PsychOpen
publishing psychology

Research Reports

Validation of a Turkish Version of the ICD-10 Symptom Rating (ISR)

Jan Ilhan Kizilhan*[a], Antje Roniger[b], Friedrich von Heymann[c], Karin Tritt[cd]

[a] Cooperative State University BW, Villingen-Schwennigen, Germany. [b] University of Freiburg, Freiburg, Germany. [c] Quality Development in Psychotherapy and Psychosomatic Medicine (IQP), Munich, Germany. [d] University of Regensburg and Quality Development in Psychotherapy and Psychosomatic Medicine (IQP), Munich, Germany.
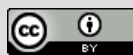
Abstract

Numerous psychiatric and psychosomatic clinics in Turkey and Germany use the Symptom Checklist 90 Revised (SCL-90-R) developed by Derogatis (1977) or the validated Turkish version by Dag (1991) for assessing psychological symptoms. Many patients informed us during numerous studies and visits to these clinics that this test with its 90 questions took too long and that they were unable to sufficiently concentrate on it. In the meantime, the much more economical ICD-10 Symptom Rating (ISR) (Tritt et al., 2008) self-rating questionnaire, comprising 29 questions, has been developed in Germany in 2008. In 2008 and 2009 we therefore decided to translate the ISR into Turkish, to analyse it for its reliability and validity and compare it with the SCL-90-R and the BDI. In an analysis of 277 Turkish subjects – 127 of whom were inpatients, 36 outpatients and 104 clinically unremarkable healthy participants – very good psychometric characteristics were achieved in terms of high internal consistency of individual, additional and overall scales. The results of the factor analysis conducted showed that the ISR Measure has satisfactory construct validity. In a random sample of inpatients, the Cronbach's alpha values ranged from 0.66 (scale: *Compulsive syndrome*) to 0.93 (overall scale). The advantage of this instrument over BDI and SCL-90-R lies in its shorter processing time. The German version of the ISR promises lesser use of time and good empirical quality, which we double-checked with a translated Turkish version tested on persons of Turkish origin in Germany.

Keywords: psychometrics, Turkish version, ISR, validity, reliability

Introduction

Psychometric symptom capture is a widely used method in psychotherapeutic practice (Freyberger & Stieglitz, 2005), quality assurance (Grawe & Braun, 1994; Heymann, Zaudig, & Tritt, 2003) and research (Hill & Lambert, 2000). It enables those symptoms of a disorder that are important for the entire treatment process to be determined from the initial visit, and it assists the process of arriving at a diagnosis and assessing the therapy.

At the same time, regular methods of diagnosing psychiatric disorders using psychometric tests for each individual patient can be too time-consuming and impracticable (Wittchen & Perkonigg, 1997). For this reason, when developing psychometric tests, efforts are made to bear in mind not only validity and reliability issues, but also their application in terms of time and practical implementation.

It is precisely thanks to the fact that it is economical and practical to implement that the SCL-90-R is one of the most frequently used self-assessment instruments for patients with psychological disorders undergoing inpatient or outpatient treatment as well as in research (Hill & Lambert, 2000). A number of different symptoms are polled in a relatively short time using its 90 items. Apart from nine individual scales and a Global Symptom Index (GSI), the SCL-90-R provides a characteristic value for the severity of psychological impairment, which also takes psychological comorbidity into account. The drawback of the SCL-90-R is that its suitability for patients with somatoform disorders, whose symptoms are mainly of a physical nature, is reported to be limited (Tritt, Peseschkian, & Bidmon, 2002). It has been reported that the test's 90 questions are too much to handle, especially for patients with moderate to severe depression and that they feel that the application of the test is stressful (Kizilhan, 2010).

The ICD-10 Symptom Rating (ISR) instrument as an economical and comprehensive symptom survey across the spectrum of disorders is designed to image psychological complaints that the patients themselves are able to assess and to rate their degree of severity.

An attempt is made, using a time- and resource-saving approach with the newly developed ICD-10 Symptom Rating (ISR) instrument (which is available free of charge), to bridge the gap between capturing psychiatric symptoms across the spectrum of disorders, on the one hand, and provide diagnoses that are standardised to a maximum extent on the other (Tritt et al., 2008).

