


How to Evaluate Online Education for General Practitioners: Development of a Tailored Questionnaire for Heart Failure Education

Journal of Medical Education and Curricular Development
Volume 11: 1–10
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/23821205241232497



Willem Raat¹ , Evelyne Housiaux¹, Miek Smeets¹, Stefan Janssens², Birgitte Schoenmakers¹ and Bert Vaes¹

¹Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium. ²Department of Cardiovascular Diseases, University Hospitals Leuven, KU Leuven, Leuven, Belgium.

ABSTRACT: Physician-oriented online education could be a pathway to improve care for patients with heart failure, however, it is difficult to measure the impact of such education. Self-efficacy is a potential outcome measure. In this article, we develop a methodology for analyzing an educational intervention for general practitioners (GPs) using self-efficacy as a concept. This study was partly conducted within the setting of an observational study, IMPACT-B, where we developed online education for GPs. We designed and refined a 24-item questionnaire using item analysis, and exploratory and confirmatory factor analysis. Ninety-one GPs completed the questionnaire before and after the online education. Follow-up data after 6 months was available for 13 GPs. Item analysis revealed a high degree of internal consistency (coefficient alpha 0.95) and validity. Each additional year of experience was associated with an average baseline self-efficacy score of 0.50 points (95% CI [0.21-0.80]), and each additional patient in HF follow-up with an average score of 2.0 points (95% CI [0.48-3.5]). Items that differentiated most between GPs with high and low self-efficacy were the treatment of congestion as well as titrating medication and MRA in heart failure with reduced ejection fraction. Factor analysis reduced the number of questions to 14, mapping to three factors (diagnosis, treatment, and follow-up), and improved the model fit as measured by the goodness-of-fit indicator comparative-fit-index (from 0.83 to 0.91). We demonstrated a method to assess the impact of online education on general practitioners. This led to a questionnaire that was reliable, valid, and convenient to use in an implementation context.

KEYWORDS: Heart failure, PROM, general practitioner, self-efficacy, qualitative design

RECEIVED: November 18, 2023. **ACCEPTED:** January 29, 2024.

TYPE: Methodology

FUNDING: The authors received no financial support for the research, authorship, and/or publication of this article.

DECLARATION CONFLICTING INTERESTS: The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Willem Raat, Evelyne Housiaux, Miek Smeets, Birgitte Schoenmakers, and Bert Vaes report no conflict of interest. S. Janssens is the holder of a named chair in Cardiology at the University of Leuven financed by AstraZeneca.

CORRESPONDING AUTHOR: Willem Raat, Department of Public Health and Primary Care, KU Leuven, Kapucijnenvoer 7, blok D bus 7001, 3000 Leuven, Belgium.
Email: willem.raat@kuleuven.be

Introduction

Heart failure (HF) is an important and growing health concern.¹ Patients are mainly older and fragile. Thus, family physicians or general practitioners (GPs) are ideally placed to deliver care.² However, the adequate diagnosis and treatment of HF in this older and fragile primary care population presents GPs with unique difficulties, and they report a lack of knowledge of recent guidelines as well as diagnostic uncertainty.^{3–6} Several studies conducted in primary care settings therefore sought to improve HF outcomes at the patient level by offering tailored HF education to GPs in place of or in addition to patient education and self-management support, albeit with limited success.^{7,8}

This type of physician-oriented medical education is an integral part of quality improvement initiatives targeting the healthcare provider aspect of the so-called quadruple aim.⁹ It is, however, difficult to measure the effectiveness of these interventions. One particular outcome could be self-efficacy, defined as an individual's belief in their capacity to act in the ways necessary to reach specific goals.¹⁰ A high level of self-efficacy indicates a higher ability to use certain skills, even in more demanding environments and vice versa. This outcome has been defined and used as a proxy for HF patient self-management for several decades^{11–13} and has recently been defined for caregivers as well.¹⁴ Self-efficacy is also

related to competence, since individuals with high self-efficacy are more likely to engage in challenging tasks, persist in the face of setbacks, and achieve their goals, leading to a positive feedback cycle on actual competence. In occupational settings, several meta-analyses demonstrated an association between higher self-efficacy and workplace motivation and performance.^{15–17} As with other self-beliefs, this relationship can be prone to overestimation.¹⁸

However, to our knowledge, there is no instrument to measure GPs' (or other physicians') self-efficacy in heart failure, nor is there a clear methodology for developing instruments to assess educational interventions' effectiveness. We therefore aim to develop a methodology for assessing educational interventions using self-efficacy as a concept.

Methods

Design and Recruitment

This study took place in Belgium from January 2020 until February 2021 and was partly conducted within the setting of an observational study, the IMPACT-B study.¹⁹ This project aims to improve HF care by implementing, among others, GP HF education and reimbursement of natriuretic peptides. Each participating GP had to follow an online HF education session (a narrated slideshow video lasting half an hour).



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without

further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

This education session (Supplemental File 1) summarized current HF guidelines (based on the guidelines available in 2020) and was reviewed by several general practitioners with expertise in heart failure. Inclusion criteria for participation in the study were simple: participants needed to be active GPs and needed to give informed consent. We had no predefined exclusion criteria. We recruited GPs from the IMPACT-B study, as well as by email and online from a network of different vocational training and professional institutions to complete the online HF education and test the corresponding questionnaire. Participants entered informed consent and data on the survey platform Qualtrics®. A first set of questions (pretest data) was followed by the education session after which a second set of questions was completed (posttest score). Completing this entire trajectory took about 45 min. After 6 months, we asked GPs participating in the IMPACT-B study to fill out the survey again via email or telephone and sent 2 reminders via email.

Development of Questionnaire

Design. We used Bandura's concept of self-efficacy, defined as an individual's belief in their capacity to act in the ways necessary to reach specific goals.¹⁰ The design of our questionnaire to assess the impact of an educational intervention consisted of 2 steps. First, a literature review of articles that reported the results of online training on self-efficacy of healthcare providers to look for validated questionnaires (MEDLINE database, 2000-2021, varying Mesh terms based on the concepts of health personnel, online education, and self-efficacy). Second, the design of a first questionnaire based on the findings of the review, including a preliminary test survey among five GPs (without special expertise in HF) to check the questions' clarity. We found 17 articles that explicitly reported self-efficacy questions (Supplemental File 2). They were heterogeneous in form and content. There was no standardized and/or validated self-efficacy questionnaire for evaluating the impact of online education. We therefore based the design of our self-efficacy questionnaire on the principles reported by Bandura²⁰ (and the contents of the educational session). We formulated 24 questions concerning 6 domains of heart failure care (see Supplemental File 3 for the full questionnaire). Five questions regarded the diagnosis of HF. Six questions regarded the treatment of all types of HF, 6 regarded HF with reduced ejection fraction (HFrEF) specifically. Three pertained to follow-up, 3 to patient communication and one to advanced care planning.

