

A very short Symptom Checklist-90-R version for routine outcome monitoring in psychotherapy; The SCL-3/7

Reinier Timman¹  | Willem A. Arrindell² 

¹Section of Medical Psychology and Psychotherapy, Department of Psychiatry, Erasmus Medical Centre, Rotterdam, The Netherlands

²University of Social Sciences and Humanities, Vietnam National University, Ho Chi Minh City, Vietnam

Correspondence

Reinier Timman, Section of Medical Psychology and Psychotherapy, Department of Psychiatry, Erasmus Medical Centre, Wytewaweg 80, 3015 CN Rotterdam, P.O.Box 2040, NL-3000 CA Rotterdam, The Netherlands.
Email: r.timman@erasmusmc.nl

Funding information

The authors declare no financial disclosures.

Abstract

Objective: Routine outcome monitoring (ROM) is applied in many physical and mental health treatments. The treatment course is monitored with patient reported outcome measures (PROMs). A potential problem with PROM is response burden. This can be decreased by presenting such measures with less and better selected items. The SCL-90-R is an often used PROM for psychotherapies and a number of very short forms have been developed; the SCL-5, SCL-8, SCL-9 and SCL-10. This study aims to develop a new very short form, the symptom checklist 3 out of 7 (SCL-3/7) and to evaluate the effectiveness of these PROM with the precision relative to the complete SCL-90-R score.

Methods: Item Response Theory analysis was applied to select the 7 best discriminating items, evenly distributed over the latent trait. A routing serves that patients only need to administer 3 items.

Results: In a sample of 15,055 cases, the relative precisions of the SCL-3/7 were best for outpatients (122.7%), day care patients (111.8%) and inpatients (108.3). The SCL-5 was best for juvenile patients (110.0%), and the SCL-9 was best for addicted patients (107.2%).

Conclusion: The SCL-3/7 decreases patient burden in ROM and has a better precision in adult therapies than other SCL-90 short forms.

KEYWORDS

IRT analysis, patient reported outcome measures, response burden, routine outcome monitoring, SCL-90-R

1 | INTRODUCTION

In mental health services, the Symptom Checklist-90-Revised (SCL-90-R)¹ is an often used tool for measuring progress of treatments in routine outcome monitoring (ROM). In routine outcome monitoring, the patient's perceived mental health state is estimated with patient reported outcome measures (PROMs). Patient reported

outcome measures in mental health treatment are considered a valuable addition to medical outcomes in effectiveness and cost-effectiveness evaluations in clinical trials and quality improvement. Also, PROMs are regularly applied by therapists to monitor the treatment effect from the patient's perspective. It stimulates patient participation and shared clinical decision-making.² PROMs are also considered to make the quality of care more transparent

to patients, the government and financing bodies such as health insurers. In psychotherapy, systematic routine measurement of patient reported outcome measures has been taking place for decades. In fact, some elements of routine clinical outcome measurement (that of clinical change, intervention or context) have been described to have been implemented in mental health services in the United Kingdom and elsewhere for at least 150 years.³ For the routine use of measuring devices in regular care, these should be short and easy to administer. Most of the available patient reported outcome measures are valid and reliable for research, but are generally too long for frequent use.

1.1 | The Symptom Checklist-90-Revised and short forms

The SCL-90-R^{1,4} is a widely used patient reported outcome measure in clinical trials in psychiatry and is applied for routine outcome monitoring by clinical practitioners. However, the SCL-90-R with its current length of 90 items is too time-consuming and cumbersome for patients for routine use in regular care. This may result in response burden. Although response burden is difficult to precisely define and operationalise, the length of surveys may obviously be its most important cause. Response burden could represent a potential barrier for clinical practitioners to implement a standardised outcome-assessment strategy (e.g. Hatfield & Ogles⁵). Besides length, the quality of the content also plays an important part.⁶ Response burden occurs when respondents' motivation drops as a result of the length of a survey and hence the data quality begins to deteriorate. A number of very short forms of the SCL-90-R have been developed. A 5-item version was presented in 1993 by Tambs & Moum,⁷ an 8-item version was reported by Fink et al.,⁸ a 9-item version by Petrowsky et al.⁹ and a 10-item version by Strand et al.¹⁰

1.2 | Aims of the study

This study serves two aims; to construct a very short version of the Symptom Checklist, suitable for routine use in regular psychotherapeutic care with a minimal loss of information and to compare its effectiveness with the other very short forms of the SCL-90-R. For the first aim, the focus is to retain the range of the latent trait(s) as wide as possible, while the scale will still be sensitive for patients with severe as well as mild mental problems. As such, our methods differed from those in previous studies. Fink et al.,⁸ for example, also used IRT analysis, but only used the discriminative value of the items, not the position on the latent trait.

Significant outcomes

- The SCL-3/7 has a better relative precision than the complete SCL-90-R and other very short forms of the SCL-90-R.
- The response burden for patients in ROM is minimised with only three most relevant questions to be answered.