Thus, unlike the SCL-90-R, the ISR is limited to the non-psychotic disorders. The ISR, with its compulsive and eating disorder syndrome scales and screening items relating to further disturbances and impediments, goes beyond such instruments as the SCL-90-R and PHQ (Patient Health Questionnaire), imaging the most common psychosomatic disorders (Gräfe, Zipfel, Herzog, & Löwe, 2004). Despite this, the questionnaire and its 29 items are decidedly shorter than the PHQ with 78 items or the SCL-90-R with 90 items. The objective of the instrument is to assess the severity of psychological disorder at overall and syndrome-specific levels (Tritt et al., 2008).

The instrument, based on the worldwide consensus of the ICD-10, describes which symptoms are relevant to the description and evaluation of psychological disorders and consistently adheres to the diagnostic criteria. The questionnaire comprises five syndrome scales (depression, anxiety, compulsion, somatisation, eating disorders) and an additional scale with a screening function for various other symptoms, which a panel of experts found to be possible to evaluate through self-assessment (Fischer, Tritt, Klapp, & Fliege, 2010). The syndrome scales do not correspond to specific clinical ICD diagnoses, but rate those symptoms that frequently accompany various disorders (e.g., anxiety syndromes in specific phobias, panic disorders or hypochondriacal disorder).

Empirical studies conducted to date in the German-speaking part of the world rate the ISR's quality (validity and reliability) as being in the middle range (Brandt, 2009; Fischer et al., 2010; Schirmer, 2009; Tritt et al., 2008).

The Turkish versions of the questionnaires in use for the general collection of symptoms on patients of Turkish origin in Germany, such as SCL-90-R, PHQ (Patient Health Questionnaire) or GHQ (General Health Questionnaire), are also problematic due to their large number of items and their insufficient coverage of the most frequent disorders of psychosomatic patients (Goldberg, 2008; Kizilhan, 2011).

However, we believe that the present study would help to address the issue of providing empirical evidence for the validity of the ISR. Furthermore, the attempt to test the factor structure of a measure of symptoms in a Turkish-

speaking sample would help to demonstrate the cultural fairness and linguistic independence of the ISR in general and its more economical use of time in particular.

Sometimes, direct translations of psychometric measures – particularly dealing with psychiatric syndromes, which depends on culture and local understandings of illness (Kizilhan, 2012) – into another language may fail to capture the meaning of the construct in the new culture and language.

For instance, Haasen and colleagues (2005) observed that the construct of psychotic symptoms is sufficiently different between German and Turkish cultures and literal translation of the German scale in Turkish may not reflect the way in which Turkish participants regulate and understand psychotic symptoms. Further, evaluation of the psychometric properties of self-report measures of psychiatric syndromes across different cultures becomes important in light of the observed cultural differences in terms of affect, cognition, behaviour and physiological reactions (Kizilhan, 2011). The connection between attitudes on health and illness, on the one hand, and cultural norms, on the other, are particularly strong, which also means that any statement about an illness is culture-specific to a certain degree (Kizilhan, 2012).

For instance, the association of religious and magic notions (such as the existence of spirits, jinn, symbols or rituals) on the one hand and illness on the other is still of great relevance to many traditional societies in Middle East and Africa (Heine & Assion, 2005).

The present study will test the validity of ISR and its usability in the Turkish cultural context in addition to paving the way for future research in the area of transcultural diagnostic and psychometric tests by making available a Turkish version of the ISR.

Method

Sample

The sample of this study comprised a total of 277 participants, 127 of whom were inpatients, 36 outpatients (therapeutic treatment once a week by a therapist) and 104 clinically (healthy) unremarkable subjects.

The number of female subjects taking part in the study was almost twice as high (64.1 percent) as that of the male ones (35.9 percent). This proportion is homogeneous in all three groups ($\chi2$ (2, N = 276) = 0.165; p = 0.92; ns; M.V. (missing value) = 1).

The groups differed with respect to the participants' age (F(2, 272) = 180.12; $p < 0.01$). Inpatient subjects (M = 42.92; SD = 9.39) and outpatient ones (M = 45.21; SD = 7.48) were older than the healthy participants (M = 26.93; SD = 7.11); this was a strong effect (η^2 = 0.57). The difference between the outpatients and inpatients was not significant statistically (p = 0.43).

The variable education was formed by scoring the highest completed level of education stated by the participant (0 = no education; 1 = elementary school; 2 = upper division of elementary school; 3 = junior high school; 4 = senior high school/vocational school; 5 = university of applied sciences; 6 = university). It made no difference whether the education was completed in Turkey or in Germany.

