Qual Life Res (2013) 22:1391–1403
DOI 10.1007/s11136-012-0268-6

Assessment of proximal outcomes of self-management programs: translation and psychometric evaluation of a German version of the Health Education Impact Questionnaire (heiQ^{TM})

Michael Schuler · Gunda Musekamp · Hermann Faller · Inge Ehlebracht-König · Christoph Gutenbrunner · Roland Kirchhof · Jürgen Bengel · Sandra Nolte · Richard H. Osborne · Monika Schwarze

Accepted: 7 September 2012 / Published online: 18 September 2012

Abstract
Purpose This paper describes the translation, cultural adaption, and psychometric evaluation of a German version of the Health Education Impact Questionnaire (heiQ^{TM}), a widely used generic instrument assessing a wide range of proximal outcomes of self-management programs.
Methods The translation was carried out according to international standards and included forward and backward translations. Comprehensibility and content validity were tested using cognitive interviews with 10 rehabilitation inpatients. Psychometric properties were examined in rehabilitation inpatients ($n = 1{,}202$) with a range of chronic conditions. Factorial validity was assessed using confirmatory factor analysis; concurrent validity was explored by correlations with comparator scales.
Results The items of the German heiQ^{TM} were well understood by rehabilitation inpatients. The structure of the eight heiQ^{TM} scales was replicated after minor adjustment. heiQ^{TM} scales had higher correlations with comparator scales with similar constructs, particularly mental health concepts than with physical health. Moreover, all heiQ^{TM} scales differentiated between individuals across different levels of depression.
Conclusion The German heiQ^{TM} is comprehensible for German-speaking patients suffering from different types of chronic conditions; it assesses relevant outcomes of self-management programs in a reliable and valid manner. Further studies involving its practical application are warranted.

Keywords Self-management · Assessment · Chronic disease · Translation · Validation

M. Schuler (✉) · G. Musekamp · H. Faller
Department of Psychotherapy and Medical Psychology, Rehabilitation Sciences Section, University of Wuerzburg, Klinikstr. 3, 97070 Wuerzburg, Germany
e-mail: m.schuler@uni-wuerzburg.de
URL: www.psychotherapie.uni-wuerzburg.de

I. Ehlebracht-König
Rehabilitation Center Bad Eilsen, Bad Eilsen, Germany

C. Gutenbrunner · R. Kirchhof · M. Schwarze
Coordination Centre Applied Rehabilitation Research, Department of Rehabilitation Medicine, Medical School Hannover, Hannover, Germany

J. Bengel
Department of Psychology, University of Freiburg, Freiburg, Germany

S. Nolte · R. H. Osborne
School of Health and Social Development, Deakin University, Melbourne, VIC, Australia

S. Nolte
Medical Clinic for Psychosomatics Charité, University Medicine Berlin, Berlin, Germany

Introduction

Self-management and patient education programs attempt to promote self-management competencies, empowerment, and participant's acceptance of their chronic condition(s). This is achieved through health professionals imparting knowledge and insight, and providing participants with training on how to incorporate new behaviors into their lives [1–3]. However, efficacy studies of self-management programs often do not address the aforementioned outcomes. Instead, standard clinical or socio-medical outcomes are measured, for example, somatic parameters, quality of life, or return to work. These more distal outcomes depend on

factors that are not directly influenced by self-management programs. For example, distal outcomes may be strongly influenced by the severity of a somatic disease [4, 5]. Accordingly, systematic reviews of the impact of self-management or patient education programs often show little or no change in distal outcomes [6–16].

When only distal outcomes are assessed, the efficacy of self-management programs may be underestimated, and moreover, effects on key early outcomes may be overlooked [17]. Therefore, it is important that researchers and program managers incorporate proximal outcome measures in the evaluation of self-management programs [3, 18]. Proximal outcomes are more directly affected by the intervention than distal outcomes [19] and can be clearly deduced from the contents and goals of self-management programs. As expected, stronger effects in proximal outcomes are often demonstrated empirically [12, 15, 20, 21].

In response to an observed lack of valid measures of the intended proximal outcomes of self-management programs, Osborne and colleagues developed the Health Education Impact Questionnaire (heiQ^{TM}) in Australia [22]. Originally, this generic instrument contained 42 items, assessed on a 6-point Likert response scale, to measure eight independent constructs: *Positive and active engagement in life*, *Health-directed activities*, *Skill and technique acquisition*, *Constructive attitudes and approaches*, *Self-monitoring and insight*, *Health service navigation*, *Social integration and support*, and *Emotional distress.* The items and scales were developed through careful consultation with patients, healthcare professionals, researchers, healthcare managers, and policymakers; the constructs were subsequently validated using rigorous psychometrics. Studies have demonstrated that heiQ^{TM} can be used to display the effects of self-management programs in outpatient and community settings [12, 21, 23, 24]. Since its development, the heiQ^{TM} has become widely applied and has required only minor refinements. The original heiQ^{TM} had 42 items with a 6-point response scale. Analyses during the construction of heiQ^{TM} showed that some items had disordered thresholds, that is, some respondents were unable to differentiate between the two midpoints "slightly agree" and "slightly disagree." Further analysis (unpublished) also suggested two items could be removed without compromising content validity. As a result, the response format was simplified to a 4-point response scale ("strongly disagree" to "strongly agree") and 2 items removed. Generally, higher values in the heiQ^{TM} scales indicate better status, except for *Emotional distress*, in which higher values indicate higher distress. Further information on the heiQ^{TM} can be found elsewhere [2, 22] and on: www.heiQ.org.au.

The eight independent heiQ^{TM} scales were designed to be sensitive to the immediate or proximal outcomes of self-management [17]. Longer-term outcomes of an intervention might be a reduction in disability, improved health-related quality of life, or even prolonging survival. The proximal outcomes were conceptualized as those impacts that are likely to be observable soon after participation in a self-management education program, such as improvements in attitudes associated with the chronic illness (*Constructive Attitudes and approaches*) or particular skills taught in a diabetes or weight loss education program (*Skill and technique acquisition*). Group-based interventions that promote connectedness between participants are likely to result in immediate improvements in *Social integration and support*.

Until now, robust and sensitive questionnaires to comprehensively assess outcomes of self-management programs across chronic disorders have been absent in the German language. Therefore, the heiQ^{TM} was translated and culturally adapted. The rigorous analyses of its factorial validity and reliability are reported in this paper. Furthermore, we conducted a first approach to test concurrent validity of the heiQ^{TM}. Thus far, no studies have systematically examined correlations between heiQ^{TM} scales and other instruments used in self-management program evaluation.

