

RESEARCH ARTICLE

Open Access



# The German version of the helping alliance questionnaire: psychometric properties in patients with persistent depressive disorder

Hannah Sophie Eich<sup>1\*</sup> , Levente Kriston<sup>2</sup>, Elisabeth Schramm<sup>3</sup> and Josef Bailer<sup>1</sup>

## Abstract

**Background:** The Helping Alliance Questionnaire (HAQ) is a frequently used and highly relevant instrument to assess the therapeutic alliance. The questionnaire was translated into German by Bassler and colleagues (1995) and is available for patients (HAQ-P) and therapists (HAQ-T). Whereas the HAQ-P has been tested regarding psychometrics, the HAQ-T has not. This study aimed at further investigating the psychometric properties of both the HAQ-P and HAQ-T. We hypothesized that the instrument is reliable and shows factorial as well as convergent validity.

**Methods:** Within the framework of a multisite, randomized-controlled clinical trial, comparing the efficacy of Cognitive Behavioral Analyses System of Psychotherapy (CBASP) and supportive psychotherapy (SP) in the treatment of early onset persistently depressed outpatients, the HAQ was filled out by patients ( $n = 255$ ) and therapists ( $n = 81$ ). 66.0% of patients were female; average age at randomization was 44.9 years ( $SD = 11.8$ ). Several confirmatory factor analyses were conducted to test different structures for the HAQ. In addition, correlations between the HAQ and the Inventory of Interpersonal Problems (IIP) were calculated to test for convergent validity.

**Results:** Goodness of fit indices for both a model with two different but strongly related factors named 'relation to the patient/ therapist' and 'satisfaction with therapeutic outcome' and a second model with only one global helping alliance factor were comparable: Chi-Square-based indices rejected the models; RMSEA closely approached the threshold of good model fit, and CFI/ TLI and SRMR suggested that both models sufficiently fit the data. The internal consistency (Cronbach's  $\alpha$ ) calculated for the different scales of the HAQ ranges between questionable to good. Finally, the HAQ scores were significantly related to some of the IIP scores.

**Conclusions:** The German versions of the HAQ offer sufficient reliable instruments for the quick assessment of different facets of the therapeutic alliance. The HAQ global scores can be used as indicators for the global impression of the patients and therapists perception of the quality of the therapeutic alliance. However, the small correlations found between the IIP and the HAQ puts the question of external validity into perspective.

**Trial registration:** This study analysed data from a RCT which was registered on ClinicalTrials.com (NCT00970437). First submitted on September 1, 2009.

**Keywords:** Helping alliance questionnaire (HAQ), Helping alliance, Therapeutic alliance, Psychometrics, Persistent depression

\* Correspondence: [hannah-sophie.eich@zi-mannheim.de](mailto:hannah-sophie.eich@zi-mannheim.de)

<sup>1</sup>Central Institute of Mental Health, Department of Clinical Psychology, Medical Faculty Mannheim/Heidelberg University, J5, 68159 Mannheim, Heidelberg, Germany

Full list of author information is available at the end of the article



## Background

The relationship between patient and therapist is one important factor in psychotherapy which predicts therapy outcome [1, 2]. One perspective on this relationship is the concept of alliance which was originally defined by Bordin [3] and “describes the degree to which the therapy dyad is engaged in collaborative, purposive work” [4]. Today, it is the most studied process variable in psychotherapy research [5]. Three internationally often used instruments to measure alliance are the Helping Alliance Questionnaire [6], the Working Alliance Inventory [7] and the California Psychotherapy Alliance Scales [8, 9]. The Working Alliance Inventory [7] is directly derived from Bordin’s theory of alliance [4]. It measures the agreement of patient and therapist on goals for and tasks in therapy as well as the affective bond between patient and therapist [7]. The California Psychotherapy Alliance Scales, which incorporates several perspectives on alliance, assesses the (i) patient’s commitment to therapy, (ii) the working capacity of the patient, (iii) the therapist’s understanding and involvement and (iv) the agreement of patient and therapist on goals and tasks [10]. The Helping Alliance Questionnaires (HAQ), of which one was investigated in this study, were developed by Alexander and Luborsky [11]. The first version of the HAQ was designed so that it encompasses two dimensions: HA1, i.e., the patient perceiving the therapist as helpful and supportive and HA2, i.e., working together towards common goals. Despite the two dimensions, authors themselves worked with the sum score of all items [12]. Later, Luborsky and colleagues developed a revised version of the HAQ, the HAQ-II, in which the authors removed 6 items on early symptomatic improvement and added 14 items on the collaboration between patient and therapist, on how the patient perceived the therapist and on how the patients perceives the therapist’s feeling toward him or her [13]. The revised version however lacks sound psychometric testing and application in research [14]. Hence, while there is the HAQ-II (for patients, therapists and observers), the HAQ is still widely used and recommended for research [15, 16].

The factor structure and psychometric properties of the HAQ have been investigated in six studies, which included in- and outpatient samples with heterogeneous diagnoses and receiving different forms of psychotherapy [9, 14, 17–20]. The studies generally confirmed the HAQ’s quality and its two factors: one related to the relationship and the other to outcome. Yet, the assignment of items to factors as well as the labelling of factors differed between studies. The authors attributed these discrepancies to differing statistical approaches (e.g., relying solely on exploratory factor analysis or allowing for correlated errors in confirmatory factor analysis), linguistic, cultural and scaling influences on responses, as well as

differences in study setting such as study sample and treatment [9, 14, 17–20].

## German version of the HAQ

### Factor structure

Only the first version of the HAQ was translated into German [18]. Like Alexander and Luborsky [11], the German authors found two factors [17, 18]. Yet, the assignment of items to factors and the number of items per factor differed to the originally proposed two dimensional structure by Alexander and Luborsky [11, 15, 16].