The level of education of the clinically unremarkable subjects was higher than that of those from the other two groups. There was also a statistically significant difference between outpatients and inpatients (F(2, 274) = 194.93;

$p < 0.01$; $N = 277$); this, too, was a strong effect ($\eta^2 = 0.58$). In the inpatient group, men ($M = 2.84$; $SD = 1.35$) had a markedly higher level of education than women ($M = 1.74$; $SD = 1.21$, $T = 4.67$; $p < 0.01$; $N = 126$, M.V. = 1); this was also a strong effect ($d = 0.86$).

Research Instruments

ICD-10-Symptom Rating: A detailed description of the design of the ICD-10 Symptom Rating can be found in Tritt et al. (2008). The licence-free questionnaire is available for download free of charge. The questionnaire lists a total of 29 items, which comprise five syndrome scales and an additional scale with a screening function. A five-level rating scale from 0 (do not agree) to 4 (agree absolutely) is provided as an answer format.

The *Sociodemographic Questionnaire* by Koch (1997) provides important background information about a person (origin, religion, migration status, family, work, financial situation, duration of illness, treatment, etc.).

Beck Depression Inventory (BDI). The severity of depressive symptoms was captured with the Turkish version of the questionnaire (Hisli, 1989) of BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Numerous validation studies have been published, which assume the capture of the construct of depressivity to be highly valid and reliable for patients of Turkish descent (Aktürk, Dagdeviren, Türe, & Tuglu, 2005).

The *Symptom Check List* (*SCL-90-R*; Derogatis, 1977) measures a person's subjectively perceived impairment through physical and psychological symptoms. Dag (1991) analysed and confirmed the reliability and validity of the Turkish version of the SCL-90-R.

Procedure

The Turkish version of the ICD-10 Symptom Rating (ISR) instrument was developed by researchers working in small groups. Initially, a group of four researchers in the field (who spoke both Turkish and German) were requested to translate the ISR into Turkish. This group was briefed about the concept of affect intensity severity and the purpose of the original scale and was requested to pay attention to the grammatical form as well as the psychological content of each item they translated. These four translations were evaluated by the first author and a preliminary Turkish version of the ISR was prepared and accepted by the four researchers. The final Turkish translation of the ISR was then translated back into German by another researcher specialising in this field. The back translation was compared with the original form by the first and second author and the translation was found satisfactory. The final Turkish version of the ISR was administered to a group of 10 members of a Turkish migrant association and a group of 25 Turkish inpatients in a psychosomatic clinic to evaluate their subjective understanding of the item content and identify any ambiguity in instructions or meaning of the items. This pilot study revealed no difficulty in understanding either the instruction or the item content of the ISR. Finally, the ISR was administered to the 277 participants of the present study.

The collection of data was carried out with three groups of subjects of Turkish descent living in Germany:

Group 1: Clinically unremarkable subjects.

Group 2: Patients undergoing outpatient psychotherapeutic/psychiatric treatment.

Group 3: Persons undergoing inpatient psychosomatic treatment.

A total of six clinics and four outpatient psychotherapists, who provide treatment to Turkish-speaking patients in their native tongue, took part in the study. Data on the clinically unremarkable subjects were collected in the period from October 2008 to January 2009. To this end, Turkish associations were contacted in person or in

writing to draw their attention to the investigation. They were then provided with the test questionnaires and asked to complete them. Participants who had already sought psychological help in the past were therefore excluded a priori from this group. The clinical outpatient subjects were informed about the investigation, asked to complete the questionnaire in a room where they would not be disturbed in and to return it when they were finished. Participants in an inpatient setting were selected within the framework of the standardised test diagnostics at the initial visit or admission using the test instruments that had been prepared in advance.

Statistical Investigation

The collected data were processed using the SPSS 15.0 for Windows (2006) and Amos 7.0 software. The concurrent validity of ISR was analysed using Pearson correlations of the ISR scales with the established Turkish translations of the SCL-90-R and BDI instruments, while confirmatory factor analysis was used to check the model fit of the ISR. Comparisons of Fischer's z transformed correlations were made using z tests and weighting of the random sample size.

Results

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis was used to check if the questionnaire's structure corresponded to the theoretical model (see Table 1). The individual items were understood to be manifest variables of five latent correlated syndrome scales.