Limitations

- The SCL-3/7 is not suited for use in routine outcome monitoring therapies of addicted patients. For that goal, the SCL-9 serves best.
- The SCL-3/7 is not suited for diagnostic purposes; the SCL-3/7 is not a case-finding instrument.

The operationalisation of our first aim was to create the 'SCL-3 out of 7' (SCL-3/7), by reducing the SCL-90-R to seven items. By using smart routing, patients are only required to answer 3 of the 7 items, as items out of range would not be presented to them. For instance, if the first item already indicated that the patient had severe mental problems, items about minor problems would not be presented.

2 | MATERIAL AND METHODS

2.1 | Study sample

A sample of 14,036 administrations was collected in the Standard Evaluation Project.¹¹ This was a project grounded by the Stichting Klinische Psychotherapie (SKP) with 10 participating mental clinics in the Netherlands. Four subgroups were distinguished within this sample: outpatients, day care patients, inpatients and juvenile patients. Outpatients had one or more individual hours of therapy per week, fortnight or month. Day care patients had one or more days of therapy each week, and inpatients stayed overnight and were released in the weekends. All data were administered with paper and pencil. In principle, administration was at five time points: at start of the treatment, at the end of treatment and follow-ups at 6 months and 1 year. Some participating mental clinics also had an interim administration during the treatment. Juvenile patients were younger than 20 years of age. Another sample comprised 1019 consecutive applicants to outpatient treatment facilities for addiction (substance use and impulse-control disorders) at Novadic-Kentron in Roosendaal and Bergen op Zoom in The Netherlands. This resulted in a total sample of 15,055 participants. The background variables are presented in Table 1.

2.2 | Measure

The SCL-90-R is a 90-item self-report symptom inventory designed to reflect psychological symptom patterns of psychiatric and medical patients. Each item of the questionnaire is rated on a 5-point scale of distress ranging from 1 (not at all) to 5 (extremely). The SCL-90-R consists of the following nine primary symptom dimensions: somatisation, which reflects distress arising from bodily perceptions; obsessive-compulsive, which reflects obsessive-compulsive symptoms; interpersonal sensitivity, which reflects feelings of personal inadequacy and inferiority in comparison with others; depression, which reflects depressive symptoms, as well as lack of motivation; anxiety, which reflects anxiety symptoms and tension; hostility, which reflects symptoms of negative affect, aggression and irritability; phobic anxiety, which reflects symptoms of persistent fears as responses to specific conditions; paranoid ideation, which reflects symptoms of projective thinking, hostility, suspiciousness, fear of loss of autonomy; and psychoticism, which reflects a broad range of symptoms from mild interpersonal alienation to dramatic evidence of psychosis. A total score termed general psychological distress is calculated by summing across all 90 items for obtaining an overall index of an individual's mental state.^{12,13} The Dutch form of the SCL-90-R was administered.⁴

2.3 | Ethics

The local ethics committees of all participating clinics approved of the data collection. The SCL-90-R was administered for means of diagnosis and all data entry was performed locally.

The Medical Ethics Committee of the Erasmus Medical Centre (Rotterdam, the Netherlands) judged that according to Dutch law the current study did not require a formal approval as the data were anonymous and had been collected in previously approved studies.

2.4 | Data analysis

2.4.1 | SCL-3/7 construction

For the construction of the SCL-3/7, we applied a methodology very similar to the one used for the development of the Visual Function Questionnaire (VFQ-3007).¹⁴

In a first selection, we excluded items with more than 1.0% missing values. This was a strict criterion, as for the proposed 'routing procedure' (see below) missing data would have been problematic.