Two studies explicitly investigated the factor structure of the German version of the HAQ which is available in two versions – one for the patient (HAQ-P) and one for the therapist (HAQ-T). The earlier study [18] tested the HAQ-P in a sample of 239 psychodynamically treated inpatients with diverse diagnoses and found two factors, which they called ‘satisfaction with therapeutic outcome’ (items 2, 3, 4, 5, 11) and ‘relation to the therapist’ (items 1, 6, 7, 8, 9, 10). Item 2 and item 3 did not load clearly on either factor. Based on the items’ semantics, the authors assigned these items to the factor labelled ‘satisfaction with therapeutic outcome’. This assignment was later confirmed by test theoretical examination. Cronbach’s  $\alpha$  of the global alliance scale was 0.89; the internal reliability of the subscales was similarly high (‘relation to the therapist’: Cronbach’s  $\alpha = 0.89$ , ‘satisfaction with therapeutic outcome’: Cronbach’s  $\alpha = 0.84$ ). The intercorrelation between the factors was  $r = 0.43$ . Recently, a study by Nübling et al. [14] generally supported the two-factorial structure in a combined sample of three studies with in total 4626 in- and outpatients. Yet, items 1, 2 and 3 loaded inconsistently on the factors. Moreover, the fit indices of the confirmatory factor analyses suggested that the two items should be removed from the questionnaire. The authors however retained the two-dimensional structure of the HAQ-P including all 11 items for content-related reasons and because it is commonly used. The factors correlated between  $r = 0.45$  and  $r = 0.76$ .

Hence, considering the equivocal findings with regards to the factor structure of the HAQ, the call for a sound psychometric foundation of the widely used HAQ-P [14] and its appropriateness to measure the helping alliance [21] remains open to discussion. Furthermore, psychometric examination of the German HAQ-T is entirely lacking. Therefore, this study aimed at adding to the already existing literature by testing the reliability and the factorial structure, of the German version of the HAQ-P and HAQ-T.

### Validity

The German HAQ-P was found to have satisfactory convergent and discriminant validity as assessed by a number of variables which directly or indirectly measure

therapy outcome and motivation for therapy [14, 17, 18]. We sought to expand the research on the validity of the HAQ by correlating the HAQ-P and HAQ-T with patients' pre-treatment interpersonal problems. Patients with friendly-submissive behaviours facilitate a positive alliance with the therapist, whereas patients with hostile-dominant behaviours negatively impact alliance [22]. Also, McCullough [23] postulates that the hostile and hostile-submissive behaviours of persistently depressed patients impedes the interactions between patient and therapist. Hence, alliance scores should be negatively related to patients being overly hostile-dominant, hostile and hostile-submissive. Patients being friendly-submissive should be positively related to alliances scores.

### Hypotheses

We hypothesised that the two theoretically proposed factors, i.e., 'satisfaction with therapeutic outcome' and 'relation to the patient/ therapist', would be confirmed and that the resulting subscales would be reliable and convergently valid.

### Methods

The hypotheses were tested with data from a randomized controlled trial (RCT) comparing the efficacy of the Cognitive Behavioural Analysis System of Psychotherapy (CBASP) to an active control group, i.e. supportive psychotherapy (SP). The RCT was registered on [ClinicalTrials.com](https://www.clinicaltrials.com) (NCT00970437). This prospective and observer-blind study was conducted at eight university centres throughout Germany. For details on the procedures, methodology, and outcome of the RCT see [24, 25].

### Participants

Patients between the age of 18 and 65 years, who had been diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric [26]) with an early onset (before the age of 21) diagnosis of chronic Major Depressive Disorder (MDD), current MDD superimposed on a pre-existing dysthymic disorder ("double depression") or recurrent MDD with incomplete remission between episodes and scoring a minimum of 20 points on the 24-item Hamilton Rating Scale for Depression (HRSD) [27] were eligible to participate. Patients on antidepressant medication had the opportunity to discontinue it (at least two weeks of washout) before entering the trial. Unless any of the following exclusion criteria were met, patients were invited to take part in the study. Exclusion criteria were (i) acute suicidality, (ii) a history of psychotic symptoms, bipolar disorder, or organic brain disorder, (iii) a comorbid primary diagnosis of another axis I disorder, substance use disorder, (iv) anti-social, schizotypal, or borderline personality disorder, (v) severe cognitive impairment, (vi) non-response to CBASP

and/or (SP) in an earlier trial, (vii) ongoing psycho-/pharmacotherapy, and (viii) a serious medical condition [25]. This sampling procedure resulted in 268 participants of whom 66% were female and who were on average 44.91 ( $SD = 11.82$ ) years old (Table 1).

### Treatments

Psychotherapy ran in parallel in both conditions (CBASP and SP) and included an acute therapy phase (20 weeks, 24 individual sessions) followed by eight continuation sessions over the next 28 weeks. CBASP is a highly structured intervention in which patients learn to recognize the effects of their behaviours on others, to actively deal with interpersonal problems and to strengthen self-efficacy by reaching their desired outcomes with other people. SP is a supportive, non-specific, client-centred approach to psychotherapy including elements of psychoeducation and facilitation of affect [28].

### Therapists

CBASP and SP were conducted by two groups of psychotherapists ( $n = 81$ ), all of whom had either completed a three-year psychotherapy training or were in an advanced stage of training. In addition, therapists had been trained in a two-day workshop and had at least one practice day in either CBASP or SP. Before therapists began working with patients, they had to meet the criteria for mastery in CBASP or SP. During the therapy, all sessions were videotaped and supervision took place regularly.

### Measures

#### *Demographic variables and early maltreatment*

During the initial screening, sociodemographic variables, such as sex, age, nationality, marital status, education, occupation, and employment were recorded. Additionally, the Childhood Trauma Questionnaire (CTQ; [29]) was used to assess early traumatization in terms of emotional and physical abuse or neglect and sexual violence.