Table 1

Fit Indices of the Tested Models of the ICD-10 Symptom Rating

Model	NPAR	χ^2	Df	X2/df	RMSEA	NFI	CFI	SRMR
Total sampling (N = 277)	61	287.64	109	2.64	.078	.91	.94	.07
Inpatients (N = 127)	61	223.91	109	2.05	.09	.82	.89	.08
Clinically conspicuous patients (N = 163)	61	225.05	109	2.07	.08	.87	.93	.07
Clinically inconspicuous patients (N = 104)	61	279.69	109	2.57	.12	.72	.80	.12

The testing was carried out on a total random sample of $N = 277$ subjects as a recursive model. Parameters were assessed using the maximum likelihood method. The partial standardized regression weights ranged from 0.47 (item 15) to 0.97 (item 16). The non-standardized regression weights became significant for all items at α level of $p < 0.001$. A test for collinearity of the items showed no item correlation > 0.85.

The variance of the items clarified by the factor ranged from 0.22 (item 15; *Eating habits*) to 0.93. All error variances except item 16, *Thinking about food*, ($p = 0.31$), became significant at α level of $p < 0.001$ (see Figure 1).

Particularly striking was the connection between the latent variables *Compulsion* and *Anxiety* ($r = 0.89$), which in turn correlated strongly with the latent variable *Somatoform disorder* ($r = 0.85$).

Confirmatory factor analysis revealed a significant deviation of the data set from the specified model (χ^2 (109, $N = 277$) = 287.64). Since all items except item 9 (*Compulsion*) and 17 (*Thinking about weight reduction*) clearly violated the multivariate normal distribution ($z > 1.96$) in the Mardia Test, a Bollen-Stine-Bootstrap Correction of the p value was made. This too resulted in a clear rejection of the model ($p = 0.001$).
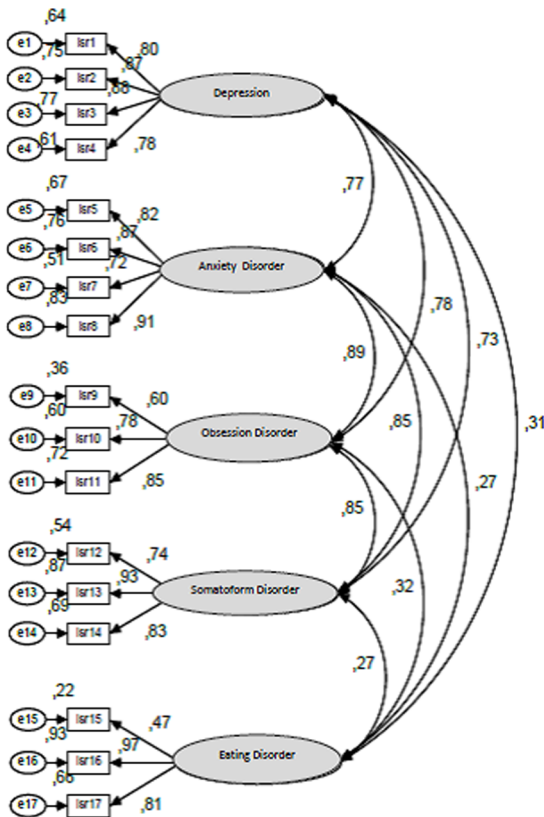
Figure 1. Model of the CFA of the overall sample with factor loading and factor correlations.

Based on the classic, descriptive goodness-of-fit measure (see Table 1), the model fit can be seen as satisfactory.

Gender Effects and Education

The gender effect ($T = 2.70$; $p < 0.01$) was found to be moderate on the *Eating disorder syndrome* scale at the level of the groups of clinically unremarkable patients (see Table 2). Women ($M = 0.66$; $SD = 0.96$; N = 66) had higher values than men ($M = 0.29$; $SD = 0.43$; $N = 38$; $d = 0.50$) and a weaker effect at the inpatient level (women: $M = 0.71$; $SD = 0.47$; men: $M = 0.55$; $SD = 0.47$; $d = 0.47$). However, after a Bonferroni correction of the error, these effects were no longer statistically significant. The same applied to the weak gender effect on the *Anxiety syndrome* scale in the case of inpatients ($d = 0.38$), to a weak gender effect on the *Depressive syndrome* scale ($d = 0.34$) and to the overall score ($d = 0.24$) in the case of clinically unremarkable patients.

