴ DIF was examined for each item across gender (male vs. female) and across NYHA classification (NYHA II vs. NYHA III and IV), and a DIF contrast (i.e., the logit of Group 1 minus logit of Group 2) > 0.5 logits suggests substantial DIF.³⁵

Descriptive statistics were analyzed using IBM SPSS 23.0 (IBM Corp., Armonk, NY.), CFA and ordinal alpha using the R software (lavaan package for CFA³⁶ and psych for ordinal alpha³⁷), and Rasch models using WINSTEPS[®] software (Winsteps, Chicago, IL).

Results

The response rate was 87% (n=380). Table 1 presents the demographics of participants. In brief, the participants were relatively young (mean age=61.7±9.1 years) with limited education (mean years=6.4±3.7 years), and less than half of participants were receiving guideline-based care (e.g., ACE inhibitors [42.6%] and Beta blockers [40.5%]). The internal consistency was good for the entire EHFScB-9 (ordinal α =0.87) and the two factors of the originally proposed two-factor EHFScB-9 (ordinal α =0.81 [*Adherence to regimen*] and 0.89 [*Consulting behavior*]). All the corrected item-total polyserial correlations were acceptable with the range between 0.46 and 0.67.

(Insert Table 1 here)

The originally proposed two-factor model demonstrated better model fit than the one-factor model and the three-factor model suggested by Lambrinou et al.¹⁷ The three-factor model described by Vellone et al.¹⁸ showed all fit indices close to the two-factor model; however, it contained a Heywood case: the correlation between Factors 2 and 3 was larger

than 1 (Table 2). Based on these results, the following Rasch analyses and concurrent validity were analyzed using the two-factor structure: Items 1, 5, 7, 8, and 9 were embedded in Factor 1 (*Adherence to regimen*); Items 2, 3, 4, and 6 were in Factor 2 (*Consulting behavior*).

(Insert Table 2 here)

Then, we conducted two Rasch models, one for each EHFScB-9 factor. The unidimensionality of each factor was supported by the first PCA of the residuals: the first component's eigenvalue was 1.6 for *Adherence to regimen* and 1.5 for *Consulting behavior*. All items fit well in their underlying construct. Although the range of item difficulties was slightly narrow, the difficulty ranges were widely distributed according to the item thresholds (ranged between -2.16 and 2.42 for *Adherence to regimen*; between -4.33 and 4.04 for *Consulting behavior*). Person separation reliability was satisfactory; item separation reliability was excellent (Table 3). Moreover, most of the item residuals showed low local independency (See Appendix B for details). The categorical functioning of each item was supported by average and step measures: both monotonically increased with categories (See Appendix C for details).

No items displayed substantial DIF across gender. Except for Item 3 (*Contact doctor or nurse if legs/feet are swollen*), no items displayed substantial DIF across NYHA. As for Item 3, the DIF contrast (-0.69) suggested that the item description was easier for the participants with NYHA II than those with NYHA III or IV (Table 3). That is, participants with NYHA II tended to score Item 3 lower than those with NYHA III or IV.

(Insert Table 3 here)

Discussion

Generally speaking, our results corroborate the findings of Østergaard et al.¹⁹ and Jaarsma et al.¹⁴: the two-factor structure model fits best for the EHFScB-9. Our excellent person and item separation reliabilities for both factors further corroborate the results of the two-factor

structure. The internal consistency found in our Iranian sample ($\alpha = 0.84$) is satisfactory and consistent with those in the samples of European countries ($\alpha=0.68-0.87$).^{14,19} Our results additionally demonstrated the promising psychometric properties of the EHFScB-9. First, Rasch analyses supported the unidimensionality of the EHFScB-9 for each factor based on the two-factor structure. Second, each item was well-embedded in its underlying construct. Third, the categorical functioning of each item was in order followed by the difficulty, which suggests that the respondents could well distinguish the levels between scores 1 and 2, 2 and 3, and so on. Fourth, no items displayed substantial DIF across gender, and only Item 3 (*Contact doctor or nurse if legs/feet are swollen*) displayed substantial DIF across the classes of NYHA (II vs. III and IV).