#### *Diagnoses and depressive symptomatology*

Before therapy commenced, clinical diagnoses had been derived from the Structured Clinical Interviews for DSM-IV (SCID I and II; [30, 31]) and severity of depressive symptoms had been quantified by the 24-item version of the HRSD.

#### *Therapeutic alliance*

Alliance was assessed at the beginning of therapy, i.e. after session 1. If the alliance questionnaire was not distributed and/ or was not returned after session 1, alliance was assessed after session 2 or 3. Both patient and therapist filled out the HAQ-P and HAQ-T [18], respectively. The 11 items of these self-report instruments

**Table 1** Patient characteristics at baseline

Variable	Participants (n = 268)
Age, <i>M</i> ( <i>SD</i> )	44.91 (11.82)
Gender, <i>n</i> (%)	
Male	91 (34.0)
Female	177 (66.0)
Age at onset, <i>M</i> ( <i>SD</i> )	13.00 (4.41)
Marital status, <i>n</i> (%)	
Married; cohabiting	106 (39.6)
Single	117 (43.7)
Divorced, widowed	45 (16.8)
Educational level, <i>n</i> (%)	
≤ 11 years	96 (35.8)
≥ 12 years	172 (64.2)
Diagnosis <sup>a</sup> , <i>n</i> (%)	
Double Depression	119 (45.8)
Chronic Major Depression	82 (31.5)
Recurrent Major Depression without complete remission between episodes	59 (22.7)
Early trauma <sup>a, b</sup> , <i>n</i> (%)	194 (74.6)
CTQ global sum score <sup>a</sup> , <i>M</i> ( <i>SD</i> )	52.82 (16.03)
Emotional abuse <sup>a</sup> , <i>M</i> ( <i>SD</i> )	13.82 (5.57)
Physical abuse <sup>a</sup> , <i>M</i> ( <i>SD</i> )	7.83 (4.21)
Sexual abuse <sup>d</sup> , <i>M</i> ( <i>SD</i> )	6.52 (3.03)
Emotional neglect <sup>e</sup> , <i>M</i> ( <i>SD</i> )	16.18 (5.02)
Physical neglect <sup>a</sup> , <i>M</i> ( <i>SD</i> )	8.43 (3.17)
HRSD score <sup>c</sup> , <i>M</i> ( <i>SD</i> )	27.07 (5.61)

Notes. <sup>a</sup> *n* = 260 (different to Schramm et al., 2017, we included belatedly collected (session 3) CTQ data for 4 participants for whom data was missing at baseline); <sup>b</sup> at least moderate to severe in 1 of 5 dimensions assessed with the Childhood Trauma Questionnaire; <sup>c</sup> HRSD Hamilton Rating Scale of Depression; <sup>d</sup> *n* = 258, <sup>e</sup> *n* = 259

take maximally 10 min to fill out. All items were rated on a 6-point Likert Scale from “strongly agree” (3 points) to “strongly disagree” (− 3 points).

### Pretreatment interpersonal problems

We measured interpersonal problems by means of the German version of the Inventory of Interpersonal Problems (IIP-64; [32]). Its 64 items (5-point Likert scale) assess several aspects of social malfunctioning on 8 subscales which are correlated in the form of a circumplex: (i) domineering/controlling, (ii) vindictive/ self-centred, (iii) cold/ distant, (iv) socially inhibited, (v) non-assertive, (vi) overly accommodating, (vii) self-sacrificing, (viii) intrusive/ needy. The subscales 1 through 4 describe problems with being too dominant, hostile, hostile-dominant, and hostile-submissive. The other subscales deal with problems concerning friendly submissiveness or friendly dominance. The

questionnaire has been found to be a reliable and valid research instrument in English and German populations [32].

### Data analysis

To test the factor structure of the German HAQ for patients and therapists, we performed confirmatory factor analyses (CFA) by means of structural equation modelling with diagonally weighted least squares estimation (WLSMV). In the analyses items were modelled as ordinally scaled. Based on the recommendation to consider several tests when evaluating model fit [33], we included the following indices: Normed Chi-Square (Chi-Square Test of Model Fit divided by Degrees of Freedom), Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker Lewis Index (TLI) and the Standardized Root Mean Square Residual (SRMR). In order to interpret the fit indices we relied on the same cut-off values (Table 3) as Nübling et al. did in their study on the HAQ [14]. The models were investigated with regards to internal consistency and external validity. The internal consistency of the HAQ was assessed by use of Cronbach's  $\alpha$ . To externally validate the German HAQ as a measure to assess the therapeutic alliance between patient and therapist, we performed Pearson correlations (two-tailed) between HAQ and IIP (total score and subscale scores). All descriptive statistics, analyses on internal consistency and external validation were performed using IBM SPSS Statistics 24 [34]; lavaan for R [35] was used to perform the CFA.

### Results

The sample comprised of 268 patients (Table 1), 177 (66.0%) of whom were female, with an average age of 44.91 (*SD* = 11.82). Almost half of the sample was suffering from double depression (45.8%). The other patients had either chronic major depression (31.5%) or recurrent major depression without complete remission between episodes (22.7%). Patients reported a mean age of onset of 13 years (*SD* = 4.41). The average HRSD score at baseline was 27.07 (*SD* = 5.61). Most patients were single (43.7%) or married (39.6%), 16.8% were divorced or widowed. About one third (35.8%) of patients had been in formal education for at least 12 years. Over 70% reported early childhood maltreatment. Of the 255 patients, who returned the HAQ-P, 254 also filled out the IIP.