Economy and Practicability

In contrast to the ISR questionnaire, the SCL-90-R comprises only 29 items which, on average, can be processed in ten to twelve minutes. The participants needed significantly less time ($F(2,156) = 27,43$; $p < 0.1$) to complete the ISR ($M = 11,40$; $SD = 4.0$) than the BDI ($M = 16,36$; $SD = 6.13$) and the SCL-90-R ($M = 20.11$; $SD = 7.57$). This effect was strong ($\eta^2 = 0.25$). All individual comparisons were significant at α level of $p < 0.01$. Missing variance homogeneity was corrected in accordance with Tamhane-t2.

Table 2

Gender Differences for the ISR Scale Means of Inpatients and Clinically Inconspicuous Test Persons

| | Inpatients | | | | | Clinically unremarkable persons | | | | |
| | Men (n = 44) | | Women (n = 82) | | | Men (n = 38) | | Women (n = 66) | | |
	M	SD	M	SD	Cohen's d	M	SD	M	SD	Cohen's d
Depressive syndrome	2.09	.91	2.18	.91	.10	.75	.81	1.06	1.01	.34
Anxiety syndrome	1.68	.97	2.05	.97	.38	.40	.47	.40	.497	.00
Anankastia syndrome	1.62	.92	1.55	.92	.08	.48	.59	.44	.75	.06
Somatoform syndrome	1.71	.995	1.65	.99	.06	.27	.50	.26	.50	.02
Eating disorder syndrome	.55	.71	.89	.71	.47	.29	.43	.66	.96	.497
Additional scale	1.50	.65	1.60	.65	.15	.44	.51	.51	.41	.15
Overall score	1.52	.67	1.65	.67	.197	.44	.44	.55	.46	.24

Reliability

In inpatients, Item 15 (*Weight control*) had the lowest mean value (0.43). The means of the remaining items ranged from 0.77 (item 21; scale: additional scale) to 2.36 (item 3; scale: *Depressives syndrome*). The distribution of the item peaks was predominantly broad.

The item means of the Turkish random sample (M = 1,62; SD = 0.51) were distinctly higher than those of the German random samples obtained by Fischer, 2009 (M = 1,12; SD = 0.50). This difference was statistically significant (t = 3,75; $p < 0.01$); this was a strong effect (d = 0.97). The differences ranged between 0.02 (item 17) and 1.12 (item 29). Only items 15 and 21 relating to eating habits were below the values of the German patients.

The internal consistency of the scale with Cronbach's alpha of 0.78 was seen as good. All other scales, too, especially the *Depressive syndrome* scale (Cronbach's alpha: 0.90), *Anxiety syndrome* (Cronbach's alpha: 0.89), the additional scale (Cronbach's alpha: 0.91) and the overall scale (Cronbach's alpha: 0.96) proved stable.

Concurrent Validity

Compared with the results on patients undergoing inpatient psychosomatic care obtained by Tritt (Tritt et al., 2008) and other Turkish participants in psychosomatic clinics, a higher correlation was observed on the ISR scales *Anxiety syndrome, Somatoform syndrome* and *Compulsive syndrome* and the corresponding subscales *Anxiety, Somatisation* and *Compulsiveness* of the SCL-90-R. The *Depressivity* scale of SCL-90-R correlated better with the ISR than the BDI overall value did, also at the item level, as can be seen in Table 3.

Discussion

The present study analysed the Turkish translation of the ICD-10 symptom rating questionnaire (ISR) for its validity and reliability.

Using confirmatory factor analysis, the intended dimensional structure of the instrument could not be unambiguously confirmed for the overall sample or at the level of individual groups. However, the various goodness-of-fit measures confirmed a "good" fit of the theoretically specified model based on the overall sample.

The ISR showed good psychometric characteristics, with high internal consistency, weak age and gender effects, and weak to moderate education effects for female participants. The instrument proved objective, reliable and

Table 3

Correlations of the ISR Scales BDI and the SCL-90-R

			German version of ISR		
Scale	Group	Current Study	Tritt et al.[a]	Brandt[b]	Franke[c]
Overall ISR score & SCL-90-R-GSI	total	r = .89**	r = .79**		
ISR: Depressive syndrome & SCL-90-R: depressivity	total	r = .81**	r = .76**	r = .75**	
ISR: Depressive syndrome & BDI: sum value	total	r = .73**		r = .73**	
BDI: Sum value & SCL-90-R: depressivity	total	r = .88**		r = .76**	r = .75**
ISR: Anxiety syndrome & SCL-90-R: phobic anxiety	total	r = .82**	r = .72**		
ISR: Anxiety syndrome & SCL-90-R: Anxiousness	total	r =. 84**	r = .66**		
ISR: Anankastia syndrome & SCL-90-R: anankastia	total	r = .71**	r = .49**		
ISR: Somatoform syndrome & SCL-90-R: somatisation	total	r = .74**	r = .37**		