TABLE 1 Participant characteristics

| | Outpatients | | Day care | | Inpatients | | Juvenile | | Addicted | | Total | |
|-----------------------------------|-------------|-----------|----------|-----------|------------|-----------|----------|-----------|----------|-----------|----------|-----------|
| | n | % | n | % | n | % | n | % | n | % | n | % |
| Female | 405 | 60.7 | 1029 | 70.9 | 1991 | 68.8 | 357 | 81.9 | 231 | 25.4 | 4013 | 63.1 |
| Male | 262 | 39.3 | 423 | 29.1 | 902 | 31.2 | 79 | 18.1 | 678 | 74.6 | 2344 | 36.9 |
| Single | 232 | 67.6 | 855 | 80.3 | 1257 | 86.3 | 262 | 99.6 | NA | NA | 2606 | 83.3 |
| Married/ cohabited | 111 | 32.4 | 210 | 19.7 | 200 | 13.7 | 1 | 0.4 | NA | NA | 522 | 16.7 |
| Baseline | 664 | | 1595 | | 3558 | | 488 | | 906 | | 7211 | |
| Follow-ups | | | | | | | | | | | | |
| 1 | 96 | | 40 | | 298 | | 0 | | 0 | | 434 | |
| 2 | 216 | | 732 | | 1818 | | 244 | | 102 | | 3112 | |
| 3 | 43 | | 471 | | 960 | | 163 | | 10 | | 1647 | |
| 4 | 66 | | 378 | | 836 | | 137 | | 1 | | 1418 | |
| 5 | 94 | | 270 | | 749 | | 120 | | | | 1233 | |
| Total | 1179 | | 3486 | | 8219 | | 1152 | | 1019 | | 15055 | |
| | M | SD | M | SD | M | SD | M | SD | M | SD | M | SD |
| Age | 35.7 | 11.3 | 32.2 | 9.2 | 30.7 | 10.0 | 17.2 | 1.4 | 37.5 | 13.0 | 31.6 | 11.1 |
| Education (years) ^a | 12.2 | 3.0 | 12.3 | 2.8 | 12.2 | 3.0 | 8.4 | 2.5 | NA | NA | 11.9 | 3.0 |
| SCL-90-R total | 191.5 | 58.9 | 211.7 | 55.4 | 208.3 | 62.7 | 211.8 | 69.7 | 174.4 | 62.3 | 204.0 | 62.5 |

Abbreviation: NA, not available.

^aElementary school 6 years; advanced elementary 8 years; lower vocational 10 years; advanced vocational 12 years; higher vocational 15 years and university 17 years.

A principal components analysis (PCA) was performed to check for potential uni-dimensionality and select items. Uni-dimensionality was predicted on the basis of recent findings with the SCL-90-R showing that (a) the general distress component emerges as a potent component which is substantially loaded by each item^{15,16}; and related herewith (b) McDonald's ordinal omega as a measure of overall scale quality reflects values in excess of 0.90.^{15,17} Accordingly, the following observations were required: First, the first unrotated component in PCA should have an eigenvalue and corresponding explained variance much larger than equivalent values for the remaining components. Second, all items should have a positive and at least medium-effect sized loading on the first unrotated component. Note that component loadings resemble correlations, and according to Cohen,¹⁸ the magnitude of a correlation can be expressed in terms of an effect size. Cohen states that correlations higher than 0.30 correspond with a medium-effect size and higher than 0.50 reflect a large effect size. To aid in selecting the most powerful items, however, it was required that an item should load at least 0.50 to reflect a large effect size according to Cohen,¹⁸ on the first unrotated component. Third, McDonald's omega should be at least 0.90. Fourth, a scree test should confirm the number of components.

Next, cases with more than 5 missing items were excluded. The remaining missing items were imputed, using the mean of ten linear regression imputations.

All items complying with these unidimensional requirements were analysed with a generalised Partial Credit Model (gPCM). This is a two parameter Item Response Theory (IRT) model for ordered categories. In IRT and Rasch models, items are ordered on their position on a latent trait, such as intelligence or psychoneuroticism. The Rasch model is used for binary items and calculates one parameter, namely the position on the latent trait. Other models such as the generalised Partial Credit Model (gPCM) and the Graded Response Model (GRM) can also calculate discriminative properties of items. More information on IRT and the Rasch model can be found elsewhere.^{19–23}

The generalised Partial Credit Model assumes equal differences between the answer categories over the items. This makes an ordering of the items on the latent trait possible, based on the item measure, and provides item differentiation parameters. IRT analysis also allows to express the respondent's performance on this same latent trait, the person measure.¹⁹ The data included pre-treatment baseline scores and one or more follow-up measures after treatment. This is not in accordance with the independence of measurement assumption. Generally, this can be overcome by performing multilevel analyses, where the

persons form the upper level and their repeated measures the lower level. Unfortunately, the Winsteps programme is not capable of performing multilevel analyses. Therefore, we applied a procedure to estimate the effect of bypassing the multilevel structure.²⁴ This procedure is described more detailed in Appendix S1, from which it was concluded that violation of measurement independence had no appreciable influence on the outcome of the analyses.

Seven items were deemed sufficient for a broad classification in a(n) (computerised) administration, where routing reduced the number of presented items to three (Figure 1). The selection was done by classifying the latent variable into seven classes, and from each class the best discriminating item was selected for the final version of the SCL-3/7.

2.4.2 | Routing of the SCL-3/7

The first item to be filled in is in the middle of the latent trait, the second on a quarter or three quarters, depending on the answer on the first item. The routing was determined by the medians. The answer on the second item determined which of the remaining four items would be presented as the third item (Figure 1).

2.4.3 | Validating the SCL-3/7

To test the statistical validity of the SCL-3/7, the following analyses were performed.