Methods

The research was undertaken in two phases. First, the heiQ^{TM} was translated and culturally adapted to German, and its comprehensibility was tested. Second, psychometric properties were examined.

Phase 1

The translation and cultural adaptation of the heiQ^{TM} was undertaken using a strict protocol conforming to international standards [25–27]. After translating the questionnaire, cognitive interviews [28] were conducted with members of the target population. Ten patients (35–55 years, eight females, all native German speakers) with either orthopedic conditions or heart disease from two hospitals were interviewed. This procedure checked for semantic equivalence [29], comprehensibility, and content validity.

The translation process included one forward and one backward translation with the aid of two professional translators. The forward translation was checked by bilingual researchers (MS, MS, GM, RK, CG, IE, and HF) and was slightly modified in consultation with the forward translator, a native German speaker. The modified version was translated back to English by a native English speaker (backward translator) who had no knowledge of the original heiQ^{TM}. The back translation was then compared with the original heiQ^{TM} by the Australian-based researchers (SN—bilingual) including the author (RHO). A

consensus meeting generated a preliminary German translation. While emphasis was placed on the equivalence between the English and the German version, when discrepancies arose, cultural and conceptual adaptations were preferred over the literal translations.

Phase 2

Sample

Patients from seven rehabilitation hospitals with a range of medical conditions (cancer, chronic pain, heart disease, inflammatory bowel disease, obesity disorder, orthopedic condition, and respiratory disease) were included. Patients completed the survey, that is, heiQ™ as well as other questionnaires to further assess construct validity, at the beginning (T1), at the end (T2), and 3 months after inpatient rehabilitation (T3). A subgroup also completed the heiQ™ 3 weeks before inpatient rehabilitation (T0). Only patients that were able to complete the questionnaires independently were included in the study. All analyses presented in this paper were based on data from T0 and T1.

Factorial validity

Confirmatory factor analyses (CFA) were initially conducted separately for each scale. Evaluation of model accuracy was based on chi-square test and model fit indices such as Comparative fit index (CFI), Root mean square error of approximation (RMSEA), and Standardized root mean square residual (SRMR) [30]. As small and essentially unimportant discrepancies of the data from postulated models are likely to result in statistically significant chi-square values if sample sizes are large as in the present case [31, 32], a significant chi-square was always interpreted in conjunction with other fit indices. For model fit to be interpreted as "acceptable," CFI needed to be above 0.95, RMSEA below 0.06, and SRMR below 0.08 [31, 33]. If a model test exceeds one or more of the cutoff values, expected parameter changes (EPCs) and modification indices (MI) were calculated to estimate type and magnitude of model misspecification [34]. Akaike's Information Criterion (AIC) and Bayes Information Criterion (BIC) were used to compare non-nested models [31].

To test factorial validity, the total sample was divided into a calibration ($N = 603$) and a validation sample ($N = 599$) using a stratified randomization procedure whereby the condition was the grouping variable. First, the total sample was used to test the assumed one-factor measurement models. If evaluation of respective model fit was positive, the model was accepted. Otherwise, a modified model was tested in the calibration sample. To modify a model, statistical criteria (EPC, MI) [34] and content-related considerations were used. If model fit was accepted, it was then tested in the validation sample (cross-validation). Eventually, all final one-factor models were again tested in the total sample. After all one-factor models were confirmed, the full eight-factor model was tested in all three samples.

Reliability

Reliability of each scale was estimated using Raykov's Composite Reliability Coefficient (CRC) score [35, 36]. CRC values can be interpreted like Cronbach's alpha; it requires only a congeneric measurement model [37] and takes the effects of correlated error variances into account [31]. Based on CRCs, Standard Error of Measurement (SEM) [38, 39] was computed. Furthermore, test–retest reliability [intraclass correlation coefficient (3,1)] of each scale was computed [39] in a subsample of $N = 69$ patients with orthopedic disease who had completed the heiQ™ at T0 and T1.

Concurrent validity

To study concurrent validity, the following comparison scales were used: (1) SF-36 [40, 41], a widely used generic instrument for assessing health status with eight subscales divided into Physical and Mental Health scales; (2) IRES-24 [42], a short-form of the IRES 3 [43], a widely used instrument in Germany for assessing subjective health in patients with chronic conditions; (3) Illness Perception Questionnaire-Revised (IPQ-R) [44, 45], an instrument based on the Common Sense Self-Regulation-Model [46] assessing cognitive and emotional representations of an illness; (4) Patient Health Questionnaire (PHQ-9), a short screening instrument for depression that allows a categorical analysis (no depression—other depression—major depression) based on criteria of the Diagnostic and Statistical Manual of Mental Disorders [47]; and (5) Generalized Anxiety Disorder Scale (GAD-7) [48], a short screening instrument to measure anxiety. The latter two instruments are used worldwide for patients with different chronic conditions [49–53].

We made the following hypotheses:

1. Overall heiQ™ scales would have low to moderate associations with the comparator scales with the majority of correlations expected to be below $r = 0.6$, given that they were intended to measure something different than most available scales (for exceptions see below).
2. Most heiQ™ scales will show low to moderate correlations with scales of subjective health, depression, and anxiety; correlations between most heiQ™ scales and with mental health scales are expected to be higher than those with physical health scales given the item content.
3. Most heiQ™ scales will show low to moderate correlations with the following IPQ-R scales: *Personal*

Control, *Coherence*, *Consequences*, and *Emotional Representation*. Especially, *Self-monitoring and Insight* and *Skill and technique acquisition* will show at least moderate correlations with the IPQ-R scale *Personal Control*. Furthermore, *Emotional distress* will show high correlations with the IPQR-Scale *Emotional Representation.* No hypotheses were formulated about correlations between heiQ™ scales and other IPQ-R scales.

4. The heiQ™ scales *Emotional distress*, *Constructive attitudes and approaches*, and *Positive and active engagement in life* will show at least moderate to high correlations with depression, anxiety, and mental health.