Different factorial structures have been proposed for the EHFScB-9.^{14,16-18} Østergaard et al.¹⁹ evaluated all the existing factorial structures using CFA and found that all the factorial structures did not have all fit indices being satisfactory, though the most favorable factorial structure was the two-factor structure. Somewhat unsatisfactory fit indices were found in other studies across all of the proposed factorial structures;^{14,16-18} however, we considered that the unsatisfactory fit indices could be attributable to the estimators used in the studies: except for Lee et al.,¹⁶ other studies^{14,17-19} did not apply the DWLS estimator to account for the ordinal data and ordinal indicator variables in the EHFScB-9 (i.e., the Likert-type scale). Although Lee et al.¹⁶ used the DWLS estimator, they only tested the one-factor structure with somewhat promising fit indices (RMSEA=0.12, WRMR=1.00, CFI=0.94, TLI=0.92), which were in accord with the fit indices of our one-factor structure (RMSEA=0.13, WRMR=1.69, CFI=0.97, TLI=0.96). Consequently, we adopted the DWLS estimator to retest all the proposed factorial structures to corroborate the findings of Østergaard et al.,¹⁹ and we strongly recommend using the two-factor structure of the EHFScB-9 in future studies. However, cross-culture validation is warranted for further corroboration, i.e., to investigate

whether the two subscales are invariant measures of self-care behavior across different language versions.

In addition to using an appropriate estimator to reconfirm the factorial structure for EHFScB-9, another important finding in our study was the DIF items. Gender and severity levels are commonly used factors to test DIF. Taking the EHFScB-9 item *Contact doctor or nurse if gaining weight (more than 2 kg)* as an example, the difference in the item score could be because different genders (or different severity levels) have different self-care behaviors on this item. However, there is another possibility: men and women (or different severity levels) may have different perceptions of “*gaining weight.*” If men (or patients with less severe HF) think that gaining 2.4 kg is not more than 2 kg because 0.4 kg may be due to error, and indicates no need to contact doctor or nurse; females (or more severe patients) think that gaining 2.01 kg indicates the need to contact doctor or nurse, comparing the item scores between genders (or different severity levels) would be inappropriate. Therefore, before using the EHFSc-B-9 score for comparisons, we need to investigate whether genders (or different severity levels) interpret the items similarly. Because most of the items did not display substantial DIF, we recommend using the EHFScB-9 items to compare self-care behavior between males and females, or between those with NYHA II and those with NYHA III and IV.

However, when a researcher wants to compare self-care behavior across different NYHA classes, attention should be paid to Item 3 (*Contact doctor or nurse if legs/feet are swollen*). Our Rasch results indicate that participants with NYHA II may more easily detect their legs/feet to become swollen than those with NYHA III or IV. Because patients with NYHA III and IV more frequently have severe symptoms than those with NYHA II,⁵ we postulated that patients with NYHA III and IV are more used to symptoms and therefore may have more difficulties detecting changes in their swollen legs/feet. In contrast, patients with NYHA II—

who more seldom perceive these symptoms—may more easily observe such changes. Two limitations in our DIF results, however, should be noted. First, our analyses did not allow us to test whether DIF was real or artificial. It is important to identify DIF as real or artificial because it may violate the conclusion of invariant properties of an instrument.³⁸ Second, non-uniform DIF was not possible to evaluate. Future studies are needed to evaluate the importance of this DIF and whether it will affect the interpretation of scores between groups of different gender and NYHA class.

There are limitations in the study. First, we did not collect the retest data for the EHFScB-9; therefore, we cannot evaluate the reproducibility (test-retest reliability). The responsiveness and minimal clinically important differences are also left unknown because we did not collect such data. Second, we did not recruit patients with NYHA I, thus the generalizability of our CFA results may be limited.

In conclusion, despite one item displaced DIF across the NYHA classes, the EHFScB-9 demonstrated sound psychometric properties. Two-factor structure (*adherence to regimen* and *consulting behavior*) of the EHFScB-9 was confirmed and supported by both CFA and Rasch. Because the psychometric properties of the Persian EHFScB-9 were supported, our results could be generalizable to other Persian-speaking countries (Iran, Afghanistan, and Tajikistan, with ~120 million Persian-speaking people). A substantial amount of people may thus get benefits.