Note. Correlations of the ISR scales for the complete group (N_{total} = 267) with the sum value of the BDI and the SCL-90-R-GSI in comparison to the results of the German version of the ISR.
[a]Correlation and significance level according to Tritt et al. (2008), N = 22. [b]Correlation and significance level according to Brandt (2009) N = 968, correlations before treatment. [c]Correlation and significance level according to Franke (2001, 2002), N = 5057.
**$p < .01$.

valid on all scales with respect to the criteria set except the *Eating disorder syndrome* scale, which was not examined. Decidedly less time-consuming, the ISR achieved reliability values comparable to those obtained with the more extensive SCL-90-R. The same holds true for the performance of the ISR scale *Depression syndrome* with only four items compared with the BDI's 21 items. In terms of comparability, this speaks in favour of using ISR.

The validity of the results for patients with a Turkish background undergoing inpatient psychosomatic treatment was considered good. A generalization of the results for the unremarkable subjects needed to be qualified only to the extent that only a selective portion of the normal population could be reached through Turkish associations and personal contact. Also, since the reliability of the data obtained on the outpatients was limited owing to the small number of subjects (n = 39) (Bühner, 2006), they were not incorporated into or discussed in the results. A minimum of n = 100 subjects is considered necessary confidently to assess the reliability of a test (Mendoza, Stafford, & Stauffer, 2000).

The SCL-90-R proved more suitable for identifying and filtering out psychologically challenged persons, giving it a slight advantage over the ICD-10 as a global screening instrument (Fischer et al., 2010).

The correlation between the SCL-90-R scales was markedly higher than that of ICD-10 scales (Tritt et al., 2008); this difference in relations was statistically significant for both the group of unremarkable subjects and that undergoing inpatient treatment. An explanation for this can be found at the items level. For instance, the SCL compulsion scale comprises items relating to memory difficulties and a feeling of lack of motivation. These items are equally suited for assessing depression and/or cognitive deficits (Tritt et al., 2008).

Hitherto, it has not been possible, with the aid of the SCL-90-R, to establish the intended independence of individual symptom-related scales in different populations such as a normal population or psychosomatic patients (Olsen, Mortensen, & Bech, 2004; Rief, Greitermeyer, & Fichter, 1991; Schmitz, Hartkamp, Kiuse, Franke, Reister, & Tress, 2000); by in contrast, by using the ISR, it is more likely to be possible to diagnose disorder-specific conditions beyond the overall severity factor (Fischer et al., 2010).

As regards linguistic comparability, it can be said that, in general, comparable reliability values of the German and Turkish versions have been achieved. In both languages, the instrument has apparently evoked similar concepts in the respondents. It was possible sufficiently to confirm the criterion validity in all cases via correlations with the two quasi criteria of SCL-90-R and BDI. Moreover, the strong correlation between the overall ISR score and the SCL-90-GSI spoke in favour of a valid capture of the criterion of overall severity.

Further investigations with valid and reliable diagnostic test instruments would be required to assess the criterion validity of the *Eating disorder syndrome* scale, which was not analysed.

It would also be advisable to analyse the criterion validity of all ICD-10 Symptom Rating scales with regard to real external criteria and target criteria.

Larger samples are required to reach more reliable findings. Further objectives are a five-to ten-fold parameter number of a model or samples with *n* > 250 (Scholderer & Balderjahn, 2006).

The stability of the individual ICD-10 Symptom Rating scales has been substantiated by the present investigation. However, larger samples are needed to reach more reliable findings regarding the reliability of the scales for their use in the outpatient setting and regarding the intended syndrome structure of the instrument. Validations of the ISR scales using other established diagnostic test instruments, clinical diagnoses and interviews would be advisable in order to assess the validity and diagnostic precision of the Turkish-language instrument.

It should be mentioned that the ISR is not a replacement for more intensive and lengthy diagnostic protocols, but serves as a reliable and valid screening measure. The ISR is of course not the only method for arriving at a diagnosis and there is a need of more research in different kind of groups and in different cultures.