Statistical analysis

Confirmatory factor analyses were computed using Mplus 6.1 [54] with Robust Maximum Likelihood (MLR-estimator). To handle missing data, the Full Information Maximum Likelihood (FIML) algorithm was used [55]. Computations with manifest variables were conducted with IBM PASW Statistics 18. In these analyses, missing data were estimated using multiple imputations. Five complete data sets were imputed, and the results of each were combined to build the overall results [56]. The amount of missing data per item was low (0.1–3.0 %). A p value <0.05 was regarded statistically significant unless otherwise stated. Effect sizes for between-group effects were estimated using Cohen's d (with pooled standard deviations of the compared groups as denominator), with $d = 0.2/0.5/0.8$ indicating small/medium/large effects. Correlation coefficients of 0.1/0.3/0.5 were regarded as small/medium/large [57].

Results

Phase 1

After finalizing the translation process, a preliminary German heiQ™ was established. Cognitive interviews showed that items were generally well understood by interviewees in the intended manner. However, based on the responses, 12 items (30 %) required further refinement. All changes were discussed with all project members and the author of the heiQ™.

Phase 2

Sample

The total sample comprised 1,202 patients from seven clinics. A large proportion had rheumatic/orthopedic conditions (40.9 %) or respiratory conditions (28.4 %); 11.3 % were diagnosed with cancer (rectum, colon, or bladder cancer), 11.8 % with inflammatory bowel disease, 4 % with heart disease, and 4 % with other chronic conditions. Sample characteristics are shown in Table 1. No substantive differences between calibration and validation sample were observed regarding socio-demographic parameters (age, sex, education, and income).

Factorial validity

Table 2 displays the results of the CFA for the total sample (results of calibration and validation sample are available on request). The postulated measurement models of *Positive and active engagement in life*, *Constructive attitudes and approaches*, and *Skill and technique acquisition* showed good fit. In contrast, the remaining five scales showed inadequate fit in at least one fit index. When freeing an error covariance in respective measurement models, fit indices improved in a way that model fit was acceptable. For the measurement model of *Emotional distress*, two possibilities for improving model fit were found: A good model fit ($\chi^2 = 24.81$, $df = 8$, $p = 0.002$; CFI = 0.993; RMSEA = 0.042) was achieved by freeing the error covariance between items 4 and 18, but a superior model fit ($\chi^2 = 5.22$, $df = 5$, $p = 0.390$; CFI = 1.00; RMSEA = 0.006) was achieved by deleting one item (item 18). Since this item should not be deleted from the scale prematurely, it was maintained in subsequent analyses involving this scale; the error covariance was freed instead. All eight heiQ™ scales showed good factorial properties.

Factor loadings of all tested models in the total sample are shown in Table 3. Most loadings were between 0.5 and 0.9, indicating a good representation of the items by the underlying factors. The only exception was *Self-monitoring and insight* which had some coefficients between 0.4 and 0.5.

Based on the results shown above, a full eight-factor model was tested in all three samples, whereby latent factors were allowed to correlate. No additional associations between items or between items and factors (cross-loadings) were allowed. As results of all three samples were similar, only the results of the total sample are reported herein. The model exhibits acceptable fit values ($\chi^2 = 2223.96$, $df = 670$, $p < 0.001$; CFI = 0.918; RMSEA = 0.044; SRMR = 0.054).

Correlations between heiQ™ factors and those between manifest heiQ™ scales are displayed in Table 4. Positive correlations were observed between all factors, with correlation coefficients ranging from $r = 0.17$ to $r = 0.95$. Noticeable are the high correlations between *Skill and technique acquisition* and *Self-monitoring and insight* ($r = 0.95$), and *Active engagement in life* and *Constructive attitudes and approaches* ($r = 0.85$). However, testing alternative models such as allowing cross-loadings between

Table 1 Sample characteristics

	Respiratory disease ($n = 343$) %	Inflammatory bowel disease ($n = 130$) %	Heart disease ($n = 49$) %	Cancer ($n = 135$) %	Orthopedic condition[a] ($n = 492$) %	Other[b] ($n = 53$) %	Total ($n = 1,202$) %
Sex							
Male	63.6	35.4	87.8	60.7	32.7	26.4	46.9
Female	36.4	64.6	12.2	39.3	67.3	73.6	53.1
Marital status							
Single	13.8	36.9	4.1	5.3	11.3	24.5	14.4
Married	68.0	47.7	79.6	67.9	68.4	50.9	65.7
Separated/ divorced	14.7	13.1	14.3	6.9	17.2	20.8	14.9
Widowed	3.5	2.3	2.0	19.8	3.1	3.8	4.9
Level of education							
Vocational training	67.6	56.6	36.7	56.8	55.3	61.5	58.6
Technical college	9.6	13.2	28.6	14.4	14.8	11.5	13.5
Higher education	6.3	20.2	26.5	12.8	11.0	11.5	11.5
Other	7.5	7.0	6.1	6.4	8.7	9.6	7.9
No vocational educ.	9.0	3.1	2.0	9.6	10.2	5.8	8.5
Employment status							
Working	75.9	81.5	91.8	12.5	74.6	66.0	69.3
Unemployed	10.0	4.6	0	3.9	11.3	18.9	9.2
Pension	7.1	2.3	2.0	64.8	3.3	1.9	10.8
Other (housewife…)	7.1	11.5	6.1	18.8	10.8	13.2	10.6
Occupation							
Laborer	55.1	9.2	12.2	17.2	27.3	9.4	30.8
Clerk/civil servant	39.0	86.2	73.5	61.2	65.2	86.8	60.9
Self-employed	3.0	3.1	12.2	17.2	4.0	0	5.1
Other	3.0	1.5	2.0	4.3	3.5	3.8	3.2
Income (monthly)							
<1,000 €	12.7	12.4	2.1	7.8	14.5	24.5	13.0
1,000–3,000 €	73.2	57.0	58.3	79.1	66.7	55.1	67.9
>3,000 €	14.0	30.6	39.6	13.0	18.9	20.4	19.1
Age [mean (SD)]	50.9 (9.9)	43.0 (10.3)	42.4 (5.8)	67.2 (11.4)	49.6 (8.7)	47.6 (10.0)	51.3 (11.3)

Due to rounding errors, not all percentages add up to 100 %; [a] $n = 312$ rheumatism; $n = 180$ chronic back pain, [b] $n = 26$ obesity disorder patients, $n = 27$ depressive patients with chronic pain

single items and both factors did not lead to a significant reduction in the factor correlations. In a further assessment, an alternative model with only one factor for all items from both scales was tested and compared with the two-factor models. For *Active engagement in life* and *Constructive attitudes and approaches*, the one-factor model (CFI = 0.923; AIC = 22,599.82; BIC = 22,752.57) shows worse fit values than the two-factor model (CFI = 0.954; AIC = 22,467.61; BIC = 22,625.45). For *Skill and technique acquisition* and *Self-monitoring and insight*, results of the one-factor model (CFI = 0.980; AIC = 22,931.70; BIC = 23,094.64) and the two-factor model are very similar (CFI = 0.980; AIC = 22,933.03; BIC = 23,101.06).