Refinement. We then conducted extensive testing of the questionnaire in the broader GP population and combined item and factor analysis to reduce the number of items in our questionnaire and test internal consistency. The goal of the item analysis was to describe the contribution of different test items to our instrument and the validity of our instrument. The goal of

the factor analysis was to identify possible different subdomains or factors within our instrument (exploratory factor analysis) and to conduct formal hypothesis testing (confirmatory factor analysis). Finally, we explored the questionnaire's validity within a nested observational study.

Outcomes

Our primary objective was to develop a methodology to assess an educational intervention among physicians, in our case by developing a physician self-efficacy questionnaire for HF. The outcome for each survey item was a Likert scale between 0 (*no competence*) and 100 (*full competence*). Our main outcomes for the questionnaire were internal consistency, validity, availability, and usability. A secondary outcome was the impact of an educational module on self-efficacy.

Data Collection and Analysis

We used descriptive statistics (means, standard deviations, medians, and interquartile ranges where appropriate) for participant characteristics.

We used pretest data for item and factor analysis. We tested 3 statistics for the initial item analysis: item difficulty, discrimination and alpha-if-item deleted. Item difficulty was expressed as the mean score and standard deviation of each response item. Item discrimination was estimated using the correlation between item responses and construct scores and expressed in item-total (ITC) and corrected item-total correlations (CITC). Internal consistency was estimated using alpha coefficients and different item's contributions to internal consistency were measured by estimating alpha with that item removed from the set (alpha-if-item-deleted). We computed a matrix of Pearson's correlation coefficients based on all possible pairs of questionnaire items and calculated *p*-values. We used a hierarchical clustering method with 4 possible clusters to visualize items in a correlation plot. We then conducted a factor analysis to identify multiple unobserved variables, or *factors*, that explained the correlations among our observed variables (in this case test item scores).²¹ In an exploratory factor analysis we explored the possible number and types of factors that explained correlations within the questionnaire and how these factors related to each other by calculating eigenvalues (numerical values that express the total variance explained by each factor). We first fitted a 10-factor structure and evaluated the results in a scree plot, which visualizes factors in decreasing order of eigenvalue, where values of 1 or more are considered the cutoff for acceptability.²¹ We then decided upon the optimal number of factors based on the strength of underlying constructs (scree effect) and how well each item related or loaded, on these hypothesized factors. We therefore consecutively fitted a 1-factor, 2-factor, 3-factor, and 4-factor analysis respectively and discussed which structure was most appropriate to explain possible constructs or subdomains within the

questionnaire. In confirmatory factor analysis, we investigated the appropriateness of a given factor structure and formally compared different hypotheses as to which items best assess the factors we're interested in measuring. Our main metrics for comparing model fits were comparative-fit-index (CFI), and root mean square error of approximation (RMSEA). We conducted factor analyses using the "lavaan" package and "factanal" function in R 4.2.1 using a varimax rotation method and maximum likelihood estimator, with nonlinear minimization to optimize the confirmatory factor analysis.²² We tried to improve model fits using modification indices (MI). These are values that indicate how much a model fit would improve when a particular path (such as a covariance), were added into the model. Since these are a form of post hoc model modification, we judiciously assessed any identified MIs on plausibility. Finally, we investigated validity and usability by testing the association between self-efficacy measures at baseline and experience with HF care, as well as longitudinal measurements on the self-efficacy instrument. These last measurements were

assessed in an exploratory boxplot, depicting the mean and interquartile range for all given question responses across 3 different sections.

Ethics

Participants needed to explicitly provide informed consent to analyze their responses before completing the survey. This procedure was approved by the ethical committee of the University Hospital Leuven (MP012922) on April 21st, 2020.

Results

Step 1—Questionnaire Design and Testing

Participant Characteristics. Table 1 depicts the characteristics of participating GPs. There were 91 respondents, of which the majority were female. More than half of all participants ($n=50$) were recruited from within the IMPACT-B study. Participants had experience with heart failure patients, with about half of all participants ($n=43$) counseling 6 HF patients or more, and more than half ($n=47$) having diagnosed at least 2 HF patients in the previous 3 months. Thirteen GPs from the IMPACT-B study repeated the questionnaire after 6 months.

Table 1. Participant characteristics.

CHARACTERISTIC	$N=91^a$
Sex	
Female	63 (69%)
Male	28 (31%)
Formation	
GP	65 (71%)
GP trainee	26 (29%)
Experience	
More than 10 years	27 (30%)
Between 5 and 10 years	23 (25%)
Less than 5 years	41 (45%)
Source of referral	
IMPACT-B study	50 (55%)
Other	41 (45%)
Number of HF patients	
Between 2 and 5 patients	47 (52%)
More than 6 patients	43 (47%)
Less than 2 patients	1 (1.1%)
Number of new HF diagnoses in the past 3 months	
>2 times	5 (5.5%)
1-2 times	45 (49%)
Never	41 (45%)

Abbreviations: GP, general practitioner; HF, heart failure.

^a n (%).

Item analysis. Table 2 depicts the item analysis for all questionnaire items. The consistency on individual items and on the entire questionnaire was very high with a total coefficient alpha of 0.95. The highest discriminating items (highest CITC scores) were the treatment of HFrEF with MRA, titration of HF medication in patients with HFrEF and adequate treatment of congestion.

Figure 1 demonstrates a correlation plot of the 24 different item responses, based on a correlation matrix (Supplemental File 4) computed by calculating the correlations (Pearson's correlation coefficients) between each pair of questionnaire items. In this plot, items have been ordered according to a hierarchical clustering algorithm with 4 possible clusters. Hues of blue indicate the level of correlation (Pearson's correlation coefficient) between items and asterisks the corresponding p -values. This correlation plot indicates one strong (dark blue) cluster for items 8-15 and 18, as well as 2 weaker (lighter blue) ones. Item 4 (coding) correlated poorly with most other response items.