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. *Circulation*. 2017;135(10):e146-e603.
2. Grady K. Self-care and quality of life outcomes in heart failure patients. *J Cardiovasc Nurs*. 2008;23 (3):285-292.
3. Clark AM, Savard LA, Spaling MA, Heath S, Duncan AS, Spiers JA. Understanding help-seeking decisions in people with heart failure: a qualitative systematic review. *Int J Nurs Stud*. 2012;49:1582-1597.
4. Hwang B, Luttik ML, Dracup K, Jaarsma T. Family caregiving for patients with heart failure: Types of care provided and gender differences. *J Card Fail*. 2010;16:398-403.
5. McMurray JJ, Adamopoulos S, Anker SD, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2012;33:1787-1847.
6. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37:2129-2200.
7. Bahrami M, Etemadifar S, Shahriari M, Farsani AK. Caregiver burden among Iranian heart failure family caregivers: a descriptive, exploratory, qualitative study. *Iran J Nurs Midwifery Res*. 2014;19(1):56-63.
8. Riegel B, Jaarsma T, Stömberg A. A middle-range theory of self-care of chronic illness. *ANS Adv Nurs Sci*. 2012;35(3):194-204.
9. Riegel B, Lee CS, Dickson VV, Carlson B. An update on the Self-Care of Heart

- Failure Index. *J Cardiovasc Nurs.* 2009;24:485-497.
10. Buck HG, Lee CS, Moser DK, et al. The relationship between self-care and health related quality of life in older adults with moderate to advanced heart failure. *J Cardiovasc Nurs.* 2012;27(1):8-15.
 11. Lee CS, Moser DK, Lennie TA, Tkacs NC, Margulies KB, Riegel B. Biomarkers of myocardial stress and systemic inflammation are lower in patients who engage in heart failure self-care management. *J Cardiovasc Nurs.* 2011;6(4):321-328.
 12. Vellone E, Riegel B, Cocchieri A, et al. Psychometric testing of the self-care of heart failure index version 6.2. *Res Nurs Health.* 2013;36(5):500-511.
 13. Jaarsma T, Strömberg A, Mårtensson J, Dracup K. Development and testing of the European Heart Failure Self-Care Behaviour Scale. *Eur J Heart Fail.* 2003;5(3):363-370.
 14. Jaarsma T, Arestedt KF, Mårtensson J, Dracup K, Strömberg A. The European Heart Failure Self-care Behaviour scale revised into a nine-item scale (EHFScB-9): a reliable and valid international instrument. *Eur J Heart Fail.* 2009;11:99-105.
 15. Sedlar N, Socan G, Farkas J, et al. (2017). Measuring self-care in patients with heart failure: a review of the psychometric properties of the European Heart Failure Self-Care Behaviour Scale (EHFScBS). *Patient Educ Couns.* 2017; Advance online publication. doi: 10.1016/j.pec.2017.02.005.
 16. Lee CS, Lyons KS, Gelow JM, et al. Validity and reliability of the European Heart Failure Self-care Behavior Scale among adults from the United States with symptomatic heart failure. *Eur J Cardiovasc Nurs.* 2013;12:214-218.
 17. Lambrinou E, Kalogirou F, Lamnisis D, et al. The Greek version of the 9-item European Heart Failure Self-care Behaviour Scale: a multidimensional or a uni-dimensional scale? *Heart Lung.* 2014;43:494-499.