Reliability

Reliability estimates using Raykov's CRC for the accepted models can be classified as moderate (e.g., CRC = 0.71 for *Self-monitoring and insight*) or good (e.g., 0.87 for *Constructive attitudes and approaches*) (Table 2). Test–retest reliability coefficients were somewhat lower ($r_{tt} = 0.60$ for

Table 2 Model fit and reliability indices of the measurement models

Model[a]	χ^2	*df*	*p*	CFI	RMSEA	SRMR	r^b	CRC[c]	r^d_{t0-t1}	SEM[e]
Health-directed activities										
Original	13.729	2	0.001	0.991	0.070	0.015	–	0.83		
1 with 13	2.249	1	0.134	0.999	0.032	0.005	0.25	0.81	0.60	0.33
Positive and active engagement in life										
Original	15.991	5	0.007	0.988	0.043	0.018	–	0.75	0.72	0.27
Emotional distress										
Original	52.690	9	<0.001	0.982	0.064	0.021	–	0.88		
4 with 18	24.807	8	0.002	0.993	0.042	0.013	0.19	0.87	0.76	0.35
Without 18	5.216	5	0.390	1.000	0.006	0.008	–	0.85	0.77	
Self-monitoring and insight										
Original	94.950	9	<0.001	0.915	0.089	0.045	–	0.74		
3 with 17	21.646	8	0.006	0.987	0.038	0.022	0.29	0.71	0.63	0.21
Constructive attitudes and approaches										
Original	23.230	5	<0.001	0.988	0.055	0.016	–	0.87	0.77	0.26
Skill and technique acquisition										
Original	1.044	2	0.593	1.000	<0.001	0.006	–	0.77	0.72	0.27
Social integration and support										
Original	89.778	5	<0.001	0.962	0.119	0.030	–	0.88		
28 with 22	30.346	4	<0.001	0.988	0.074	0.017	0.32	0.86	0.83	0.26
Health service navigation										
Original	53.252	5	<0.001	0.964	0.090	0.027	–	0.87		
29 with 32	21.943	4	0.002	0.987	0.061	0.016	0.25	0.80	0.68	0.24

[a] Numbers of items that are allowed to correlate or are deleted in the modified models; [b] additional correlation between two items in modified model; [c] composite reliability coefficient; [d] test–retest reliability [ICC (3,1)], based on manifest scale values; [e] standard error of measurement

Health-directed activity to $r_{tt} = 0.83$ for *Social integration and support)*.

Concurrent validity

As hypothesized, the heiQ™ scales showed generally low to moderate correlations with most comparator scales. Only one correlation coefficient exceeded 0.6 (see below). Correlations with scales of mental health were slightly higher than those with physical health scales. For example, the range of the correlations between heiQ™ scales and IRES-24 *Subjective health* was between 0.21 and 0.60, while correlations with IRES-24 *Physical health* were between 0.11 and 0.36.

Most heiQ™ scales showed low to moderate correlations with IPQ-R scales *Coherence*, *Consequences*, and *Emotional Representation* (Table 5). The highest correlation was seen between *Emotional distress* and *Emotional representation* ($r = 0.73$). However, only very low correlations with the IPQ-R scale *Personal control* were observed, even no correlations with heiQ™ scales *Self-monitoring and insight* ($r = 0.01$) and *Skill and technique acquisition* ($r = 0.02$).

Further, heiQ™ scales showed low to high correlations with PHQ-9 and GAD-7. As shown in Table 6, patients with major depression or other depression (according to PHQ-9) had significantly lower values in heiQ™ scale scores than those without depression. Effect sizes were moderate or high between patients with major depression and patients without depression.

As expected, moderate to high correlations were found between heiQ™ scales *Emotional distress*, *Constructive attitudes and approaches,* and *Positive and active engagement in life* and measures of anxiety, depression, and mental health.

Discussion and conclusion

Discussion

In this study, the heiQ™ was translated and culturally adapted to German. Comprehensibility of the items was confirmed using cognitive interviews; comparison with other relevant constructs yielded meaningful associations. Using robust and highly restricted CFA procedures, the

Table 3 Item content and factor loadings of original and modified[a] models (total sample)

Item number	Item content	Factor loadings	
		Original model	Modified model
Health-directed activities			
1	On most days of the week I do at least one activity to improve…	0.78	0.72
9	I do at least one type of physical activity every day for…	0.61	0.64
13	On most days of the week I set aside time for healthy activities	0.86	0.81
19	I walk for exercise, for at least 15 min per day, most days…	0.72	0.77
Positive and active engagement in life			
2	Most days I am doing some of the things I really enjoy	0.59	–
5	I try to make the most of my life	0.56	–
8	I am doing interesting things in my life	0.67	–
10	I have plans to do enjoyable things for myself…	0.56	–
15	I feel like I am actively involved in life	0.69	–
Emotional distress			
4	I often worry about my health	0.58	0.56
7	My health problems make me very…	0.71	0.71
12	I often feel angry when I think about my health	0.77	0.77
14	I feel hopeless because of my health…	0.68	0.68
18	I get upset when I think about my health	0.79	0.78
21	If I think about my health, I get depressed	0.87	0.87
Constructive attitudes and approaches			
27	I try not to let my health problems stop me from …	0.72	–
34	My health problems do not ruin my life	0.78	–
36	I feel I have a very good life even when I have health …	0.72	–
39	I do not let my health problems control my life	0.80	–
40	If others can cope with problems like mine, I can too	0.73	–
Self-monitoring and insight			
3	As well as seeing my doctor, I regularly monitor changes…	0.40	0.46
6	I know what things can trigger my health problems…	0.54	0.52
11	I have a very good understanding of when and why…	0.45	0.45
16	When I have health problems I have a clear understanding…	0.74	0.72
17	I carefully watch my health and do what is necessary to keep…	0.49	0.54
20	With my health in mind, I have realistic expectations…	0.69	0.69
Skill and technique acquisition			
23	I have effective ways to prevent my symptoms…	0.55	–
25	I have a very good idea of how to manage my…	0.86	–
26	When I have symptoms, I have the skills that help me cope	0.87	–
30	I am very good at using aids and devices to…	0.45	–
Social integration and support			
22	If I need help, I have plenty of people I can rely on…	0.78	0.73
28	I have enough friends who help me cope with my health…	0.80	0.76
31	When I feel ill, my family and carers really understand…	0.68	0.71
35	Overall, I feel well looked after by friends and family	0.81	0.83
37	I get enough chances to talk about my health…	0.81	0.82
Health service navigation			
24	I have very positive relationships with my healthcare…	0.71	0.69
29	I communicate very confidently with my doctor about…	0.72	0.77
32	I confidently give healthcare professionals the information…	0.63	0.68
33	I get my needs met from available healthcare resources…	0.55	0.55