Fit statistics

Infit and outfit measures are mean squares provided by Winsteps for detecting poorly fitted items. Mean squares greater than 1.0 indicate an underfit to the model and mean squares less than 1.0 indicate an overfit, where values between 0.7 and 1.3 are considered acceptable.²⁵

Differential item functioning

Differential item functioning (DIF) may occur when a test item does not have the same relationship to a latent variable across two or more groups.¹⁹ That means that persons from different groups who have the same position on the latent trait will have a different outcome. In the present study, DIF was discerned for the five treatment groups on the 7 applied items. For large samples, the DIF t-value is unduly often significant.²⁶ To compensate for this, we applied the normalising procedure described at the IRT Organisation site, and adjusted the standard errors with $\sqrt{(N/100)}$. We preferred to construct one general SCL-3/7 for all treatment groups, but in case of severe DIF, we considered to construct separate versions.

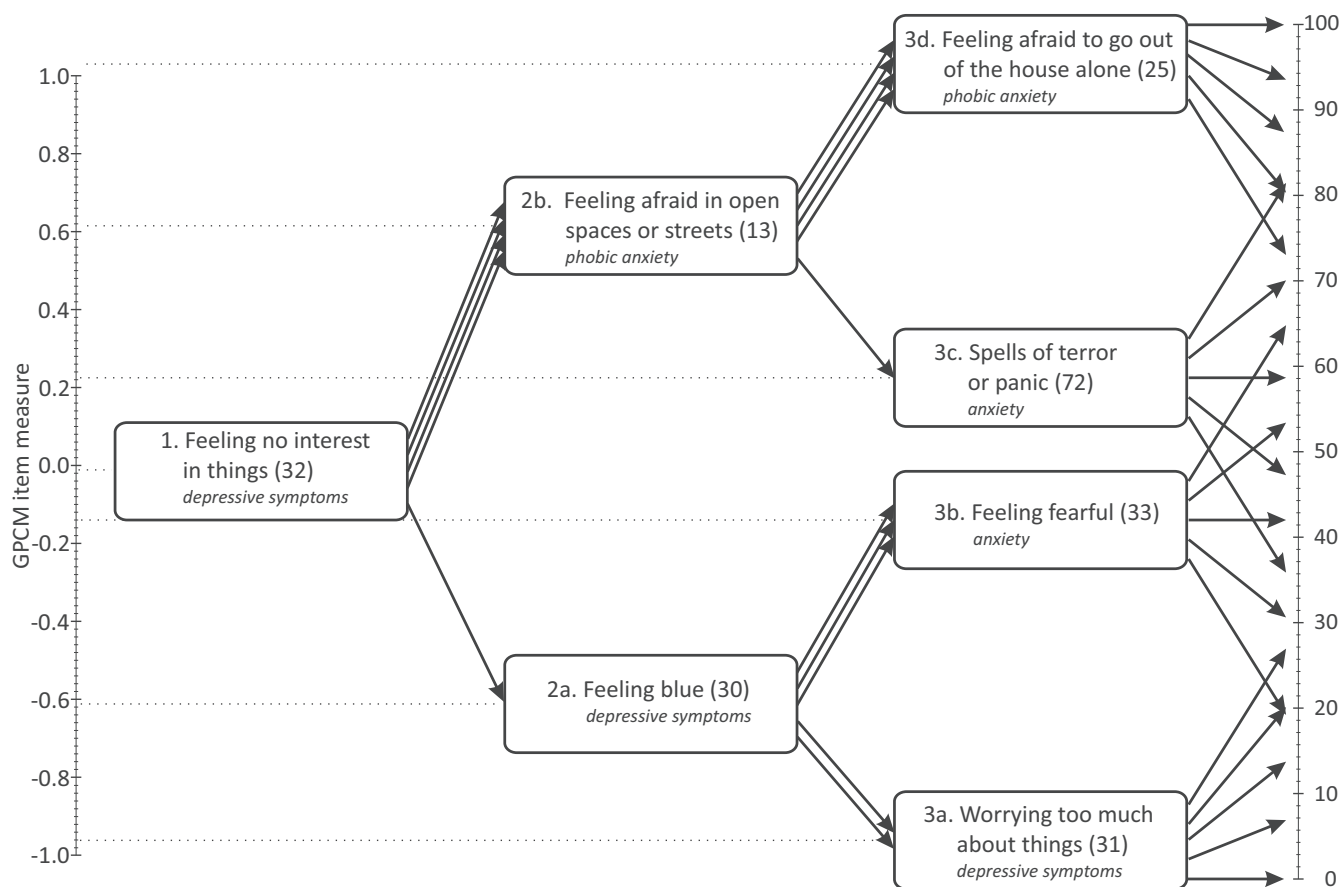


FIGURE 1 Schematic presentation of the SCL-3/7. Rasch analysis allows expressing the respondents' performance on the same latent trait as the item measure. First item 32, in the middle of the latent trait, is administered. Depending on the answer, the respondent is routed through the questionnaire. Every arrow represents an answer category, and the split is determined by the median of the item. In the end, only three out of the seven items are used, where the answers navigate patients to a fitting trait level

Regimen specific analyses

For practical reasons and optimisation of generalisation, one uniform SCL-3/7 is preferred; however, we performed separate analyses for the five treatment regimens, leading to different versions of the SCL-3/7. We applied sensitivity analyses within the various samples in order to decide whether it was worthwhile to have different versions for each particular regimen.