This study has several limitations. We gathered the data from psychosomatic patients and not psychiatric patients in this study. Data collection from a group of psychiatric patients will be necessary in future studies. When a cut-off point is set, utility is high for clinical application.

Furthermore, we believe that standardisation data from patients of Turkish descent living in Germany and from those living in Turkey from several regions in Turkey as a reference population are absolutely necessary to have a culture-sensitive approach in accordance with the Test Adaption Guidelines of the International Test Commission (Hambleton, 2001; Moosbrugger & Höfling, 2007).

References

Aktürk, Z., Dagdeviren, N., Türe, M., & Tuglu, C. (2005). The reliability and validity analysis of the Turkish version of Beck Depression Inventory for primary care. *Türkiye Aile Hekimliği Dergisi, 9*(3), 117-122.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571. doi:10.1001/archpsyc.1961.01710120031004

Brandt, J. (2009). *Kreuzvalidierung des ICD-10-Symptom-Rating (ISR) Selbstbeurteilungsfragebogens an 969 depressiven Patienten mit der SCL-90-R und dem BDI I* (Unpublished doctoral dissertation). Universität Regensburg, Regensburg.

Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.

Dag, I. (1991). Reliability and validity of the Symptom Checklist (SCL-90-R) for university students. *Türk Psikiyatri Dergisi, 2*, 5-12.

Derogatis, L. R. (1977). *SCL-90-R, administration, scoring & procedures manual-I for the R(evised) version*. Baltimore: John Hopkins University School of Medicine.

Fischer, H. F., Tritt, K., Klapp, B. F., & Fliege, H. (2010). Faktorstruktur und psychometrische Eigenschaften des ICD-10–Symptom-Rating (ISR) an Stichproben psychosomatischer Patienten. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 60*, 307-315. doi:10.1055/s-0029-1214419

Franke, G. H. (2001). *Brief Symptom Inventory (BSI) von L. R. Derogatis: Deutsches Manual*. Göttingen: Beltz.

Franke, G. H. (2002). *SCL-90-R. Die Symptom-Checkliste von Derogatis: Deutsche Version* (2nd ed.). Göttingen: Beltz.

Freyberger, H. J., & Stieglitz, R.-D. (2005). Krankheitsbilder, Klassifikation, Dokumentation. In W. Senf, & M. Broda (Eds.), *Praxis der Psychotherapie* (pp. 134-143). Stuttgart: Thieme.

Goldberg, D. P. (2008). General Health Questionnaire (GHQ). In A. Rush, M. First, D. Blacker (Eds.), *Handbook of psychiatric measures* (pp. 67-70). Washington, DC: American Psychiatric Publishing.

Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem „Gesundheitsfragebogen für Patienten (PHQ-D)": Ergebnisse der deutschen Validierungsstudie. *Diagnostica, 50*, 171-181. doi:10.1026/0012-1924.50.4.171

Grawe, K., & Braun, U. (1994). Qualitätskontrolle in der Psychotherapiepraxis. *Psychologische Rundschau, 23*, 242-267.

Haasen, C., Kleinmeier, E., & Yadirgan, O. (2005). Kulturelle Aspekte bei der Diagnostik psychischer Störungen. In J. Assion (Ed.), *Migration und seelische Gesundheit* (pp. 145–153). Heidelberg: Springer.

Hambleton, R. K. (2001). The next generation of the ITC Test Translation and Adaptation Guidelines. *European Journal of Psychological Assessment, 17*(3), 164-172. doi:10.1027//1015-5759.17.3.164

Heine, P., & Assion, H. J. (2005). Traditionelle Medizin in islamischen Kulturen. In J. Assion (Ed.), *Migration und seelische Gesundheit* (pp. 29-42). Heidelberg: Springer.

Heymann, F., Zaudig, M., & Tritt, K. (2003). Die diagnosebezogene Behandlungsdauer in der psychosomatischen und psychotherapeutischen Medizin: Eine homogene Größe? Erste Ergebnisse der Multicenter-Basisdokumentation (Psy-BaDo-PTM) als Grundlage qualitätssichernder Maßnahmen in der stationären Psychosomatik. *Praxis der klinischen Verhaltensmedizin und Rehabilitation, 62*, 209-221.

Hill, C. E., & Lambert, M. J. (2000). Methodological issues in studying psychotherapy processes and outcomes. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 84-125). New York: Wiley.