### Confirmatory factor analysis and internal consistency

To confirm the postulated two-dimensional factor structure of the HAQ [14, 17], we assumed a model with two latent factors in confirmatory factor analysis. This model was tested for both the patient and the therapist versions of the questionnaire. Structural equation modelling showed that 6 relationship items and 5 outcome items correlated significantly with the latent factors (Table 2)

**Table 2** Standardized factor loadings of items on factors for a two- and one-factorial model

	2 factors		1 factor
	<i>Relation to the patient/ therapist</i>	<i>Satisfaction with therapeutic outcome</i>	<i>Helping alliance</i>
HAQ-P <sup>a</sup>			
1: I believe that my therapist is helping me.	.91 <sup>c</sup>		.88 <sup>c</sup>
6: I feel I can depend on my therapist.	.75**		.74**
7: I feel the therapist understand me.	.78**		.77**
8: I feel the therapist wants me to achieve my goals.	.84**		.83**
9: I feel I am working together with the therapist in a joint effort.	.88**		.87**
10: I believe we have similar ideas about the nature of my problems.	.73**		.72**
2: I believe that the treatment is helping me.		.95 <sup>c</sup>	.85**
3: I have obtained some new understanding.		.69**	.64**
4: I have been feeling better recently.		.65**	.62**
5: I can already see that I will eventually work out the problems I came to treatment for.		.76**	.70**
11: I feel now that I can understand myself and deal with myself on my own.		.22**	.21**
Correlation between factors	.83**		
<i>M (SD)</i>	1.56 (0.84)	-0.06 (1.15)	0.82 (0.88)
Cronbach's $\alpha$	.89	.75	.87
HAQ-T <sup>b</sup>			
1: I believe that I am helping my patient.	.77 <sup>c</sup>		.76 <sup>c</sup>
6: I feel that my patient relies on me.	.66**		.66**
7: I feel that my patient feels understood.	.87**		.86**
8: I feel my patient believes I am committed to the attainment of his/ her goals.	.91**		.90**
9: I feel that my patient is working together with me in a joint effort.	.79**		.78**
10: I believe my patient and I have similar ideas about the nature of his/ her problems.	.69**		.69**
2: I believe that the treatment is helping my patient.		.77 <sup>c</sup>	.71**
3: I believe that my patient has obtained some new understanding.		.71**	.66**
4: I believe that my patient has recently been feeling better.		.32**	.30**
5: I believe my patient will eventually work out the problems he/she came to treatment with.		.57**	.53**
11: I feel now that my patient can understand him/ herself and can deal with him/ herself on his/ her own.		.28**	.26**
Correlation between factors	.88**		
<i>M (SD)</i>	0.91 (0.80)	-0.32 (0.83)	0.35 (0.72)
Cronbach's $\alpha$	.85	.63	.84

Note. <sup>a</sup>  $n = 255$ ; <sup>b</sup>  $n = 260$ ; \*\*  $p < .001$ ; \*  $p < .05$ ; <sup>c</sup> reference item in the model; HAQ-P Helping Alliance Questionnaire for patients, HAQ-T Helping Alliance Questionnaire for Therapists

and that the two factors strongly correlated (HAQ-P:  $r = .83$ ; HAQ-T:  $r = .88$ ).

While the goodness of fit indices were mostly satisfactory for the HAQ-T ( $\chi^2 = 153.98$ ,  $df = 43$ ,  $p < .001$ ,  $\chi^2/df = 3.58$ , RMSEA = .10, CFI = .98; TLI = .97, SRMR = .08), the indices were inconclusive for the HAQ-P ( $\chi^2 = 213.36$ ,  $df = 43$ ,  $p < .001$ ,  $\chi^2/df = 4.96$ , RMSEA = .12, CFI = .98; TLI = .98, SRMR = .07) (Table 3). Cronbach's  $\alpha$  of the

two scales ranged from .89 to .75 on the HAQ-P and from .85 to .63 on the HAQ-T.

Because of the high correlation between the factors and cross-loadings between items and factors, we tested a competing one-factor model (Table 2). In the one-factor model all items load onto one latent global factor. This model had comparable or slightly inferior fit indices than the two-factor model: HAQ-P ( $\chi^2 = 268.42$ ,  $df = 44$ ,  $p < .001$ ,  $\chi^2/df =$

**Table 3** Goodness of fit indices for a two- and one-factorial model

Model	$\chi^2$	df	p	$\chi^2/df$	RMSEA	CFI/ TLI	SRMR
Two-factorial model							
HAQ-P	213.36	43	< .001	4.96	.12	.98/ .98	.07
HAQ-T	153.98	43	< .001	3.58	.10	.98/ .97	.08
One-factorial model							
HAQ-P	268.42	44	< .001	6.10	.14	.98/ .97	.09
HAQ-T	163.95	44	< .001	3.73	.10	.98/ .97	.08
Cut-off			> .05	< 3	≤.10	≥ .95/ ≥ .90	< .11

Note. df degrees of freedom, RMSEA root mean error of approximation, CFI comparative fit index, TLI Tucker Lewis index, SRMR standardized root mean square residual, HAQ-P Helping Alliance Questionnaire for Patients, HAQ-T Helping Alliance Questionnaire for Therapists

6.10, RMSEA = .14, CFI = .98; TLI = .97, SRMR = .09) and HAQ-T ( $\chi^2 = 163.95$ ,  $df = 44$ ,  $p < .001$ ,  $\chi^2/df = 3.73$ , RMSEA = .10, CFI = .98; TLI = .97, SRMR = .08) (Table 3). Cronbach's  $\alpha$  of the global scale was .87 for the HAQ-P and .84 for the HAQ-T.

**External validity**

Table 4 shows the correlations between the HAQ global alliance score and its subscales with the IIP total score and its subscales.