Item weights for calculating the SCL-3/7 score

Lastly, in an iterative procedure, weights for the SCL-3/7 score were determined by maximising the Pearson correlation with the generalised Partial Credit Model measure as criterion. After a logit transformation, the scores were rescaled to a range from 0 (no problems) to 100 (severe problems).

2.4.4 | Comparing the short form versions

For all questionnaires, we reported McDonald's omega²⁷ and Cronbach's alpha,²⁸ as well as the correlation with the

classical SCL-90-R score. An adequate measure for monitoring should be sensitivity for change. Certainly, when a therapy proceeds in the wrong direction, the therapist should be warned, but also when the therapy is going right, that is useful information. The sensitivity for change of the short form patient reported outcome measures was determined with the relative precision, a method described by McHorney et al.^{29,30} Assuming that the level of psychological complaints decreases during therapy, the measure that signifies that at best is the measure that shows the most significant change. Thus, when using an F-test, the measure with the largest F-value related to time indicates the most significant effect. This relative precision was calculated for the treatment effect at the follow-up with the largest number of responses compared with baseline. In 1992, when McHorney et al.²⁹ presented their paper, repeated measures ANOVA, an F-test, was generally applied for longitudinal analyses. We adjusted this method using random effects models for the determination of the effects instead of complete cases F-tests. In doing so, all data were incorporated, including those from persons without follow-up measures. We applied data of all datasets together

and per treatment apart and calculated the relative precision of the random effect model t-values. The value of the original method for computing the SCL-90-R was applied as reference value; thus, for the SCL-90-R, the relative precision was set at 100%. Note that the development of the SCL-3/7 is performed independently of testing its effectiveness. Thus, there was no need to split the data into an exploratory and a confirmatory part.

2.4.5 | Applied statistical programs

Item Response Theory generalised Partial Credit Model analyses were performed with Winsteps version 4.5.5 (Linacre, J. M. [2020]. Winsteps® Rasch measurement computer program. Winsteps.com). McDonald's omega was calculated in R, version 4.0.2 using packagePsych.³¹ All other analyses were performed with SPSS version 25 (IBM SPSS Statistics for Windows: IBM Corp.)

3 | RESULTS

3.1 | SCL-3/7 construction

The exclusion of items with 1.0% missing values or more concerned item 35 (1.6%), item 10 (1.4%) and item 5 (1.0%, Table 2, Figure 2). Forty-four cases had more than 5 missing values and were deleted. Data of the remaining 15,011 cases were imputed and entered into the generalised Partial Credit Model analysis. The first unrotated component had an explained variance of 37.5% and 22 items with a too low component loading (<0.50, Table 3), even though all items loaded positively and in excess of 0.3. The explained variance of the second component was 3.7%. The scree test clearly suggested a unidimensional structure (Figure 3). Accordingly, 65 items were selected for generalised Partial Credit Model analyses. Ordering of the items on the basis of the latent trait, classification and selection of the most discriminating items per class, resulted in the selection of the items 31, 30, 33, 32, 72, 13 and 25 (Table 3).

3.2 | SCL-3/7 validation

3.2.1 | Fit statistics

The item infit mean square measure for the 65-item generalised Partial Credit Model analysis was 1.05, and the outfit measure was 1.02. For the seven item analysis (including all scored categories), the corresponding measures

TABLE 2 Original SCL-90-R items, PCA loadings, percentage of missing values and items included in SCL-90-R short forms