Validity and Usability. We hypothesized that self-efficacy would be greater in those physicians with a higher degree of clinical experience, particularly with HF. Figure 2 demonstrates the relationship between average self-efficacy scores per item and clinical experience.

Table 3 quantifies the associations. Each year of extra experience was associated with a 0.50-point average score increase per item, each extra HF patient with a 2.0-point increase and more than 2 recent HF diagnoses with an average increase of 7.3 points.

Table 2. Item analysis of the POSE-HF questionnaire.

QUESTION ITEM	ITEM ABBREVIATION	M	SD	n	na	ITC	CITC	Aid
<i>Diagnostics of heart failure</i>								
1. Recognizing symptoms and signs of heart failure	signs	64.45	15.50	91	0	0.61	0.59	0.95
2. Requesting the appropriate work-up when there is a suspicion of heart failure	exam	61.37	17.16	91	0	0.69	0.66	0.95
3. Using the biomarker NT-proBNP in the diagnosis of heart failure	ntprobnp	58.19	24.89	91	0	0.53	0.48	0.95
4. Registering the diagnosis "heart failure" in the electronic health record (EHR)	coding	74.78	22.83	91	0	0.37	0.32	0.96
5. Registering parameters linked to heart failure (NYHA, ejection fraction) in the EHR)	parameters	45.55	26.05	91	0	0.57	0.52	0.95
<i>Treatment of all types of heart failure</i>								
6. Giving concrete instructions regarding salt restriction and weight follow-up	salt	49.18	21.40	91	0	0.62	0.58	0.95
7. Monitoring influenza and pneumococcal vaccination in patients with heart failure	vaccination	68.68	22.68	91	0	0.58	0.54	0.95
8. Starting diuretics in the presence of signs of congestion	congestion	65.22	20.34	91	0	0.82	0.80	0.95
9. Requesting necessary laboratory tests in patients under treatment with diuretics	labs	67.69	20.57	91	0	0.72	0.70	0.95
10. Monitoring clinically the volume status in patients under treatment with diuretics	volume	60.33	21.56	91	0	0.81	0.78	0.95
11. Using a flexible diuretics treatment: up-titrating and tapering diuretics as needed	diuretics	58.74	23.23	91	0	0.83	0.81	0.95
<i>Treatment of HFrEF</i>								
12. Starting neurohormonal blockage (ACE inhibitor or sartan or beta-blocker) in all patients with HFrEF	ace	56.68	24.72	91	0	0.82	0.80	0.95
13. Requesting necessary laboratory tests in patients under treatment with neurohormonal blockage	acelabs	65.97	21.00	91	0	0.73	0.70	0.95
14. Starting mineralocorticoid receptor antagonists (such as spironolactone) in patients with HFrEF, who had insufficient effect from treatment with an ACE inhibitor or sartan and a beta-blocker	mra	49.68	25.23	91	0	0.87	0.85	0.95
15. Requesting appropriate laboratory tests in patients under treatment with mineralocorticoid receptor antagonists (such as spironolactone)	mralabs	59.67	24.09	91	0	0.78	0.75	0.95
16. Referring patients with HFrEF for specialist treatment, when a pharmacological treatment with ACE inhibitor, beta-blocker, and mineralocorticoid receptor antagonist (such as spironolactone) is insufficient.	hfrfreespecialist	68.81	21.33	91	0	0.65	0.62	0.95
17. Conducting a yearly check-up of iron status in patients with HFrEF and HFmrEF and treating iron deficiency appropriately	hfrfiron	45.78	26.91	91	0	0.58	0.53	0.95
<i>Follow-up of heart failure</i>								
18. Up-titrating heart failure medication to target dosage	titration	50.71	23.15	91	0	0.83	0.81	0.95
19. Conducting a clinical check-up of heart failure patients within a week after discharge from the hospital	discharge	57.09	25.45	91	0	0.80	0.77	0.95
20. Conducting a clinical check-up of a heart failure patient every 3 months in a stable phase	stable	63.46	23.55	91	0	0.77	0.75	0.95
<i>Patient communication</i>								
21. Explaining the pathophysiology of different types of heart failure (HFrEF, HFpEF, HFmrEF) to the patient	types	49.23	24.79	91	0	0.74	0.71	0.95
22. Discussing the prognosis of heart failure with the patient	prognosis	47.97	24.26	91	0	0.78	0.75	0.95
23. Discussing the importance of therapeutic compliance (both pharmacological and nonpharmacological) with the patient and promoting this	compliance	65.00	19.32	91	0	0.71	0.68	0.95

(continued)

Table 2. Continued.

QUESTION ITEM	ITEM ABBREVIATION	M	SD	n	na	ITC	CITC	Aid
<i>Advanced care planning</i>								
24. Starting and applying advanced care planning in patients with (or at risk of developing) terminal heart failure	acp	57.09	22.65	91	0	0.67	0.63	0.95

Abbreviations: NT-proBNP, N-terminal pro brain-natriuretic peptide; NYHA, New York Heart Association scale for heart failure; HF_rEF, heart failure with reduced ejection fraction; HF_pEF, heart failure with preserved ejection fraction; HF_{mr}EF, heart failure with mid-range ejection fraction; ACE, angiotensin-converting enzyme; Signifier, abbreviations for the different questions used in the manuscript; M, mean; SD, standard deviation; n, number of questionnaires; na, incomplete responses; ITC, item-total correlation; CITC, corrected item-total correlation; Aid, alpha-if-item-deleted.