Table 3 continued

Item number	Item content	Factor loadings	
		Original model	Modified model
38	I work in a team with my doctors and other healthcare…	0.85	0.82

[a] Modified models with correlated error variances (see text)

heiQ™ was found to be well replicated in German language; the psychometric properties (reliability, factorial validity, and concurrent validity) showed good fit after only minor adjustment. The German heiQ™ is therefore likely to be a useful measure of proximal outcomes of self-management and health education programs in German-speaking countries.

Overall, the translated heiQ™ was found to have good factorial validity. While three of the eight scales could be accepted immediately, the remaining five scales needed minor adjustments (freeing error covariances of distinct items) to achieve good fit indices. Across the entire questionnaire, only one item was considered problematic. In *Emotional distress* the fit indices were good after freeing the error covariance of two items; however, the deletion of item 18 ("Ich bin sehr beunruhigt, wenn ich über meine Gesundheit nachdenke") improved the model fit substantially. It may be possible that the core meaning of the original item ("upset") was not fully captured by our translation ("sehr beunruhigt"). Nonetheless, removing this item may affect the content validity of the scale; thus, the item was retained. Moreover, the reliability of the scale did not substantially improve when the item was removed (see Table 2). Further studies with different translations of the item may clarify this issue.

Although the CFI for the eight-factor model is somewhat lower than our recommended cutoff value, the fit indices for this model are still within the acceptable range for multidimensional questionnaires, particularly when interpreted in the context of the otherwise acceptable fit statistics [58]. In spite of this, the high correlation between *Skill and technique acquisition* and *Self-monitoring and insight* on the one hand and *Active engagement in life* and *Constructive attitudes and approaches* on the other hand might be problematic. The question arises whether these scales indeed measure conceptually and empirically

Table 4 Correlations between heiQ™ factors (italicized values) and between heiQ™ scales (non-italicized values)

	Health-directed activities	Positive and active engagement in life	Emotional distress[a]	Self-monitoring and insight	Constructive attitudes and approaches	Skill and technique acquisition	Social integration and support	Health service navigation
Health-directed activities	–	*0.59*	*−0.17*	*0.48*	*0.32*	*0.37*	*0.33*	*0.28*
Positive and active engagement in life	0.46	–	*−0.59*	*0.65*	*0.85*	*0.62*	*0.58*	*0.50*
Emotional distress[a]	−0.12	−0.44	–	*−0.35*	*−0.64*	*−0.41*	*−0.38*	*−0.28*
Self-monitoring and insight	0.41	0.47	−0.21	–	*0.59*	*0.95*	*0.52*	*0.70*
Constructive attitudes and approaches	0.27	0.66	−0.55	0.45	–	*0.61*	*0.67*	*0.58*
Skill and technique acquisition	0.31	0.50	−0.35	0.71	0.54	–	*0.51*	*0.63*
Social integration and support	0.27	0.45	−0.33	0.40	0.57	0.49	–	*0.61*
Health service navigation	0.23	0.40	−0.25	0.57	0.50	0.59	0.52	–

[a] In all scales, high values mean positive health-related outcome (e.g., more engagement in life), except for *Emotional Distress*, a reversed scale, where higher values mean greater emotional distress)

Table 5 Correlations between heiQ™ scales and comparator scales

	Health-directed activities	Positive and active engagement in life	Emotional distress[a]	Self-monitoring and insight	Constructive attitudes and approaches	Skill and technique acquisition	Social integration and support	Health service navigation
SF-36 physical functioning	0.10	0.26	−0.42	0.12	0.30	0.18	0.14	0.09
SF-36 role physical	−0.01	0.19	−0.34	0.09	0.25	0.15	0.09	0.08
SF-36 pain	0.09	0.23	−0.38	0.12	0.28	0.23	0.20	0.20
SF-36 general health	0.19	0.41	−0.56	0.16	0.45	0.29	0.30	0.20
SF-36 vitality	0.23	0.46	−0.47	0.23	0.45	0.34	0.33	0.24
SF-36 social functioning	0.15	0.40	−0.45	0.25	0.50	0.33	0.41	0.30
SF-36 role emotional	0.07	0.30	−0.45	0.14	0.37	0.25	0.22	0.18
SF-36 mental health	0.18	0.50	−0.62	0.26	0.57	0.37	0.42	0.32
IRES-24 subjective health	0.21	0.48	−0.60	0.24	0.51	0.37	0.38	0.29
IRES-24 physical health	0.11	0.26	−0.36	0.12	0.26	0.18	0.23	0.14
IRES-24 physical functioning	0.06	0.26	−0.38	0.10	0.29	0.19	0.12	0.11
IRES-24 pain	0.10	0.26	−0.39	0.13	0.31	0.24	0.24	0.21
IRES-24 rehabilitation status	0.15	0.39	−0.54	0.18	0.42	0.30	0.30	0.24
IPQ-R identity	−0.10	−0.25	−0.47	−0.16	−0.33	−0.27	−0.20	−0.15
IPQ-R timeline	−0.06	−0.12	0.26	0.02	−0.13	−0.07	−0.14	−0.07
IPQ-R consequence	−0.13	−0.37	−0.59	−0.16	−0.45	−0.31	−0.35	−0.23
IPQ-R personal control	−0.07	0.11	0.24	0.01	0.13	0.02	−0.01	0.01
IPQ-R coherence	−0.03	0.18	0.36	0.22	0.21	0.30	0.09	0.18
IPQ-R cycle	−0.04	−0.11	−0.29	−0.06	−0.14	−0.16	−0.11	−0.12
IPQ-R emotional representation	−0.14	−0.39	−0.73	−0.27	−0.48	−0.37	−0.27	−0.25
PHQ-9	−0.21	−0.48	0.58	−0.30	−0.54	−0.40	−0.43	−0.32
GAD-7	−0.15	−0.40	0.55	−0.23	−0.47	−0.34	−0.38	−0.32

[a] In all heiQ™ scales, high values mean positive health-related outcome (e.g., more engagement in life), except for *Emotional Distress*, a reversed scale, where higher values mean greater emotional distress)

different constructs. Assessments of the one- and two-factor models indicate that *Active engagement in life* and *Constructive attitudes and approaches* should be modeled as two highly correlated factors. In contrast, the one-factor and two-factor models for *Skill and technique acquisition* and *Self-monitoring and insight* showed very similar fit. However, the conceptual difference between the two scales is very clear: Patients may have skills to cope with symptoms of their illness (*skills and techniques*), but at the same time, they may have little understanding of the underlying mechanisms (*insight*). Therefore, the two-factor model has been chosen for now. More studies are needed to clarify the relationship between these two scales across settings.