| | | PCA loading | % missing |
|----|--|-------------|-----------|
| 30 | ⁵⁾⁷⁾⁸⁾¹⁰⁾ Feeling blue | 0.805 | 0.1 |
| 57 | ⁹⁾¹⁰⁾ Feeling tensed or keyed up | 0.784 | 0.1 |
| 79 | ⁸⁾¹⁰⁾ Feelings of worthlessness | 0.784 | 0.1 |
| 71 | ⁸⁾¹⁰⁾ Feeling everything is an effort | 0.772 | 0.1 |
| 31 | ⁵⁾⁷⁾⁸⁾⁹⁾ Worrying too much about things | 0.769 | 0.1 |
| 77 | ⁹⁾ Feeling lonely even when you are with people | 0.768 | 0.1 |
| 54 | ⁵⁾⁸⁾¹⁰⁾ Feeling hopeless about the future | 0.763 | 0.1 |
| 51 | Your mind going blank | 0.756 | 0.1 |
| 34 | Your feelings being easily hurt | 0.752 | 0.2 |
| 33 | ⁵⁾⁷⁾⁸⁾¹⁰⁾ Feeling fearful | 0.748 | 0.2 |
| 3 | Unwanted thoughts, words or ideas that won't leave your mind | 0.747 | 0.3 |
| 90 | The idea that something is wrong with your mind | 0.742 | 0.2 |
| 29 | Feeling lonely | 0.741 | 0.1 |
| 55 | Trouble concentrating | 0.736 | 0.1 |
| 22 | Feeling of being trapped or caught | 0.735 | 0.2 |
| 26 | ¹⁰⁾ Blaming yourself for things | 0.733 | 0.2 |
| 41 | Feeling inferior to others | 0.732 | 0.1 |
| 61 | Feeling uneasy when people are watching or talking about you | 0.732 | 0.1 |
| 32 | ⁷⁾ Feeling no interest in things | 0.719 | 0.1 |
| 36 | Feeling others do not understand you or are unsympathetic | 0.717 | 0.1 |
| 72 | ^{7,8)} Spells of terror or panic | 0.716 | 0.3 |
| 28 | ⁹⁾ Feeling blocked in getting things done | 0.713 | 0.2 |
| 89 | Feelings of guilt | 0.713 | 0.2 |
| 23 | Suddenly scared for no reason | 0.700 | 0.1 |
| 69 | Feeling very self-conscious with others | 0.695 | 0.4 |
| 37 | Feeling that people are unfriendly or dislike you | 0.691 | 0.1 |
| 86 | Feeling pushed to get things done | 0.690 | 0.3 |

TABLE 2 (Continued)

| | | PCA loading | % missing |
|-------------------|--|----------------|--------------|
| 43 ⁹⁾ | Feeling that you are watched or talked about by others | 0.687 | 0.1 |
| 18 | Feeling that most people cannot be trusted | 0.683 | 0.2 |
| 2 ^{5/8)} | Nervousness or shakiness inside | 0.677 | 0.4 |
| 14 | Feeling low in energy or slowed down | 0.677 | 0.2 |
| 59 | Thoughts of death or dying | 0.676 | 0.1 |
| 46 | Difficulty making decisions | 0.663 | 0.1 |
| 11 | Feeling easily annoyed or irritated | 0.661 | 0.2 |
| 50 | Having to avoid certain things, places or activities because they frighten | 0.659 | 0.1 |
| 80 | Feeling that familiar things are strange or unreal | 0.655 | 0.1 |
| 56 | Feeling weak in parts of your body | 0.653 | 0.2 |
| 70 | Feeling uneasy in crowds, such as shopping or at a movie | 0.649 | 0.2 |
| 13 ⁷⁾ | Feeling afraid in open spaces or on the streets | 0.632 | 0.1 |
| 78 | Feeling so restless you couldn't sit still | 0.621 | 0.1 |
| 76 | Others not giving you proper credit for your achievements | 0.620 | 0.2 |
| 15 | Thoughts of ending your life | 0.619 | 0.2 |
| 75 ⁹⁾ | Feeling nervous when you are left alone | 0.616 | 0.1 |
| 83 | Feeling that people will take advantage of you if you let them | 0.616 | 0.1 |
| 68 | Having ideas or beliefs that others do not share | 0.604 | 0.3 |
| 58 ⁹⁾ | Heavy feelings in your arms or legs | 0.603 | 0.2 |
| 17 | Trembling | 0.597 | 0.1 |
| 49 | Hot or cold spells | 0.597 | 0.1 |
| 53 | A lump in your throat | 0.580 | 0.1 |
| 66 | Sleep that is restless or disturbed | 0.579 | 0.2 |

(Continues)

TABLE 2 (Continued)

| | | PCA loading | % missing |
|-------------------|--|----------------|--------------|
| 9 | Trouble remembering things | 0.578 | 0.1 |
| 73 | Feeling uncomfortable about eating or drinking in public | 0.570 | 0.1 |
| 25 ⁷⁾ | Feeling afraid to go out of your house alone | 0.561 | 0.1 |
| 40 | Nausea or upset stomach | 0.561 | 0.2 |
| 45 | Having to check and double-check what you do | 0.561 | 0.1 |
| 87 | The idea that something serious is wrong with your body | 0.557 | 0.2 |
| 44 ¹⁰⁾ | Trouble falling asleep | 0.552 | 0.1 |
| 38 | Having to do things very slowly to ensure correctness | 0.548 | 0.2 |
| 52 | Numbness or tingling in parts of your body | 0.541 | 0.1 |
| 67 | Having urges to break or smash things | 0.539 | 0.2 |
| 85 | The idea that you should be punished for your sins | 0.533 | 0.2 |
| 88 | Never feeling close to another person | 0.522 | 0.2 |
| 21 | Feeling shy or uneasy with the opposite sex | 0.515 | 0.1 |
| 48 | Trouble getting your breath | 0.505 | 0.1 |
| 47 | Feeling afraid to travel on buses, subways, trains | 0.501 | 0.2 |
| 62 | Having thoughts that are not your own | 0.497 | 0.6 |
| 4 ¹⁰⁾ | Faintness or dizziness | 0.493 | 0.8 |
| 39 | Heart pounding or racing | 0.492 | 0.2 |
| 20 ⁹⁾ | Crying easily | 0.488 | 0.1 |
| 24 ⁹⁾ | Temper outbursts that you could not control | 0.484 | 0.1 |
| 7 | The idea that someone else can control your thoughts | 0.471 | 0.2 |
| 19 | Poor appetite | 0.471 | 0.2 |
| 6 | Feeling critical of others | 0.462 | 0.3 |
| 42 | Soreness of your muscles | 0.456 | 0.2 |
| 74 | Getting into frequent arguments | 0.451 | 0.2 |