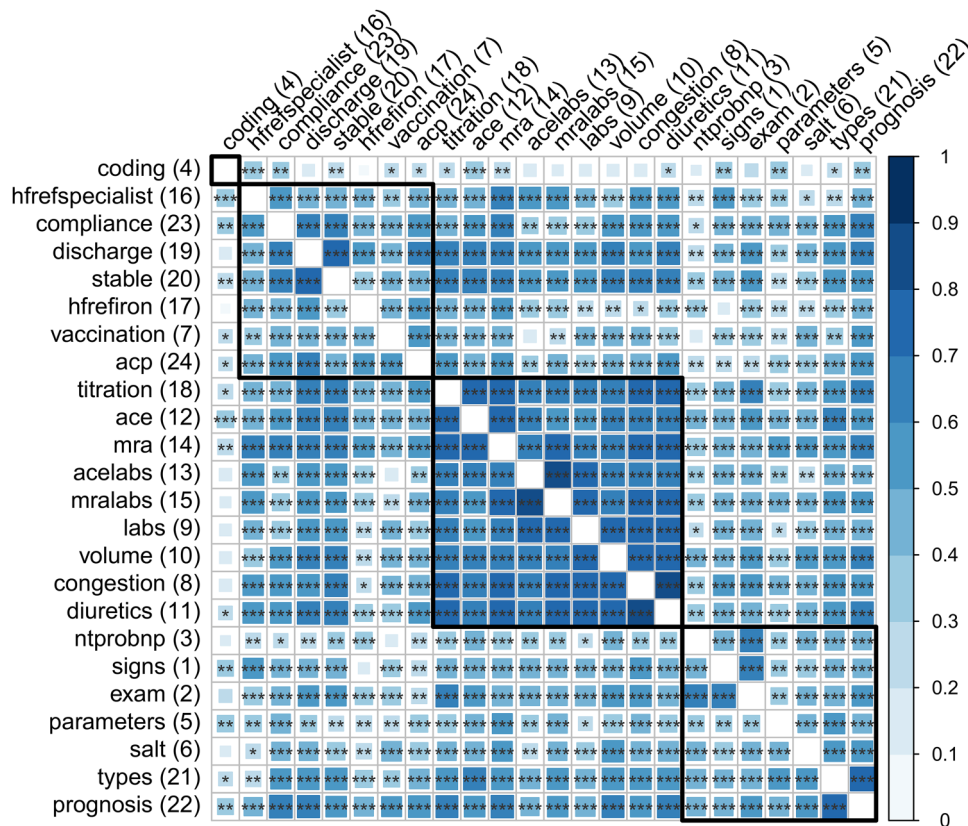


Figure 1. Questionnaire correlation plot indicating the degree of correlation between scores on each pair of questionnaire item responses. Items are ordered according to a hierarchical clustering algorithm with four different clusters (black squares). Hues of blue indicate the degree of correlation (Pearson's correlation coefficient) between items. Asterisks indicate the degree of significance: $<.001$ (***), $<.01$ (**), $<.05$ (*).

With regard to usability, we used the questionnaire to estimate the efficacy of a training module on increasing self-efficacy. Figure 3 illustrates the capacity of the questionnaire to capture an increase in perceived self-efficacy after the completion of a training session on HF (based on factors identified in the factor analysis, as explained in Step 2). The median time to complete the HF questionnaire both before and after the session (omitting the time to view the session itself) was 18 min (IQR: 7.29).

Step 2—Questionnaire Refinement

After our item analysis revealed significant clustering, we were interested in whether we could remove superfluous questions

and identify possible underlying factors, thereby improving the performance of our questionnaire in capturing a true self-efficacy construct and reducing time to completion.

Exploratory Factor Analysis. A scree plot analysis (Figure 4) revealed a strong scree effect with multiple factors having an eigenvalue higher than 1 (the cutoff for acceptability). After analyzing model fits for a 1-factor, 2-factor, 3-factor, and 4-factor exploratory factor analysis respectively, as well as the correlation plot, we hypothesized a 3-factor structure to be the most appropriate to explain possible internal constructs or subdomains within the item questionnaire. Item loadings for this 3-factor structure are depicted in Table 4.

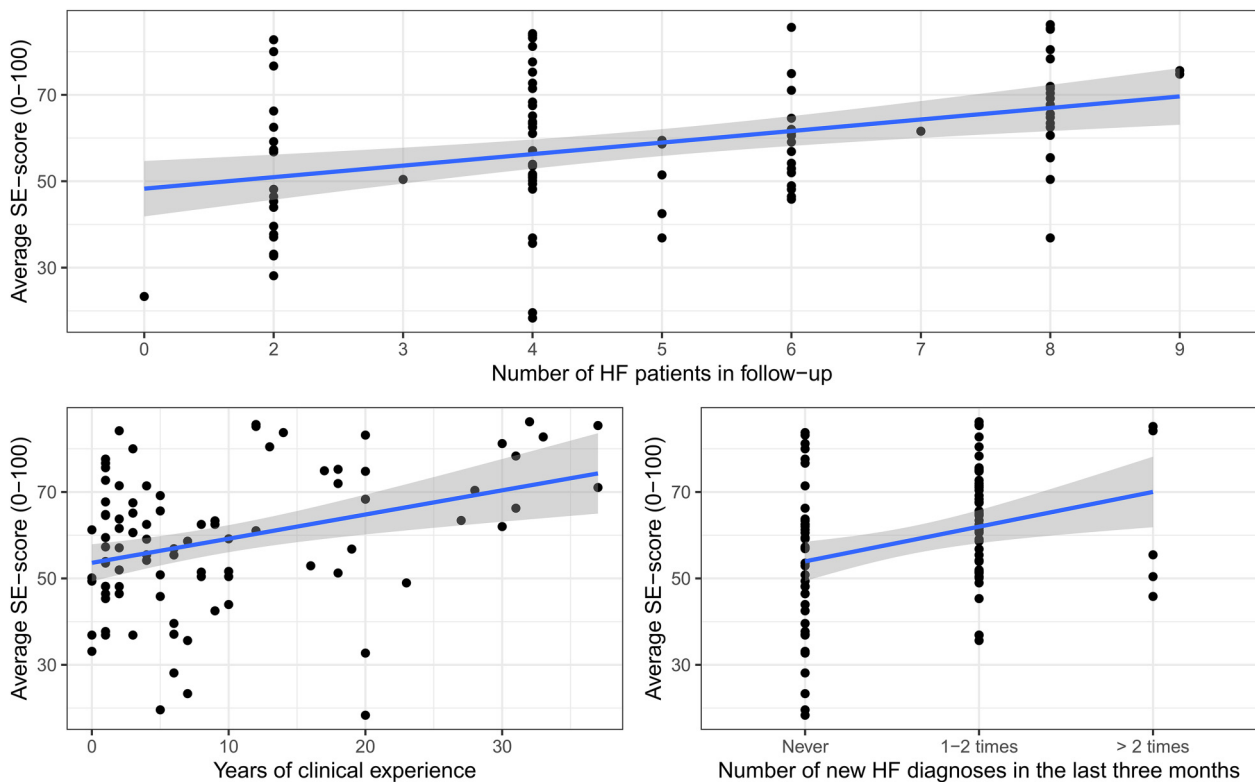


Figure 2. Association between average self-efficacy score per item and measures of clinical experience for participating general practitioners ($n = 91$). Each dot represents the mean self-efficacy score per item. HF, heart failure.