Table 6 Mean differences in heiQ™ scales between persons without depression, with other depression, and with major depression (according to PHQ-9)

	PHQ depression syndrome						Effect size[i]	
	No depression (1) ($n = 776$)		Other depression (2) ($n = 167$)		Major depression (3) ($n = 259$)		(1) versus (2)	(1) versus (3)
	M	SD	*M*	SD	*M*	SD	*d*	*d*
HA[a]	2.92	0.75	2.74	0.76	2.60	0.75	0.24	0.42
AE[b]	3.14	0.48	2.85	0.51	2.59	0.56	0.59	1.06
ED[c]	2.17	0.64	2.60	0.66	3.00	0.63	−0.67	−1.31
SI[d]	3.20	0.45	3.07	0.50	2.89	0.51	0.27	0.65
CA[e]	3.40	0.51	3.13	0.51	2.74	0.56	0.53	1.23
ST[f]	3.01	0.54	2.80	0.57	2.56	0.57	0.38	0.81
SI[g]	3.14	0.62	2.90	0.68	2.55	0.72	0.37	0.88
HN[h]	3.31	0.50	3.21	0.51	2.98	0.58	0.20	0.61

[a] Health-directed activities; [b] positive and active engagement in life; [c] emotional distress; [d] self-monitoring and insight; [e] constructive attitudes and approaches; [f] skill and technique acquisition; [g] social integration and support; [h] health service navigation; [i] all effects $p < 0.05$

Although all tested models show good model fit, the values are somewhat lower than in the validation of the original heiQ™ [22]. Several reasons may explain this discrepancy. First, this may be due to the original 42 items being selected from a large pool of items to generate the best possible model, whereas only the 40 translated items were tested in this study. As there are different possibilities to translate an item, other translation options may have led to better fit values. Second, the sample in this study differed from that of Osborne et al. [22]. For example, they did not include cancer patients or patients suffering from inflammatory bowel disease. Finally, the German translation was based on a heiQ™ version with four-point Likert scales and 40 items, while Osborne and colleagues used the six-point Likert scales and 42 items.

In general, reliability estimates of the heiQ™ scales showed acceptable to good values (0.71–0.87). As expected, retest reliability estimates were found to be slightly lower (0.60–0.83) than estimates in CRC, but they are still acceptable.

Most of our hypotheses concerning concurrent validity were confirmed. With one exception, correlation coefficients were lower than $r = 0.6$, indicating that the heiQ™ scales capture other constructs than the comparator scales. This finding confirms that the heiQ™ fills a gap in the measurement of outcomes of patient education and self-management programs.

All heiQ™ scales showed at least low to moderate correlations with measures of subjective health; correlations are slightly higher with mental health than with physical health scales. Furthermore, all heiQ™ scales showed at least low to moderate correlations with depression and anxiety. From all heiQ™ scales, *Emotional distress*, *Constructive attitudes and approaches,* and *Active engagement in life* showed the highest correlations with measures of mental health, depression, and anxiety. This result indicates that these constructs capture elements of a global mental health construct.

The very high correlation (the only one above 0.6) with the IPQ-R scale *Emotional representation* indicates a good convergent validity of heiQ™ scale *Emotional distress*. Both scales capture emotional states with clear attribution to the illness of the patient [22, 45]. The moderate correlations between the heiQ™ scales and IPQ-R scales *consequences* and *coherence* were also expected. For example, patients who feel as though their illness "doesn't make sense" or is "a mystery" (IPQ-R scale *Coherence*) understandably also have fewer skills to cope with the symptoms of the illness (heiQ™ scale *Skill and technique acquisition*). The surprisingly very low correlations between the heiQ™ scales and IPQ-R scale *Personal control* may be due to unclear psychometric properties of this particular IPQ-R scale. For example, Glattacker et al. [44] report low correlations between *Personal control* and other comparator scales (e.g., self-efficacy expectations).

Our findings have shown that the heiQ™ scales can differentiate between patients with and without major depression. Patients with high levels of distress tend to have low values on the heiQ™ constructs. Patients who have little confidence (constructive attitudes) and few self-management competencies are conceivably more likely to become depressed than other patients. An increase in self-management competencies should therefore reduce depression. Conversely, depressed patients may appraise their competencies as lower than patients without depression.

Limitations

Our study has several limitations. Although our sample represents several chronic conditions, many groups are

absent. For example, only few patients suffered from heart diseases. Some important chronic conditions, such as diabetes mellitus and common tumors (e.g., breast, prostate, lung, or skin cancer), are not represented in the sample. Further studies may appraise the generalizability of the results for patients with other chronic conditions.

Construct validity of some heiQ™ scales (e.g., *Emotional distress*) was confirmed by comparisons with related comparator scales (e.g., IPQ-R scale *Emotional representation*). However, construct validity of some other heiQ™ scales (e.g., *Health service navigation*) was less well examined since no comparator scales exist. To obtain additional information on concurrent validity, further studies should use validation scales that encompass related constructs such as doctor–patient relationship [59, 60] or patient competence [61]. Furthermore, future studies should focus on the responsiveness of the scales in groups of individuals participating in interventions that have a specific curriculum designed to improve a range of target outcomes. A more complete understanding of the construct validity of the heiQ™ will evolve through longitudinal studies where sensitivity-to-change or predictive validity is examined.