(Continues)

TABLE 2 (Continued)

| | | PCA loading | % missing |
|----|--|----------------|--------------|
| 12 | Pains in heart or chest | 0.447 | 0.1 |
| 81 | Shouting or throwing things | 0.433 | 0.1 |
| 1 | Headaches | 0.421 | 0.2 |
| 63 | Having urges to beat, injure or harm someone | 0.420 | 0.2 |
| 8 | Feeling others are to blame for most of your troubles | 0.413 | 0.1 |
| 64 | Awakening in the early morning | 0.402 | 0.1 |
| 82 | Feeling afraid you will faint in public | 0.379 | 0.1 |
| 27 | Pains in lower back | 0.376 | 0.1 |
| 65 | Having to repeat the same actions such as touching, counting and washing | 0.362 | 0.2 |
| 84 | Having thoughts about sex that bother you a lot | 0.361 | 0.2 |
| 60 | Overeating | 0.353 | 0.1 |
| 16 | Hearing voices that other people do not hear | 0.325 | 0.1 |
| 5 | Loss of sexual interest or pleasure | | 1.0 |
| 10 | Worried about sloppiness or carelessness | | 1.5 |
| 35 | Other people being aware of your private thoughts | | 1.7 |

Note: ⁵⁾ Items selected for the SCL-5 (7); ⁷⁾ Items selected for the SCL-3/7; ⁸⁾ Items selected for the SCL-8 (8); ⁹⁾ Items selected for the SCL-9 (9); ¹⁰⁾ Items selected for the SCL-10 (10).

were, respectively, 1.05 and 0.98. All measures were well within the acceptable range of 0.70 and 1.30.

3.2.2 | Differential item functioning

Outpatients showed significant DIF for the items at both ends of the latent trait (item 31, 13 and 25, Appendix S2). Day care patients showed DIF on item 72 and inpatients showed DIF for item 30, 33 and 25. Juvenile patients showed DIF for items 30, 32, 72 and 13. Addicted patients showed DIF on five of the seven items. Thus, we constructed separate individual versions within each subgroup. It turned out that these individual versions all showed lower relative precisions

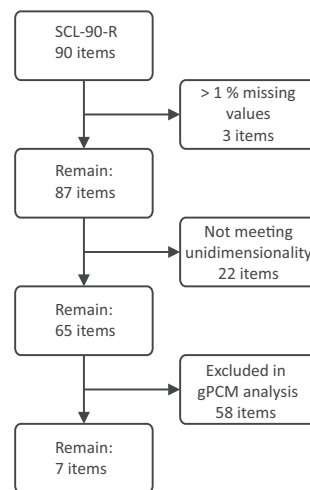


FIGURE 2 Flow chart of item selection

and lower correlations with the criterion than the general solution (Table 4).

3.2.3 | Routing of the SCL-3/7

The first item to be presented to the patients was item 32, followed by 30 or 13, and then 31, 33, 72 or 25 (Figure 1).

3.2.4 | Item weights for calculating the SCL-3/7 score

The optimal weights gained from the iterative procedure are presented in Figure 4. This solution resulted in a correlation of 0.842 with the person measures of the 65-item generalised Partial Credit Model solution.

3.2.5 | Relative precision of the general form and individual solutions

Except for the addicted patient group, the general SCL-3/7 solution had better relative precisions than the individual solutions constructed within the separate patient groups. Thus, the general solution that was based on the total sample of subjects was preferred over the separate solutions. The SCL-3/7 performed worse for juvenile patients and unsatisfactorily for addicted patients.