Table 3. Association and 95% confidence intervals (CI) for years of experience, number of HF patients in follow-up, new HF diagnoses and average total score per question item.

CHARACTERISTIC	BETA	95% CI ^a	P-VALUE
Experience	0.50	0.21, 0.80	.001
Number of HF patients in follow-up	2.0	0.48, 3.5	.012
New HF diagnoses in the last 3 months			
Never	—	—	
1-2 times	4.4	-2.4, 11	.2
>2 times	7.3	-5.7, 20	.3

^aCI, confidence interval.

We hypothesized that these different factors could be translated to have the following meaning: HF diagnosis (Factor 3), HF treatment (Factor 1), and HF follow-up (Factor 2). Figure 5 visualizes each item's loading on the 3 factors.

Items 1-3 loaded strongly on the “diagnosis” factor. Items 8-16 and 18 loaded strongly on the “treatment” factor. Items 6, 7, 17, and 19-24 loaded strongly on the “follow-up” domain. The items corresponding to the EHR (4 and 5) loaded weakly on the “follow-up” factor.

Confirmatory Factor Analysis. Based on the findings of our exploratory analysis, we investigated 3 different models to determine whether our hypothesized factor structure adequately captured certain percentages of the variability in scores. A first 3-factor model without the “coding” and “parameters” queries since they did not map strongly to different factors and were the least discriminating items. A second reduced 3-factor model with queries with loading values of 0.5 or higher on the diagnosis factor and 0.6 or higher on the treatment and follow-up factors. This excluded 8 items (salt, vaccination, ace, hfredspecialist, hfredfron, titration, stable, types), thus leaving 14 items in the questionnaire. Third, a version of this second model that includes the five highest-scoring MIs pathways. These indices seemed plausible: for example, in our second model, the highest scoring MI was the correlation between the “acelabs” and “mralabs” questions, which was a valid MI to add (Table 5).

Using factor analysis, we could improve the CFI and RMSEA to values close to the recommended cut-offs of 0.95 and 0.06 for good model data fit.²³

Discussion

We demonstrated the creation of an instrument for measuring the impact of online education on healthcare practitioners' self-efficacy following four steps: a review of the literature, design, testing and refinement. This led to a questionnaire that was reliable, valid, and convenient to use.

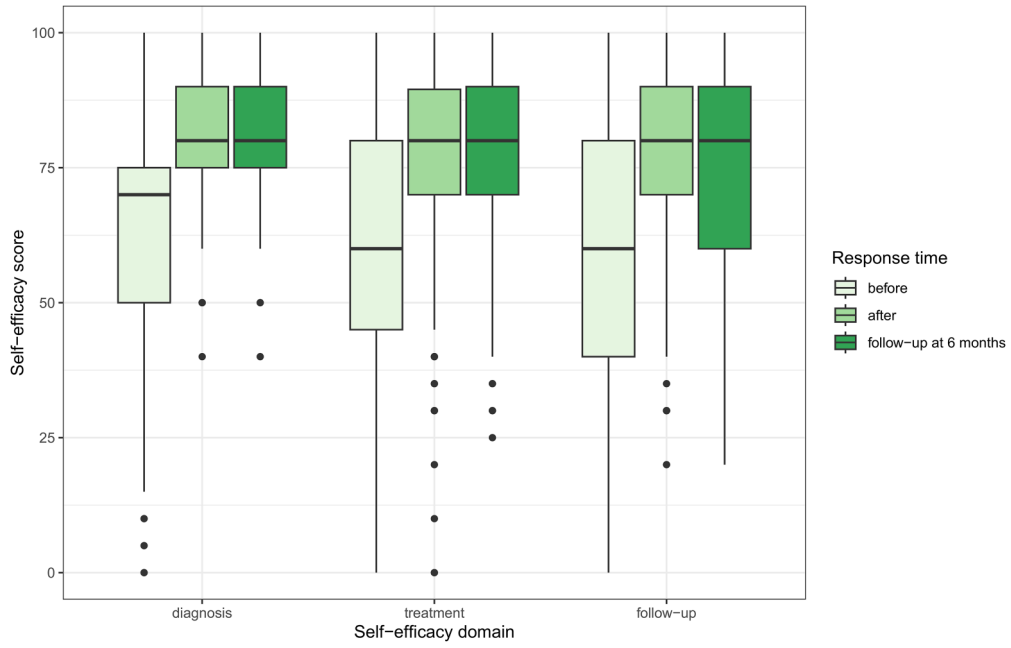


Figure 3. Longitudinal overview of average scores per item per factor across 3 different time points: before- and immediately after an online training session and after 6 months follow-up.