**HAQ-p**

The IIP total score correlated significantly negatively with the HAQ-P global helping alliance score ( $r = -0.14$ ,  $p = .03$ ) and with the subscale 'satisfaction with therapeutic

outcome':  $r = -0.15$ ,  $p = .02$ ). The HAQ-P global score also had a significant negative relationship with the following interpersonal problems: domineering/ controlling ( $r = -0.13$ ,  $p = .04$ ), vindictive/ self-centred ( $r = -0.19$ ,  $p < .01$ ), cold/ distant ( $r = -0.23$ ,  $p < .01$ ), and socially inhibited ( $r = -0.22$ ,  $p < .01$ ). Likewise, the subscale 'relation to the therapist' was negatively correlated with the same octants: domineering/ controlling ( $r = -0.14$ ,  $p = .03$ ), vindictive/ self-centred ( $r = -0.17$ ,  $p = .01$ ), cold/ distant ( $r = -0.21$ ,  $p < .01$ ), and socially inhibited ( $r = -0.16$ ,  $p = .01$ ). The HAQ-P subscale 'satisfaction with therapeutic outcome' was only related to the following IIP subscales: vindictive/ self-centred ( $r = -0.18$ ,  $p < .01$ ), cold/ distant ( $r = -0.21$ ,  $p < .01$ ), and socially inhibited ( $r = -0.23$ ,  $p < .01$ ).

**Table 4** Correlations between the HAQ and the IIP total score and its 8 subscales

	HAQ-P <sup>b</sup>			HAQ-T <sup>c</sup>			
	M (SD)	Relation to the therapist <i>r</i>	Satisfaction with therapeutic outcome <i>r</i>	Helping alliance <i>r</i>	Relation to the patient <i>r</i>	Satisfaction with therapeutic outcome <i>r</i>	Helping alliance <i>r</i>
IIP <sup>a</sup>							
total score	14.92 (3.69)	-0.10	-0.15*	-0.14*	-0.05	-0.05	-0.06
Domineering/controlling	8.31 (5.10)	-0.14*	-0.10	-0.13*	-0.12	-0.12	-0.13*
Vindictive/self-centred	11.02 (5.19)	-0.17**	-0.18**	-0.19**	-0.11	-0.11	-0.12
Cold/distant	14.79 (6.04)	-0.21*	-0.21**	-0.23**	-0.08	-0.08	-0.09
Socially inhibited	17.93 (6.71)	-0.16*	-0.23**	-0.22**	-0.11	-0.11	-0.12*
Non-assertive	20.04 (6.68)	-0.03	-0.12	-0.08	0.00	0.01	0.01
Overly accommodating	17.49 (5.83)	0.07	-0.00	0.03	0.10	0.10	0.11
Self-sacrificing	18.71 (5.29)	0.11	-0.01	0.06	0.08	0.08	0.09
Intrusive/ needy	11.10 (5.33)	0.05	-0.09	0.08	-0.01	-0.02	-0.02

Note. <sup>a</sup> n = 254; <sup>b</sup> n = 255; <sup>c</sup> n = 260; \*\* p < .01, \*p < .05; HAQ-P Helping Alliance Questionnaire for Patients, HAQ-T Helping Alliance Questionnaire for Therapists, IIP Inventory of Interpersonal Problems

### HAQ-t

For the HAQ-T, significant correlations were found between the helping alliance global scale and the octants domineering/controlling ( $r = -0.13$ ,  $p = .03$ ) and socially inhibited ( $r = -0.12$ ,  $p = .04$ ).

### Discussion

We examined the psychometric properties of the German HAQ in a large sample of early-onset persistently depressed outpatients and their therapists. By means of structural equation modelling we sought to confirm the elsewhere [14] assumed two-factorial structure of the HAQ. Fit indices were heterogeneous: Chi-Square-based indices rejected the model; RMSEA closely approached the threshold of good model fit, and CFI/TLI and SRMR suggested that this two-factorial model sufficiently fit the data. Because of the ambiguous fit indices, items cross loading on factors and the high correlation between the factors (HAQ-P:  $r = .83$ ; HAQ-T:  $r = .88$ ), we ran additional CFA on a competing one-factor model. The second analysis showed that a one-factorial model had a mostly comparable model fit. Standardized loadings of items on latent factors were generally high. Only item 11 (HAQ-P: "I feel now that I can understand myself and deal with myself on my own.", HAQ-T: "I feel now that my patient can understand him/herself and can deal with him/herself on his/her own.") had consistent loadings of  $< .30$  onto the factor 'satisfaction with therapeutic outcome' in the two-factorial model and on the global factor in the one-factorial model. This may be due to the fact that data was collected at the very beginning of treatment when agreeing to this item is unlikely. Other items can be agreed on earlier in the process of therapy. We expect that at a later point in the treatment, item 11 will load onto the factor labelled 'satisfaction with therapeutic outcome', which it has been assigned to mainly for content-related reasons.

Generally, our findings mirror that of other studies on the psychometric properties of the HAQ: the HAQ, i.e., its global scale and its subscales are internally consistent. Like in other studies, we found very high correlation between the latent factors, which indicates how close, the dimensions 'relation to the therapist/patient' and 'satisfaction with therapeutic outcome' are. While the inter-correlation between the latent factors parallels earlier findings, the magnitude of the herein reported correlation may have been overestimated due to using WLSMV estimation in a small sample [36].

Like in other studies which employed CFA to verify a theoretically proposed factor structure [37], our results failed to unambiguously confirm the HAQ scale(s). Our findings are partially in line with Nübling et al. [14]: they, too, found flaws in the two-factorial structure. In their analyses a two-factorial model without item 2 and

3 proved superior to the proposed model. Because of content related reasons and due to the dispersion of the two factor solution which includes all 11 items, the authors retained the unsatisfactory yet well-known model. No other German study has sought to confirm a one factor structure of the HAQ.

Previous research on the Dutch HAQ [20], too, compared the fit of a two- and a one-factor model of the HAQ and also found items cross-loading (two-factorial model) and correlated measurement errors in the models. Moreover, they had slightly inferior fit indices for the one- than for the two-factorial model, which is why they retained the two-factorial model.

In the literature the HAQ is used uni- and two-dimensionally. Our findings from CFA suggest that the model fit of a two- and one-factorial model is largely comparable. Hence, two models reasonably fit the data. Like in other herein cited studies, model fit was not persistently conclusive, but acceptable with regards to CFI, TLI, and SRMR. The finding that the fit indices are inconclusive is unfortunate but not surprising as they are differently susceptible to aspects of structural equation modelling (e.g., [38]).