Conclusions

Overall, the German heiQ™ is well understood by patients suffering from different types of chronic conditions; it assesses relevant outcomes of self-management programs in a reliable manner. The constructs measured by the heiQ™ scales capture different aspects than other used outcome measures and can be assigned to the defined goals of self-management programs, in particular, empowerment (e.g., *Health-directed behavior*, *Health service navigation*), self-management (e.g., *Skill and technique acquisition*), and acceptance of the chronic illness (e.g., *Constructive attitudes and approaches*). The heiQ™ constructs may serve as proximal goals of self-management programs to advance outcome assessment in this field. Further studies involving the heiQ™ and its practical application are warranted.

Acknowledgments The authors wish to thank our cooperation clinics: Rehabilitation Center Bad Eilsen, Hospital Bad Bramstedt, Hospital Bad Oexen, Hospital Bad Reichenhall, Hospital Norderney, Deegenberg Hospital Bad Kissingen, and Rehabilitation Center Bad Mergentheim Hospital Taubertal. We also wish to thank Matthias Lukasczik, Silke Neuderth, Katja Spanier, Christoph Egen, Ann Kathrin Skirde, Melanie Hawkins, and Gerald Elsworth. This project was funded by the German Federal Ministry of Education and Research (Bundesministerium fuer Bildung und Forschung). Professor Osborne was supported in part by an Australian National Health and Medical Research Council Population Health Career Development Award (#400391).

Conflict of interest The authors state that there are no conflicts of interests.

References

1. Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: A review. *Patient Education and Counseling, 48*(2), 177–187.
2. Osborne, R. H., Batterham, R., & Livingston, J. (2011). The evaluation of chronic disease self-management support across settings: The international experience of the health education impact questionnaire quality monitoring system. *The Nursing clinics of North America, 46*(3), 255–270.
3. Faller, H., Reusch, A., Vogel, H., Ehlebracht-König, I., & Petermann, F. (2005). Patientenschulung. *Rehabilitation, 44*(5), 277–286.
4. Hobbs, F. D., Kenkre, J. E., Roalfe, A. K., Davis, R. C., Hare, R., & Davies, M. K. (2002). Impact of heart failure and left ventricular systolic dysfunction on quality of life: A cross-sectional study comparing common chronic cardiac and medical disorders and a representative adult population. *European Heart Journal, 23*(23), 1867–1876.
5. Faller, H., Stork, S., Schuler, M., Schowalter, M., Steinbuechel, T., Ertl, G., et al. (2009). Depression and disease severity as predictors of health-related quality of life in patients with chronic heart failure—A structural equation modeling approach. *Journal of Cardiac Failure, 15*(4), 286–292.
6. Warsi, A., LaValley, M. P., Wang, P. S., Avorn, J., & Solomon, D. H. (2003). Arthritis self-management education programs: A meta-analysis of the effect on pain and disability. *Arthritis and Rheumatism, 48*(8), 2207–2213.
7. Warsi, A., Wang, P. S., LaValley, M. P., Avorn, J., & Solomon, D. H. (2004). Self-management education programs in chronic disease: A systematic review and methodological critique of the literature. *Archives of Internal Medicine, 164*(15), 1641–1649.
8. Chodosh, J., Morton, S. C., Mojica, W., Maglione, M., Suttorp, M. J., Hilton, L., et al. (2005). Meta-analysis: Chronic disease self-management programs for older adults. *Annals of Internal Medicine, 143*(6), 427–438.
9. Cochran, J., & Conn, V. S. (2008). Meta-analysis of quality of life outcomes following diabetes self-management training. *The Diabetes Educator, 34*(5), 815–823.
10. Jovicic, A., Holroyd-Leduc, J. M., & Straus, S. E. (2006). Effects of self-management intervention on health outcomes of patients with heart failure: A systematic review of randomized controlled trials. *BMC Cardiovascular Disorders, 6*, 43.
11. Ditewig, J. B., Blok, H., Havers, J., & van Veenendaal, H. (2010). Effectiveness of self-management interventions on mortality, hospital readmissions, chronic heart failure hospitalization rate and quality of life in patients with chronic heart failure: A systematic review. *Patient Education and Counseling, 78*(3), 297–315.
12. Cockle-Hearne, J., & Faithfull, S. (2010). Self-management for men surviving prostate cancer: A review of behavioural and psychosocial interventions to understand what strategies can work, for whom and in what circumstances. *Psycho-Oncology, 19*, 909–922.
13. Faller, H., Reusch, A., & Meng, K. (2011). DGRW-update: Patientenschulung. *Rehabilitation, 50*(5), 284–291.
14. Deakin, T. A., McShane, C. E., Cade, J. E., & Williams, R. (2005). Group based training for self-management strategies in people with type 2 diabetes mellitus. *Cochrane Database of Systematic Reviews*, (2), CD003417.
15. Foster, G., Taylor, S. J. C., Eldridge, S. E., Ramsay, J., & Griffiths, C. J. (2007). Self-management education programmes by lay leaders for people with chronic conditions. *Cochrane Database of Systematic Reviews*, (4), CD005108.
16. Newman, S., Steed, L., & Mulligan, K. (2004). Self-management interventions for chronic illness. *Lancet, 364*, 1523–1537.
17. Osborne, R. H., Spinks, J. M., & Wicks, I. P. (2004). Patient education and self-management programs in arthritis. *Medical Journal of Australia, 180*(5 Suppl), S23–S26.