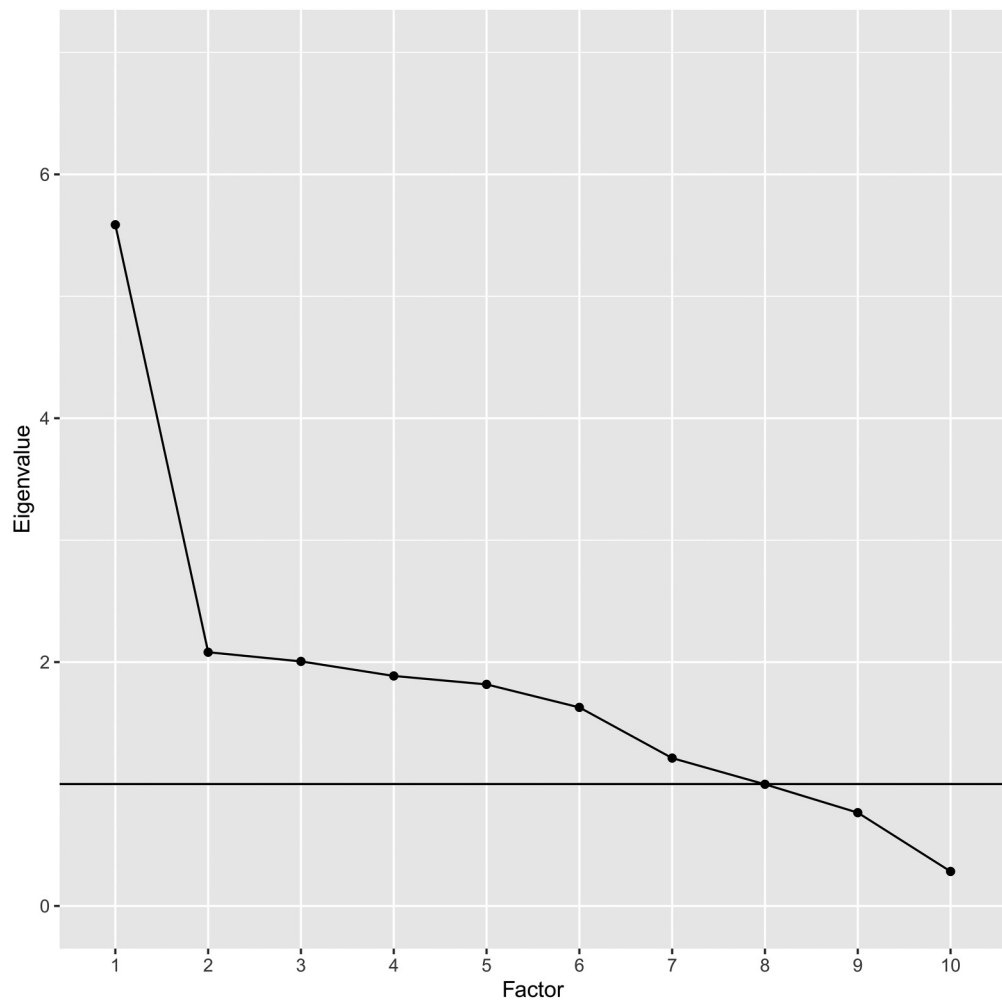


Figure 4. Scree plot indicating eigenvalues of an exploratory analysis with 10 hypothesized factors and corresponding eigenvalues.

To our knowledge, this is the first study that concretely demonstrates how to conceive and analyze a questionnaire to assess online education for health care practitioners within an implementation context. Our study differs from other studies investigating the effects of online training on self-efficacy outcomes in healthcare practitioners identified in our review. First, with regard to data analysis. Only 3 such studies used analytic methods other than simple parametric or nonparametric null-hypothesis significance testing: Blackman et al used path analysis²⁴; Heard et al used mixed-methods ANOVA.²⁵ Leung et al used mixed methods modeling.²⁶ None conducted item or factor analysis of their questionnaire to assess the discriminating value of items, their match with a hypothesized underlying construct or mapping onto meaningful factors. In our study, for example, this was particularly useful since we saw that our question items did not map clearly to our first list of 6 hypothesized factors, but rather to 3 different ones. We also saw that the items corresponding to accurate registration of HF in the electronic health record, an important topic for quality improvement,²⁷ corresponded little to the general hypothesized “HF self-efficacy” construct. Such a finding can be useful within an implementation context since it suggests that GPs’ confidence in the achievement of goals in HF management stands separate from their confidence in accurately registering HF (or other chronic conditions). In this example, this might mean that adequate registration of HF and chronic disease should be the focus of a separate education and self-efficacy instrument. Second, all studies used Likert scales ranging from 3 to 10 points rather than a continuous numerical scale, which could impact results: scales with few response options are less sensitive and reliable due to the omission of differentiating information.²⁸ A self-efficacy scale with multiple gradations of strength of self-efficacy is therefore a stronger predictor of performance than one with only a few choices.²⁹

Since self-efficacy is an important contextual element in frameworks to evaluate implementation processes such as CFIR or PariHS,^{30,31} our methods to develop a tailored questionnaire can be useful for other researchers. Firstly, with regard to an analysis of the implementation process. Before implementation, a questionnaire can be used to identify possible targets for improvement or education. Questionnaire analysis can also be useful during the implementation itself. For example, in our case of HF, it could be useful to assess the adoption of new guidelines concerning the use of SGLT2 inhibitors in HF by adding a new item to the questionnaire. One would then initially expect a low level of consistency and factor loading. Longitudinal follow-up within the context of a quality improvement initiative aimed at stimulating guideline-recommended care might then demonstrate a gradual convergence and improved coefficient alpha and factor loading on the “treatment” domain. Such a longitudinal follow-up of self-efficacy scores can then also be used as an outcome measure for the program’s impact. Secondly, with regard to the translation of research findings. Since implementation research has a cyclical

Table 4. Exploratory factor analysis with factor loadings on a three-factor structure.

VARIABLE	UNIQUENESS	FACTOR 1	FACTOR 2	FACTOR 3
Signs	0.50	0.32	0.26	0.58
Exam	0.00	0.28	0.26	0.92
Ntprobnp	0.61	0.18	0.25	0.55
Coding	0.88	0.11	0.31	0.10
Parameters	0.71	0.28	0.42	0.16
Salt	0.63	0.26	0.50	0.23
Vaccination	0.58	0.13	0.59	0.23
Congestion	0.26	0.70	0.41	0.28
Labs	0.33	0.74	0.25	0.25
Volume	0.32	0.63	0.45	0.30
Diuretics	0.24	0.70	0.48	0.19
Ace	0.35	0.55	0.54	0.25
Acelabs	0.17	0.88	0.12	0.21
Mra	0.23	0.65	0.56	0.19
Mralabs	0.18	0.85	0.24	0.18
Hfrefspecialist	0.59	0.50	0.36	0.17
Hfrefiron	0.66	0.21	0.50	0.20
Titration	0.30	0.57	0.51	0.35
Discharge	0.36	0.47	0.59	0.26
Stable	0.40	0.48	0.56	0.24
Types	0.47	0.40	0.56	0.24
Prognosis	0.35	0.31	0.69	0.27
compliance	0.35	0.21	0.76	0.17
acp	0.43	0.25	0.71	0.04

Factor loadings higher than 0.5 are in bold font. Uniqueness is an expression of the leftover unexplained variance for an item, after removing the shared variance explained by the factors.

nature, a preliminary questionnaire such as ours can be used in other programs to improve heart failure management by general practitioners as well. A first objective could be to validate the questionnaire by investigating whether the observed item analysis and factor structure generalize to other general practitioner populations. We conclude with a final note on the convenience of using the described methodology. We believe these results to be heartening to other researchers (in primary care or other specializations), since both the method of design²⁰ and analysis²¹ consist of a simple and well-described set of principles and are grounded in a subjective assessment of significance by a team of researchers and health professionals who are most acquainted with routine clinical practice. In our view, this results in a powerful combination of flexibility and precision.