18. Vellone E, Jaarsma T, Strömberg A, et al. The European Heart Failure Self-care Behaviour Scale: new insights into factorial structure, reliability, precision and scoring procedure. *Patient Educ Couns*. 2014;94:97-102.
19. Østergaard B, Mahrer-Imhof R, Lauridsen J, Wagner L. Validity and reliability of the Danish version of the 9-item European Heart Failure Self-care Behavior Scale. *Scand J Caring Sci*. 2016; Advance online publication. doi: 10.1111/scs.12342.
20. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods*. 2016;48:936-949.
21. Lin C-Y, Burri A, Fridlund B, Pakpour AH. Female sexual function mediates the effects of medication adherence on quality of life in people with epilepsy. *Epilepsy Behav*. 2017;67:60-65.
22. Chang C-C, Su J-A, Tsai C-S, Yen C-F, Liu J-H, Lin C-Y. Rasch analysis suggested three unidimensional domains for Affiliate Stigma Scale: additional psychometric evaluation. *J Clin Epidemiol*. 2015;68(6):674-683.
23. Gregorich SE. Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Med Care*. 2006;44(11 Suppl. 3):S78-S94.
24. Brodersen J, Meads D, Kreiner S, Thorsen H, Doward L, McKenna S. Methodological aspects of differential item functioning in the Rasch model. *J Med Econ*. 2007;10:309-324.
25. Tennant A, Penta M, Tesio L, et al. Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project. *Med Care*. 2004;42(1):I37-I48.

26. Beaton, DE, Bombardier, C, Guillemin, F, Ferraz, MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25(24):3186-3191.
27. Harris-Kojetin LD, Fowler FJ Jr, Brown JA, Schnaier JA, Sweeny SF. The use of cognitive testing to develop and evaluate CAHPS 1.0 core survey items. Consumer Assessment of Health Plans Study. *Med Care*. 1999;37(Suppl. 3):MS10-MS21.
28. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: a conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*. 2012;17(3):1-13.
29. Cheng C-P, Luh W-M, Yang A-L, Su C-T, Lin C-Y. Agreement of children and parents scores on Chinese version of Pediatric Quality of Life Inventory 4.0: further psychometric development. *Appl Res Qual Life*. 2016;11:891-906.
30. Muthén LK, Muthén BO. *Mplus user's guide* (2nd ed). Los Angeles: Authors; 2001.
31. Tsai M-C, Strong C, Lin C-Y. Effects of pubertal timing on deviant behaviors in Taiwan: a longitudinal analysis of 7th- to 12th-grade adolescents. *J Adolesc*. 2015;42:87-97.
32. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measurement Transactions*. 1994;8:370.
33. Duncan PW, Bode RK, Min Lai S, Perera S. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Arch Phys Med Rehabil*. 2003;84(7):950-963.
34. Jafari P, Bagheri Z, Ayatollahi SMT, Soltani Z. Using Rasch rating scale model to reassess the psychometric properties of the Persian version of the PedsQLTM 4.0 Generic Core Scales in school children. *Health Qual Life Outcomes*. 2012;10:27.
35. Shih C-L, Wang W-C. Differential item functioning detection using the multiple indicators, multiple causes method with a pure short anchor. *Applied Psychological*

Measurement. 2009;33:184-199.

36. Rosseel Y, Oberski D, Byrnes J, et al. Package 'lavaan'. 2017 Available at <https://cran.r-project.org/web/packages/lavaan/lavaan.pdf>. [accessed 31 May 2017].
37. Revelle W. Package 'psych'. 2017 Available at <https://cran.r-project.org/web/packages/psych/psych.pdf>. [accessed 31 May 2017]
38. Andrich D, Hagquist C. Real and artificial differential item functioning. *J Educ Behav Stat*. 2012;37(3):387-416.

Appendix A. Detailed information on translation procedure and pilot testing

The EHFScB-9 was translated using a standard forward–backward translation methodology. Based on recommendations proposed by Beaton et al.,²⁶ five stages were employed to translate the EHFScB-9 into Persian. First, two bilingual translators who were native Persian speakers translated the original English version into Persian. Second, the translated versions were then synthesized in an interim Persian version. Third, two bilingual translators who were native English speakers translated the interim Persian version back into English. The translators performed the procedure independently and were blinded to the original English version. Fourth, all translated versions were compared in a session with interdisciplinary experts: cardiovascular nurses, cardiologists, health psychologists, and the translators. All translated versions were reviewed by the translators as well as a coordinator; any discrepancies were discussed and resolved. Fifth, the prefinal Persian version was pilot tested on 24 patients with HF using cognitive interviewing to explore both meaning and responses of the items, and essential amendments were carried out at this stage.²⁷ Patients were systematically debriefed to explore whether they could repeat the questions in their own words, what they thought about the questions, and how they chose an answer. A think-aloud approach was used to conduct cognitive interviews. A trained research assistant asked what the patients were thinking about each question and their corresponding response.