18. Du, S., & Yuan, C. (2010). Evaluation of patient self-management outcomes in health care: A systematic review. *International Nursing Review, 57*(2), 159–167.
19. Brenner, M. H., Curbow, B., & Legro, M. W. (1995). The proximal-distal continuum of multiple health outcome measures: The case of cataract surgery. *Medical care, 33*(4 Suppl), AS236–AS244.
20. Meng, K., Seekatz, B., Roband, H., Worringen, U., Vogel, H., & Faller, H. (2011). Intermediate and long-term effects of a standardized back school for inpatient orthopedic rehabilitation on illness knowledge and self-management behaviors: A randomized controlled trial. *The Clinical Journal of Pain, 27*(3), 248–257.
21. Crotty, M., Prendergast, J., Battersby, M. W., Rowett, D., Graves, S. E., Leach, G., et al. (2009). Self-management and peer support among people with arthritis on a hospital joint replacement waiting list: A randomised controlled trial. *Osteoarthritis and Cartilage, 17*(11), 1428–1433.
22. Osborne, R. H., Elsworth, G. R., & Whitfield, K. (2007). The Health Education Impact Questionnaire (heiQ): An outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Education and Counseling, 66*(2), 192–201.
23. Francis, K. L., Matthews, B. L., Van Mechelen, W., Bennell, K. L., & Osborne, R. H. (2009). Effectiveness of a community-based osteoporosis education and self-management course: A wait list controlled trial. *Osteoporosis International: A Journal Established as Result of Cooperation Between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA, 20*(9), 1563–1570.
24. Nolte, S., Elsworth, G. R., Sinclair, A. J., & Osborne, R. H. (2007). The extent and breadth of benefits from participating in chronic disease self-management courses: A national patient-reported outcomes survey. *Patient Education and Counseling, 65*(3), 351–360.
25. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., Erikson, P., & ISPOR Task Force for Translation and Cultural Adaptation. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health, 8*(2), 94–104.
26. Acquadro, C., Conway, K., Hareendran, A., & Aaronson, N. (2008). Literature review of methods to translate health-related quality of life questionnaires for use in multinational clinical trials. *Value in Health, 11*(3), 509–521.
27. Hawkins, M., & Osborne, B. (2007). *Questionnaire translation an cultural adaption procedure. Version 1.0*. Melbourne: University of Melbourne.
28. Willis, G. B. (2004). *Cognitive interviewing. A tool for improving questionnaire design*. Thousand Oaks: Sage.
29. Stewart, A. L., & Napoles-Springer, A. (2000). Health-related quality-of-life assessments in diverse population groups in the United States. *Medical Care, 38*(9 Suppl), II102–II124.
30. Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275–340). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
31. Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
32. Goffin, R. D. (2007). Assessing the adequacy of structural equation models: Golden rules and editorial policies. *Personality and Individual Differences, 42*(5), 831–839.
33. Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55.
34. Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal, 16*(4), 561–582.
35. Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173–184.
36. Raykov, T., & Shrout, P. E. (2002). Reliability of scales with general structure: Point and interval estimation using a structural equation modeling approach. *Structural Equation Modeling, 9*(2), 195–212.
37. Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement, 66*(6), 930–944.
38. Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of strength and conditioning research/National Strength & Conditioning Association, 19*(1), 231–240.
39. Rousson, V., Gasser, T., & Seifert, B. (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Statistics in Medicine, 21*(22), 3431–3446.
40. Bullinger, M. (1995). German translation and psychometric testing of the SF-36 Health Survey: Preliminary results from the IQOLA project. *Social Science and Medicine, 41*(10), 1359–1366.
41. Bullinger, M., & Kirchberger, I. (1998). *Der SF-36-Fragebogen zum Gesundheitszustand. Handanweisung*. Göttingen: Hogrefe.
42. Wirtz, M., Farin, E., Bengel, J., Jaeckel, W. H., Haemmerer, D., & Gerdes, N. (2005). IRES-24 Patientenfragebogen—Entwicklung der Kurzform eines Assessmentinstrumentes in der Rehabilitation mittels der Mixed-Rasch-Analyse. *Diagnostica, 51*(2), 75–87.
43. Buehrlen, B., Gerdes, N., & Jaeckel, W. H. (2005). Entwicklung und psychometrische Testung eines Patientenfragebogens fuer die medizinische Rehabilitation (IRES-3). *Rehabilitation, 44*(2), 63–74.
44. Glattacker, M., Bengel, J., & Jaeckel, W. H. (2009). Die deutschsprachige Version des Illness Perception Questionnaire-Revised. Psychometrische Evaluation an Patienten mit chronisch somatischen Erkrankungen. *Zeitschrift fuer Gesundheitspsychologie, 17*(4), 158–169.
45. Moss-Morris, R., Weinman, J., Petrie, K. J., Horne, R., Cameron, L. D., & Buick, D. (2002). The revised Illness Perception Questionnaire (IPQ-R). *Psychology and health, 17*(1), 1–16.
46. Leventhal, H., Brissette, I., & Leventhal, E. A. (2003). The common-sense model of self-regulation of health and illness. In L. D. Cameron & H. Leventhal (Eds.), *The self-regulation of health and illness behaviour* (pp. 42–65). New York, NY: Routledge.
47. Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613.
48. Spitzer, R. L., Kroenke, K., Williams, J. B., & Lowe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097.
49. Bair, M. J., Wu, J., Damush, T. M., Sutherland, J. M., & Kroenke, K. (2008). Association of depression and anxiety alone and in combination with chronic musculoskeletal pain in primary care patients. *Psychosomatic Medicine, 70*(8), 890–897.
50. Rosemann, T., Laux, G., Szecsenyi, J., Wensing, M., & Grol, R. (2008). Pain and osteoarthritis in primary care: Factors associated with pain perception in a sample of 1, 021 patients. *Pain Medicine, 9*(7), 903–910.
51. Stafford, L., Berk, M., & Jackson, H. J. (2007). Validity of the Hospital Anxiety and Depression Scale and Patient Health

Questionnaire-9 to screen for depression in patients with coronary artery disease. *General Hospital Psychiatry, 29*(5), 417–424.
52. Lamers, F., Jonkers, C. C., Bosma, H., Penninx, B. W., Knottnerus, J. A., & van Eijk, J. T. (2008). Summed score of the Patient Health Questionnaire-9 was a reliable and valid method for depression screening in chronically ill elderly patients. *Journal of Clinical Epidemiology, 61*(7), 679–687.
53. Kroenke, K., Spitzer, R. L., Williams, J. B., & Lowe, B. (2010). The Patient Health Questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345–359.
54. Muthén, L. K., & Muthén, B. (2010). *Mplus user's guide*. Los Angeles: Muthén & Muthén.
55. Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430–457.
56. Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549–576.
57. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2., 2. print. ed.)*. Hillsdale, NJ: Erlbaum.
58. Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated? *Personality and Social Psychology Review, 14*(3), 332–346.
59. Maly, R. C., Frank, J. C., Marshall, G. N., DiMatteo, M. R., & Reuben, D. B. (1998). Perceived efficacy in patient-physician interactions (PEPPI): Validation of an instrument in older persons. *Journal of the American Geriatrics Society, 46*(7), 889–894.
60. Gramm, L., Kosiol, D., & Farin, E. (2010). Evaluation der deutschen Version des Fragebogens "Perceived efficacy in patient-physician interactions" (PEPPI). *DRV-Schriften, 88*, 108–110.
61. Giesler, J. M., & Weis, J. (2008). Developing a self-rating measure of patient competence in the context of oncology: A multi-center study. *Psycho-Oncology, 17*(11), 1089–1099.