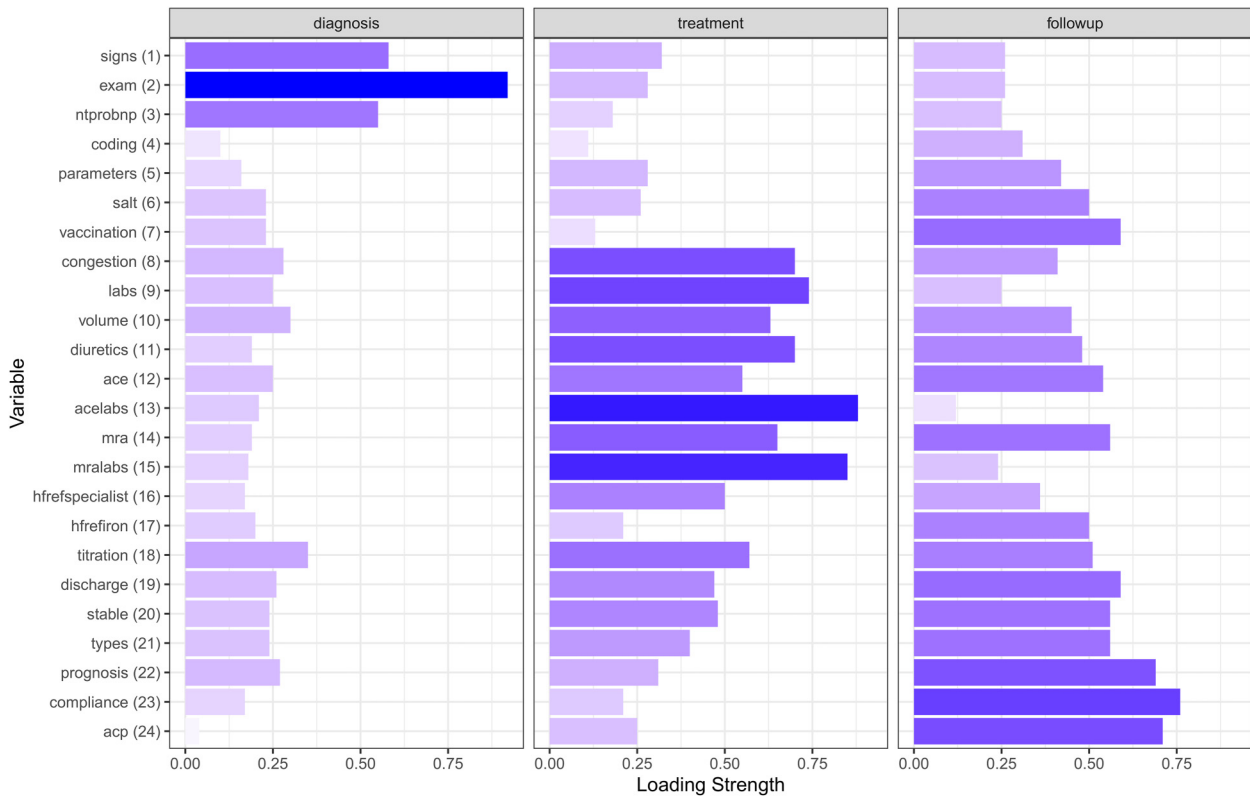


Figure 5. Bar chart of factor loadings for each item on 3 hypothesized factors.

Strengths and Limitations

Our study has several strengths. First, we designed our questionnaire after a systematic literature review and tested it in a large and nationally representative sample of general practitioners. Second, we used and reported a method of data analysis that is original in research surrounding online education and self-efficacy in health care practitioners. Our research has three limitations. First, we did not translate and test our questionnaire outside Belgium. However, the goal was to obtain a tool to measure local implementation rather than to obtain an internationally validated questionnaire. This also limits the relevance of the relatively large fraction of GP trainees in our sample. Belgian GP trainees follow a 3-year postgraduate program³² and largely operate as autonomous physicians.³³ They are therefore also important actors to consider in health implementation processes. Second, our question items sometimes straddle the gray area between the concepts of self-efficacy and self-competence, wherein self-efficacy is defined as a person’s judgments of his or her capabilities to organize and execute courses of actions required to attain designated types of performance and perceived competence as the extent to which a person feels he or she has the necessary attributes to succeed.³⁴ However, the extent to which these dimensions overlap is unclear^{35,36} and the nuance between the 2 concepts has little relevance in the context of our research. Third, we did not perform any sample size or power analysis, since our aim was not to assess any sort of specific effect of online education,

Table 5. Confirmatory factor analysis and indices.

	THREE-FACTOR MODEL (3 F)	REDUCED THREE-FACTOR MODEL (3 FR)	REDUCED AND OPTIMIZED THREE-FACTOR MODEL
Goodness-of-fit indices			
Chisq (df)	488.27 (206)	164.20 (74)	96.12 (69)
P-value	<.01	<.01	.02
CFI	0.83	0.91	0.97
RMSEA (CI)	0.12	0.12	0.07


Abbreviations: df, degrees of freedom; CFI, confirmatory fit index; RMSEA, root mean square error of approximation; CI, confidence interval.

but rather to analyze the appropriateness of our questionnaire item structure.

Conclusion

We demonstrated the creation of an instrument for measuring the impact of online education on healthcare practitioners’ self-efficacy following four steps: a review of the literature, design, testing, and refinement. This led to a questionnaire that was reliable, valid, and convenient to use in an implementation context.

ORCID iD

Willem Raat  <https://orcid.org/0000-0002-9591-3974>

Supplemental material

Supplemental material for this article is available online.