It is known that most fit indices can be affected by sample size, but also by estimation method and other aspects [38, 39]. One study employed Maximum Likelihood (ML) and Generalized Least Squares (GLS) estimation and found that in comparison to other indices RMSEA and CFI were minimally influenced by sample size [38]. Regarding the herein employed diagonally weighted least squares estimation (WLSMV) method, there is, to our knowledge, no consensus as to how sample size affects the resulting fit indices. Rather, the WLSMV estimator has not been studied sufficiently yet [40]. Two studies investigating the effect of sample size on WLSMV estimation found that WLSMV performs equally well as ML across different sample sizes [40]. Another study [36] however found that in small samples (i.e.,  $n = 200$ ) models based on WLSMV tend to be over rejected by the Chi-Square Test. Therefore, the common assumption that Chi-Square based fit indices are lenient in small samples [38] may not hold true for our study. Rather, the unsatisfactory Chi-Square results in our study may be due to having relied on WLSMV estimation in a relatively small sample.

In light of the herein used methods and the results, looking at the global alliance scale or the subscales of the HAQ is both equally feasible. Therefore, the researcher or clinician will have to decide what approach better fits the purpose. Working with a two-factorial model holds the advantage of comparability: the subscales are well known and commonly used. Moreover, items on the scale 'satisfaction with therapeutic outcome'

are confounded with therapy outcome [14]. Therefore, assessing alliance on both subscales allows a more fine graded disentanglement of process and outcome variables in psychotherapy research. On the other hand, assessing alliance on one global score is arguably very economic. What is more, relying on one global HAQ score is frequently done (e.g. [41]).

Previously, validity of the German HAQ-P had been established through symptom-, treatment-, and health related instruments [14]. To our knowledge, this is the first study, which used a measure of interpersonal problems to validate both versions of the HAQ in a sample of persistently depressed patients. We found that patients' ratings of the global helping alliance were significantly and negatively related to a sum score of interpersonal problems. This means that the more interpersonal problems a patient had before therapy, the more negative was the evaluation of the helping alliance. Additionally, we found that the more severe problems a patient had with being too dominant, too hostile or too hostile-submissive (i.e., subscales domineering/ controlling, vindictive/ self-centred, cold/ distant, socially inhibited), the more negative the patient perceived the global alliance. These results are in line with research that found that being too hostile [21] or too hostile-dominant [42] negatively impacts the helping alliance at treatment begin. What is more, our finding that patients who are too hostile-submissive evaluate the alliance with the therapist more negatively fits McCullough's assumption that persistently depressed patients exhibit passive, submissive, and hostile behaviours towards the therapist which impedes the interaction with the therapist [23].

Looking at the subscales of the HAQ-P, the results are generally similar: the higher the interpersonal distress and the more problems a patient has with being too hostile-dominant, hostile and hostile-submissive, the more negative the patient evaluated the 'satisfaction with therapeutic outcome'. Therapists' evaluation of the global alliance however, was significantly related only to patients' interpersonal problems with being too dominant or too hostile-submissive.

These findings are principally in line with our expectations. Yet, just part of our hypotheses was confirmed: Only the patient rated global helping alliance correlated consistently with patients' pre-treatment problems with being too dominant, hostile-dominant, hostile, hostile-submissive, and the total IIP score. For the subscales and the therapist-rated alliance the correlational matrix was not so consistent. Moreover, contrary to our hypothesis, we did not find any (positive) correlations between the HAQ and interpersonal problems relating to being overly friendly-submissive.

One must acknowledge that the magnitude of the association between facets of the IIP and the HAQ is only

weak [43]. At the same time, the association is comparable in size to a study by Puschner et al. [21]. Therefore, while the results are relevant to the important question of how interpersonal problems are related to the helping alliance, the IIP may not be the most suitable instrument to establish external validity of the HAQ. This is because one may not expect medium or high correlations. Having said that, correlation coefficients in our study may reflect the homogeneity of our sample (persistent depressive patients), which may have decreased the variance in our data.

Our findings must be viewed considering some limitation the study holds. Firstly, we did not check for a socially desirable response style, i.e. evaluating the alliance more positive than actually perceived, of neither patients nor therapists. In addition to that, not taking into account the hierarchical structure of the data (i.e., several patients were treated by the same therapist) is a drawback: Theoretically, it is possible that the alliance construct is unidimensional on one level, but two-dimensional on the other level. Differences between our results and those in other studies may also be accounted for by the hierarchical structure. Yet, the relatively big sample size and the multi-centre approach of the study support the generalizability of the results. To the best of our knowledge, this was the first study to run confirmative factor analyses on both versions of the German HAQ and to employ a measure of patients' pre-treatment interpersonal problems as a criterion for convergent validity.

## Conclusion

The HAQ is a reliable instrument. CFA did not clearly recommend a two-factorial model over a one-factorial model or vice versa. Thus, our findings suggest using the instrument uni- or twodimensionally, i.e. to work with the global alliance scale or the subscales 'relation to the therapist/ patient' and 'satisfaction with therapeutic outcome'. Patients' and therapists' perception of the alliance is related to pre-treatment interpersonal problems of the patient. These findings are particularly relevant to research on and with the HAQ as it is a standard, perhaps most widely used instrument in current psychotherapy research [1, 16], especially in German speaking countries [14]. Moreover, our results contribute to the ongoing debate on the factor structure of the HAQ (e.g., [14, 20]).

In summary, the HAQ, both for patients and therapists, is an economically applicable research instrument. It assesses the therapeutic alliance via two subscales 'relation to the therapist/ patient' and 'satisfaction with therapeutic outcome' or one global scale. Moreover, its common usage in previous research, its brevity, the option to use it one- or two dimensionally, and its ability to measure changes in alliance over time [17] make the HAQ a recommendable instrument.