Appendix B. Examination of local independency.

	r
Factor 1: Adherence to regimen	
Items 1 and 5	-0.1371
Items 1 and 7	-0.3242
Items 1 and 8	-0.2087
Items 1 and 9	-0.3172
Items 5 and 7	-0.3539
Items 5 and 8	-0.3086
Items 5 and 9	-0.2825
Items 7 and 8	-0.2663
Items 7 and 9	-0.0449
Items 8 and 9	-0.2176
Factor 2: Consulting behavior	
Items 2 and 3	-0.3032
Items 2 and 4	-0.2784
Items 2 and 6	-0.2633
Items 3 and 4	-0.4244
Items 3 and 6	-0.4500
Items 4 and 6	-0.2411

Note: absolute $r < 0.4$ indicates low local independency.

Appendix C. Step and average measures

	Step measure	Average measure
Factor 1: Adherence to regimen		
Item 1		
Score 1	--	-2.39
Score 2	-1.67	-0.54
Score 3	0.10	0.58
Score 4	0.26	1.57
Score 5	1.31	3.12
Item 5		
Score 1	--	-2.32
Score 2	-1.41	-0.79
Score 3	-0.88	0.42
Score 4	0.60	1.75
Score 5	1.68	3.48
Item 7 ^a		
Score 1	--	--
Score 2	--	-3.23
Score 3	-1.53	-1.25
Score 4	0.22	0.37
Score 5	1.31	2.13

Item 8 ^a		
Score 1	--	--
Score 2	--	-3.09
Score 3	-1.28	-1.26
Score 4	0.17	0.20
Score 5	1.11	1.87
Item 9 ^a		
Score 1	--	--
Score 2	--	-2.93
Score 3	-1.80	-0.81
Score 4	0.22	1.02
Score 5	1.58	2.89
Factor 2: Consulting behavior		
Item 2		
Score 1	--	-4.31
Score 2	-3.62	-1.70
Score 3	-0.60	0.74
Score 4	1.30	2.61
Score 5	2.92	4.59
Item 3		
Score 1	--	-6.19
Score 2	-4.30	-3.02
Score 3	-0.16	0.03
Score 4	2.13	1.53
Score 5	2.33	3.02
Item 4		
Score 1	--	-5.24
Score 2	-4.33	-2.50
Score 3	-1.08	0.36
Score 4	1.37	2.95
Score 5	4.04	5.42
Item 6		
Score 1	--	-5.03
Score 2	-4.00	-2.41
Score 3	-1.00	0.44
Score 4	1.75	2.64
Score 5	3.25	4.60

^aNo responses for Score 1 in these items.

TABLE 1 Participants characteristics (N=380)

	Mean±SD	Median (IQR)	n (%)
Demographic variables			
Age (Year)	61.7±9.1	62.0 (55.0, 69.0)	
Gender (Male)			230 (60.5)
Years of education	6.4±3.7	5.0 (5.0, 8.0)	
Body mass index (kg/m ²)	28.2±5.0	27.7 (24.7, 31.5)	
Heart Failure Characteristics			
Duration of heart failure (month)	5.6±3.7	6.0 (2.4, 8.8)	
NYHA classification (II)			235 (61.8)
NYHA classification (III)			91 (23.9)
NYHA classification (IV)			54 (14.2)
Left Ventricular Ejection Fraction (%)			28.1 (11.3)
Medication n (%)			
Diuretic			183 (48.2)
Beta blockers			154 (40.5)
ACE inhibitors			162 (42.6)
Nitrate			47 (12.4)
Digoxin			50 (13.2)

NYHA: New York Heart Association; IQR: interquartile range.

What's New?

- The two-factor structure model (factors are *adherence to regimen* and *consulting behavior*) fits the best for the EHFScB-9 in our advanced statistical methods (i.e., the confirmatory factor analysis using diagonally weighted least squares).
- Except for one item displays differential item functioning (DIF) across New York Heart Association classification (NYHA), no items displayed substantial DIF across gender and NYHA.