REFERENCES

- Conrad N, Judge A, Tran J, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. *Lancet*. 2018;391(10120):572-580. doi:10.1016/S0140-6736(17)32520-5
- Groenewegen A, Rutten FH. Near-home heart failure care. *Eur J Heart Fail*. 2019;21(1):110-111. doi:10.1002/ehf.1345
- Parnicka A, Wizner B, Fedyk-Lukasiak M, Windak A, Grodzicki T. Knowledge about heart failure in primary care: need for strengthening of continuing medical education. *Cardiol J*. 2013;20(4):356-363. doi:10.5603/CJ.2013.0093
- Howlett J, Comin-Colet J, Dickstein K, Fuat A, Pözl G, Delaney S. Clinical practices and attitudes regarding the diagnosis and management of heart failure: findings from the CORE Needs Assessment Survey. *ESC Heart Fail*. 2018;5(1):172-183. doi:10.1002/ehf2.12205
- Smeets M, Roy SV, Aertgeerts B, Vermandere M, Vaes B. Improving care for heart failure patients in primary care, GPs' perceptions: a qualitative evidence synthesis. *BMJ Open*. 2016;6(11):e013459. doi:10.1136/bmjopen-2016-013459
- Peters-Klimm F, Natanzon I, Müller-Tasch T, et al. Barriers to guideline implementation and educational needs of general practitioners regarding heart failure: a qualitative study. *GMS Z Med Ausbild*. 2012;29(3):Doc46. doi:10.3205/zma000816
- Peters-Klimm F, Campbell S, Müller-Tasch T, et al. Primary care-based multifaceted, interdisciplinary medical educational intervention for patients with systolic heart failure: lessons learned from a cluster randomised controlled trial. *Trials*. 2009;10(1):68. doi:10.1186/1745-6215-10-68
- Vaillant-Roussel H, Laporte C, Pereira B, et al. Impact of patient education on chronic heart failure in primary care (ETIC): a cluster randomised trial. *BMC Fam Pract*. 2016;17(1):80. doi:10.1186/s12875-016-0473-4
- Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. *Ann Fam Med*. 2014;12(6):573-576. doi:10.1370/afm.1713
- Bandura A. Self-efficacy: toward a unifying theory of behavioral change. *Psychol Rev*. 1977;84(2):191-215. doi:10.1037/0033-295X.84.2.191
- Riegel B, Carlson B, Moser DK, Sebern M, Hicks FD, Roland V. Psychometric testing of the self-care of heart failure index. *J Card Fail*. 2004;10(4):350-360. doi:10.1016/j.cardfail.2003.12.001
- 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(1):99-105. doi:10.1093/euhj/hfn007
- 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. doi:10.1016/s1388-9842(02)00253-2
- Durante A, De Maria M, Boyne J, et al. Development and psychometric testing of the European Heart Failure Self-Care Behaviour Scale caregiver version (EHFScB-C). *Patient Educ Couns*. 2021;104(8):2106-2111. doi:10.1016/j.pec.2021.01.013
- Cherian J, Jacob J. Impact of self efficacy on motivation and performance of employees. *Int J Bus Manag*. 2013;8(14):80-88. doi:10.5539/ijbm.v8n14p80
- Bandura A, Locke EA. Negative self-efficacy and goal effects revisited. *J Appl Psychol*. 2003;88(1):87-99. doi:10.1037/0021-9010.88.1.87
- Lunenburg FC. Self-efficacy in the workplace: implications for motivation and performance. *Int J Manag Bus Admin*. 2011;14(1):1-6.
- Jones A, Sheppard L. Self-efficacy and clinical performance: a physiotherapy example. *Adv Physiother*. 2011;13(2):79-83.
- Raat W, Smeets M, Van Pottelbergh G, Van De Putte M, Janssens S, Vaes B. Implementing standards of care for heart failure patients in general practice – the IMPACT-B study protocol. *Acta Cardiol*. 2021;76(5):486-493. doi:10.1080/00015385.2020.1844504
- Bandura A. Guide for constructing self-efficacy scales. *Self-Efficacy Beliefs Adolesc*. 2006;5(1):307-337.
- Albano T. *Introduction to educational and psychological measurement using R*. Accessed January 8, 2024. <https://www.thetaminusb.com/intro-measurement-r/>
- Rosseel Y. Lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48(2):1-36. doi:10.18637/jss.v048.i02
- Xia Y, Yang Y. RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: the story they tell depends on the estimation methods. *Behav Res*. 2019;51(1):409-428. doi:10.3758/s13428-018-1055-2
- Blackman IR, Mannix T, Sinclair PM. Developing renal nurses' buttonhole cannulation skills using e-learning. *J Ren Care*. 2014;40(1):55-63. doi:10.1111/jorc.12047
- Heard R, O'Halloran R, McKinley K. Communication partner training for health care professionals in an inpatient rehabilitation setting: a parallel randomised trial. *Int J Speech Lang Pathol*. 2017;19(3):277-286. doi:10.1080/17549507.2017.1290137
- Leung YW, Wong J, Kiteley C, Ellis J, Esplen MJ. Addressing educational needs in managing complex pain in cancer populations: evaluation of APAM: an online educational intervention for nurses. *Am J Hosp Palliat Care*. 2019;36(7):587-597. doi:10.1177/1049909119832819
- Raat W, Smeets M, Henrard S, et al. Machine learning optimization of an electronic health record audit for heart failure in primary care. *ESC Heart Failure*. 2022;9(1):39-47. doi:10.1002/ehf2.13724
- Norman G, Cairney J. *Health Measurement Scales: A Practical Guide to Their Development and Use*. Vol. 117. 2015. doi:10.1093/acprof:oso/9780199231881.003.0006
- Hartley J, Valiante G, Pajares F. Response format in writing self-efficacy assessment: greater discrimination increases prediction. *Meas Eval Couns Dev*. 2001;33(4):214-221. doi:10.1080/07481756.2001.12069012
- Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. 2009;4(1):50. doi:10.1186/1748-5908-4-50
- Stetler CB, Damschroder LJ, Helfrich CD, Hagedorn HJ. A guide for applying a revised version of the PARIHS framework for implementation. *Implement Sci*. 2011;6(1):99. doi:10.1186/1748-5908-6-99
- LOI - WET. Accessed January 8, 2024. http://www.ejustice.just.fgov.be/cgi_loi/change_lg.pl?language=nl&cl=N&table_name=wet&cn=2010030102
- De Vocht K, Verheyen K, Michels NR. Learning objectives of Belgian general practitioner trainees regarding their hospital training: a qualitative study. *Eur J Gen Pract*. 2022;28(1):173-181. doi:10.1080/13814788.2022.2081319
- Kremer J, Moran A, Walker G, Craig C. *Key Concepts in Sport Psychology*. 2012. doi:10.4135/9781446288702
- Hughes A, Galbraith D, White D. Perceived competence: a common core for self-efficacy and self-concept? *J Pers Assess*. 2011;93(3):278-289. doi:10.1080/00223891.2011.559390
- Rodgers W, Markland D, Selzler AM, Murray T, Wilson P. Distinguishing perceived competence and self-efficacy: an example from exercise. *Res Q Exerc Sport*. 2014;85(4):527-539. doi:10.1080/02701367.2014.961050