### Abbreviations

CBASP: Cognitive Behavioral Analysis System of Psychotherapy; CFI: Comparative Fit Index; CTQ: Childhood Trauma Questionnaire; df: Degrees of Freedom; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, fourth edition; HA1: Helping Alliance dimension 1; HA2: : Helping Alliance dimension 2; HAQ: Helping Alliance Questionnaire; HAQ-II: revised version of the HAQ; HAQ-P: Helping Alliance Questionnaire for patients; HAQ-T: Helping Alliance Questionnaire for therapists; HRSD: Hamilton Rating Scale for Depression; IIP: Inventory of Interpersonal Problems; M: Mean; MDD: Major Depressive Disorder; RCT: Randomized Controlled Trial; RMSEA: Root Mean Square Error of Approximation; SCID: Structured Clinical Interview for DSM-IV; SD: Standard Deviation; SP: Supportive Psychotherapy; SRMR: Standardized Root Mean Square Residual; TL: Tucker Lewis Index

### Acknowledgements

The study that this analysis was based on was designed and headed by Elisabeth Schramm (University of Freiburg), Martin Härter and Levente Kriston (University Medical Center Hamburg-Eppendorf), and Martin Hautzinger (Eberhard Karls University Tübingen). It was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) with the title "A comparison of the Cognitive Behavioural Analysis System of Psychotherapy (CBASP) with the System of Supportive Psychotherapy (SYSP) for Early Onset Chronic Depression" (grant number SCHR443/11-1). In addition to the leading investigators, substantial design, coordinating, or data management tasks were performed by Ingo Zobel, Mathias Berger, Paul Bausch, Thomas Fangmaier (University of Freiburg), and Ramona Meister (University Medical Center Hamburg). Clinical trial sites were Freiburg (headed by Elisabeth Schramm), Tübingen (Martin Hautzinger), Mannheim (Josef Bailer), Marburg (Katrin Wambach), Heidelberg/Stuttgart (Matthias Backenstrass and Knut Schnell), Lübeck (Philipp Klein), Bonn (Dieter Schoepf), and Hamburg (Antje Gumz and Bernd Löwe).

### Funding

This study was part of a larger study which was funded by a grant of the German Research Association (SCHR443/11-1). We acknowledge financial support by Deutsche Forschungsgemeinschaft within the funding programme Open Access Publishing, by the Baden-Württemberg Ministry of Science, Research and the Arts and by Ruprecht-Karls-Universität Heidelberg.

### Availability of data and materials

Methodologically sound proposals to gain access to the datasets used for the presented analyses should be directed to Dr. Schramm (elisabeth.schramm@uniklinik-freiburg.de) and will be assessed by the primary investigators of the study that this analysis was based on. Upon approval, signing a data use agreement will be necessary.

### Authors' contributions

HSE formulated the research question, analyzed the data, interpreted the results, and drafted the manuscript. LK contributed to the analysis of the data and the interpretation of the results. LK and ES revised the manuscript substantially. JB supervised the conception of the study, interpreted the results, and critically revised the manuscript. All authors read and approved the final version of the manuscript.

### Ethics approval and consent to participate

This study is part of a larger RCT, which was approved by Institutional Review Board/ Institutional Ethical Committee (IRB/IEC) of the University of Freiburg. The IRB/ IECs of each participating study center confirmed the approval before the RCT commenced.

Written informed consent was given by all participants before the trial began.

The study was conducted in compliance with the guidelines for Good Clinical Practice and the applicable regulatory requirements.

### Competing interests

Dr. Schramm receives book royalties and honoraria for workshops and presentations on the Cognitive Behavioral Analysis System of Psychotherapy (CBASP). No other disclosures were reported.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Central Institute of Mental Health, Department of Clinical Psychology, Medical Faculty Mannheim/Heidelberg University, J5, 68159 Mannheim, Heidelberg, Germany. <sup>2</sup>Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinstraße 52, 20246 Hamburg-Eppendorf, Hamburg, Germany. <sup>3</sup>Department of Psychiatry, Faculty of Medicine, University of Freiburg, Hauptstraße 5, 79104 Freiburg, Germany.