Hisli, N. (1989). Beck Depresyon Envanteri'nin üniversite öğrencileri için geçerliği ve güvenilirliği. *Türk Psikoloji Dergisi, 6*, 3-13.

Kizilhan, J. I. (2010). Migration, Identität und Gesundheit. *Familiendynamik, 35*(1), 50-59.

Kizilhan, J. I. (2011). Understanding and treatment of diffuse aches and pains of patients from tradition-bound cultures. *Europe's Journal of Psychology, 7*(2), 359-373. doi:10.5964/ejop.v7i2.134

Kizilhan, J. I. (2012). Schmerzhafte Konflikte. *Gehirn & Geist, 7*, 60-68.

Koch, E. (1997). *Anamneseleitfaden für türkische Patienten am Psychiatrischen Krankenhaus Marburg.* Unpublished manuscript, Psychiatrisches Krankenhaus Marburg, Marburg, Germany.

Mendoza, J. L., Stafford, K. L., & Stauffer, J. M. (2000). Large-sample confidence intervals for validity and reliability coefficients. *Psychological Methods, 5*, 356-369. doi:10.1037/1082-989X.5.3.356

Moosbrugger, H., & Höfling, V. (2007). Standards für psychologisches Testen. In H. Moosbrugger & A. Kelava (Ed.), *Testtheorie und Fragebogenkonstruktion* (pp. 195-212). Berlin: Springer.

Olsen, L. R., Mortensen, E. L., & Bech, P. (2004). The SCL-90 and SCL-90R versions validated by item response models in a Danish community sample. *Acta Psychiatrica Scandinavica, 110*, 225-229. doi:10.1111/j.1600-0447.2004.00399.x

Rief, W., Greitermeyer, M., & Fichter, M. (1991). Die Symptom-Checklist SCL-90-R: Überprüfung an 900 psychosomatischen Patienten. *Diagnostica, 37*, 58-65.

Schirmer, N. C. (2009). *Prüfung der Reliabilität des ICD-10-Symptom-Ratings an einer klinisch unauffälligen Stichprobe* (Unpublished diploma thesis). Regensburg: Universität Regensburg, Institut für Psychologie.

Schmitz, N., Hartkamp, N., Kiuse, J., Franke, G. H., Reister, G., & Tress, W. (2000). The Symptom Check-List-90-R (SCL-90-R): A German validation study. *Quality of Life Research, 9*, 185-193. doi:10.1023/A:1008931926181

Scholderer, J., & Balderjahn, I. (2006). Was unterscheidet harte und weiche Strukturgleichungsmodelle nun wirklich? Ein Klärungsversuch zur LISREL-PLS-Frage. *Marketing, 28*, 57-70.

Tritt, K., Peseschkian, H., & Bidmon, R. K. (2002). Psychodynamische Psychotherapie bei somatoformen Störungen im symptomatischen Monitoring: Was kann der SCL-90-R hier in der Versorgungsrealität leisten? *Psychodynamische Psychotherapie, 1*, 212-224.

Tritt, K., von Heymann, F., Zaudig, M., Zacharias, I., Söllner, W., & Loew, T. (2008). Entwicklung des Fragebogens „ICD-10-Symptom-Rating" (ISR). *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 4*, 409-418.

Wittchen, H. U., & Perkonigg, A. (1997). *DIA-X-Screening-Verfahren: Fragebogen DIA-SSQ: Screening für psychische Störungen; Fragebogen DIA-ASQ: Screening für Angststörungen; Fragebogen DIA-DSQ: Screening für Depressionen.* Frankfurt: Swets und Zeitlinger.

About the Authors

Prof. Dr. **Jan Ilhan Kizilhan** is a psychologist and psychotherapist. His expertise is in psychotraumatology, social medicine (rehabilitation and retirement), culture sensitive psychiatry and psychotherapy, forensic and ethnological studies.

Dipl. Psych. **Antje Roniger**, doctorate student in clinical psychology, specifically in clinical psychology and psychometric tests

Dr. **Friedrich von Heymann**, has a background as a general doctor in Munich and is a head of Department of Quality Development in Psychotherapy and Psychosomatic Medicine (IQP) in Munich.

Prof. Dr. **Karin Tritt**, is a psychologist and psychotherapist. Her expertise is in clinical psychology, social-medicine, psychosomatic and psychometric tests.