Received: 11 December 2017 Accepted: 16 April 2018

Published online: 23 April 2018

### References

- Wampold BE, Imel ZE. The great psychotherapy debate: the evidence for what makes psychotherapy work: Routledge; 2015.
- Crits-Christoph P, Gibbons MBC, Mukherjee D. Psychotherapy process-outcome research. In: Oxford LMJ, editor. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 6 edition. England: John Wiley & Sons; 2013. p. 298–340.
- Bordin ES. The generalizability of the psychoanalytic concept of the working alliance. *Psychother Theor Res Prac*. 1979;16:252–60.
- Hatcher RL, Barends AW. How a return to theory could help alliance research. *Psychother Theor Res Prac*. 2006;43:292–9.
- Flückiger C, Horvath AO, Del Re AC, Symonds D, Holzer C. Bedeutung der Arbeitsallianz in der Psychotherapie. *Psychotherapeut*. 2015;60:187–92.
- Luborsky L. A pattern-setting therapeutic alliance study revisited. Abingdon: Taylor & Francis Group. 2000;
- Horvath AO, Greenberg LS. Development and validation of the working alliance inventory. *J Couns Psychol*. 1989;36:223–33.
- Marmar CR, Weiss DS, Gaston L. Toward the validation of the California therapeutic alliance rating system. *Psychol Assess*. 1989;1:46–52.
- Hatcher RL, Barends AW. Patients' view of the alliance in psychotherapy: exploratory factor analysis of three alliance measures. *J Consult Clin Psychol*. 1996;64:1326–36.
- Gaston L. Reliability and criterion-related validity of the California psychotherapy alliance scales—patient version. *Psychol Assess*. 1991;3:68–74.
- Alexander LB, Luborsky L: The Penn Helping Alliance Scales. In *The psychotherapeutic process: A research handbook*. Edited by Greenberg LS, Pincus WM. New York, NY: Guilford Press; 1986: 325–366. *Guilford clinical psychology and psychotherapy series*.
- Luborsky L, McLellan A, Woody GE, O'Brien CP, Auerbach A. Therapist success and its determinants. *Arch Gen Psychiat*. 1985;42:602–11.
- Luborsky L, Barber JP, Siqueland L, Johnson S, Najavits LM, Frank A, Daley D. The revised helping alliance questionnaire (HAQ-II) : psychometric properties. *J Psychother Pract Res*. 1996;5:260–71.
- Nübling R, Kraft M, Henn J, Kriz D, Lutz W, Schmidt J, Wittmann W, Bassler M. Testing the psychometric properties of the helping alliance questionnaire (HAQ) in different health care settings. *Psychother Psychosom Med Psychol*. 2017;
- Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *J Consult Clin Psychol*. 2000;68:438–50.
- Elvins R, Green J. The conceptualization and measurement of therapeutic alliance: an empirical review. *Clin Psychol Rev*. 2008;28:1167–87.
- Bassler M, Nübling R. Helping Alliance Questionnaire. In: *Diagnostische Verfahren in der Psychotherapie [Diagnostic methods in psychotherapy]*. Edited by Brähler E, Schuhmacher J, Srauß B. Göttingen: Hogrefe; 2015. p. 246–9.
- Bassler M, Potratz B, Krauthausen H. Der "helping alliance questionnaire" (HAQ) von Luborsky. *Psychotherapeut*. 1995;40:23–32.
- De Weert-Van Oene GH, Jorg F, de Jong CA. Association between interpersonal behaviour and helping alliance in substance-dependent patients. *Eur Addict Res*. 2006;12:67–73.
- Hendriksen M, Van R, Peen J, Oudejans S, Schoevers R, Dekker J. Psychometric properties of the helping alliance questionnaire-I in psychodynamic psychotherapy for major depression. *Psychother Res*. 2010;20:589–98.
- Puschner B, Bauer S, Horowitz LM, Kordy H. The relationship between interpersonal problems and the helping alliance. *J Clin Psychol*. 2005; 61:415–29.

22. Muran JC, Segal ZV, Samstag LW, Crawford CE. Patient pretreatment interpersonal problems and therapeutic alliance in short-term cognitive therapy. *J Consult Clin Psychol*. 1994;62:185–90.
23. McCullough JP. Treatment for chronic depression : cognitive behavioral analysis system of psychotherapy (CBASP). New York: Guilford; 2000.
24. Schramm E, Kriston L, Zobel I, et al. Effect of disorder-specific vs nonspecific psychotherapy for chronic depression: a randomized clinical trial. *JAMA Psychiatry*. 2017:E1–E10.
25. A Comparison of Cognitive Behavioral Analysis System of Psychotherapy (CBASP) and System of Supportive Psychotherapy (SYSP) for Early Onset Chronic Depression [<https://clinicaltrials.gov/ct2/show/study/NCT00970437?term=schramm&rank=1>].
26. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-IV*. 4th ed., text rev. edn. Washington, DC: American Psychiatric Association; 2000.
27. Hamilton M. Development of a rating scale for primary depressive illness. *British journal of social and clinical psychology*. 1967;6:278–96.
28. Markowitz JC, Manber R, Rosen P. Therapists' responses to training in brief supportive psychotherapy. *Am J Psychother*. 2008;62:67–81.
29. Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, Stokes J, Han-delsman L, Medrano M, Desmond D, Zule W. Development and validation of a brief screening version of the childhood trauma questionnaire. *Child Abuse Negl*. 2003;27:169–90.
30. First M, Gibbon M, Spitzer RL, Williams JBW, Benjamin LS. Structured clinical interview for DSM-IV Axis I personality disorders (SCID-II). Washington, DC: American Psychiatric Press; 1997.
31. First M, Spitzer RL, Gibbon M, Williams JBW: Structured clinical interview for DSM-IV-TR Axis I disorders, research version, clinical trials version (SCID-CT). New York, NY: Biometrics Research, New York State Psychiatric Institute; 2002.
32. Horowitz LM, Alden LE, Kordy H, Strauß B. *Inventar zur Erfassung interpersonaler Probleme: deutsche Version; IIP-D*. Beltz. TEST. 2000;
33. Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociol Methods Res*. 2008;36:462–94.
34. SPSS Inc: IBM SPSS statistics for windows. 24 edition. Armonk, NY: IBM Corporation; 2015.
35. Rosseel Y: lavaan: An R Package for Structural Equation Modeling. In *J Stat Softw*, vol. 48. pp. 1–36; 2012:1–36.
36. Li C-H. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods*. 2016;48:936–49.
37. Muncer S, Campbell A: Confirmatory factor analysis and the factor structure of expagg in context: A reply to forrest et al., 2002. *Aggress Behav* 2004, 30:12 S.
38. Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Struct Equ Modeling*. 1999;6:56–83.
39. Marsh HW, Balla JR, McDonald RP. Goodness-of-fit indexes in confirmatory factor analysis : the effect of sample size. *Psychol Bull*. 1988;103:391–410.
40. Beauducel A, Herzberg PY. On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Struct Equ Modeling*. 2006;13:186–203.
41. Constantino M, Smith-Hansen L. Patient interpersonal factors and the therapeutic alliance in two treatments for bulimia nervosa. *Psychother Res*. 2008;18:683–98.
42. Connolly Gibbons MB, Crits-Christoph P, de la Cruz C, Barber JP, Siqueland L, Gladis M. Pretreatment expectations, interpersonal functioning, and symptoms in the prediction of the therapeutic alliance across supportive psychotherapy and cognitive therapy. *Psychother Res*. 2003;13:59–76.
43. Taylor R. Interpretation of the correlation coefficient: a basic review. *J Diagn Med Sonogr*. 1990;6:35–9.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