3.3 | Comparing the short form versions

Four very short forms of the SCL-90-R have been constructed before; the SCL-5,⁷ the SCL-8,⁸ the SCL-9⁹ and the SCL-10.¹⁰ Tambs et al.⁷ describe that for the development

TABLE 3 Items with symptom dimensions in gPCM model sorted by item measure and grouped into seven categories

| Category | Item number | Trait position | Discriminative value | Symptom dimension |
|----------|------------------------|----------------|----------------------|-------------------|
| 1 | 48 | 1,07 | 0,94 | SOM |
| | 25 ⁷⁾ | 1,03 | 1,03 | PHOB |
| | 47 | 0,98 | 0,91 | PHOB |
| | 85 | 0,97 | 0,93 | PSY |
| | 67 | 0,86 | 0,91 | HOS |
| 2 | 15 | 0,66 | 1,05 | DEP |
| | 52 | 0,65 | 0,85 | SOM |
| | 13 ⁷⁾ | 0,61 | 1,06 | PHOB |
| | 73 | 0,56 | 0,90 | I-S |
| 3 | 75 | 0,45 | 0,98 | ANX |
| | 17 | 0,45 | 0,94 | PHOB |
| | 45 | 0,44 | 0,86 | O-C |
| | 80 | 0,40 | 1,04 | ANX |
| | 21 | 0,38 | 0,74 | I-S |
| | 87 | 0,37 | 0,79 | PSY |
| | 53 | 0,34 | 0,85 | SOM |
| | 23 | 0,33 | 1,13 | ANX |
| | 50 | 0,32 | 1,04 | PHOB |
| | 68 | 0,31 | 0,91 | PAR |
| | 49 | 0,28 | 0,83 | SOM |
| | 83 | 0,27 | 0,89 | PAR |
| | 58 | 0,26 | 0,87 | SOM |
| | 72 ⁷⁾⁽⁸⁾ | 0,25 | 1,17 | ANX |
| | 40 | 0,23 | 0,72 | O-C |
| | 38 | 0,23 | 0,75 | SOM |
| | 70 | 0,21 | 1,01 | PHOB |
| 4 | 59 | 0,17 | 1,05 | - |
| | 88 | 0,17 | 0,69 | PSY |
| | 18 | 0,14 | 1,05 | PAR |
| | 86 | 0,13 | 1,09 | PAR |
| | 43 | 0,13 | 1,05 | ANX |
| | 69 | 0,06 | 1,12 | I-S |
| | 37 | 0,02 | 1,06 | I-S |
| | 32 ⁷⁾ | -0,01 | 1,20 | DEP |
| | 76 | -0,03 | 0,84 | PAR |
| | 78 | -0,05 | 0,83 | ANX |
| | 36 | -0,09 | 1,12 | I-S |
| 56 | -0,09 | 0,91 | SOM | |
| 5 | 22 | -0,11 | 1,20 | DEP |
| | 33 ⁵⁾⁽⁷⁾⁽⁸⁾ | -0,14 | 1,22 | ANX |
| | 44 | -0,31 | 0,43 | - |
| | 9 | -0,33 | 0,59 | O-C |
| | 26 | -0,38 | 1,15 | DEP |

(Continues)

TABLE 3 (Continued)

| Category | Item number | Trait position | Discriminative value | Symptom dimension |
|----------------------|----------------------|----------------|----------------------|-------------------|
| 6 | 79 ⁸⁾ | -0,42 | 1,38 | ANX |
| | 61 | -0,42 | 1,11 | I-S |
| | 2 ⁵⁾⁸⁾ | -0,42 | 0,87 | DEP |
| | 90 | -0,44 | 1,20 | PSY |
| | 89 | -0,45 | 1,04 | - |
| | 66 | -0,45 | 0,49 | - |
| | 77 | -0,46 | 1,30 | O-C |
| | 46 | -0,46 | 0,88 | PSY |
| | 71 ⁸⁾ | -0,47 | 1,31 | I-S |
| | 34 | -0,47 | 1,16 | DEP |
| | 28 | -0,48 | 1,08 | O-C |
| | 41 | -0,49 | 1,14 | I-S |
| | 3 | -0,50 | 1,16 | O-C |
| | 51 | -0,51 | 1,25 | O-C |
| | 11 | -0,59 | 0,77 | HOS |
| | 54 ⁵⁾⁸⁾ | -0,60 | 1,29 | DEP |
| | 30 ⁵⁾⁷⁾⁸⁾ | -0,61 | 1,44 | DEP |
| | 29 | -0,66 | 1,16 | DEP |
| | 55 | -0,66 | 1,12 | O-C |
| | 7 | 14 | -0,80 | 0,86 |
| 57 | | -0,91 | 1,21 | ANX |
| 31 ⁵⁾⁷⁾⁸⁾ | | -0,96 | 1,21 | DEP |

Note: ⁵⁾ Items selected for the SCL-5 (7); ⁷⁾ Selected items for the SCL-3/7 (blue shaded); ⁸⁾ Items selected for the SCL-8 (8).

Abbreviations: ANX, anxiety; DEP, depressive symptoms; HOS, hostility; I-S, interpersonal sensitivity; O-C, obsessive-compulsive; PAR, paranoid ideation; PHOB, phobic anxiety; PSY, psychoticism; SOM, somatisation.

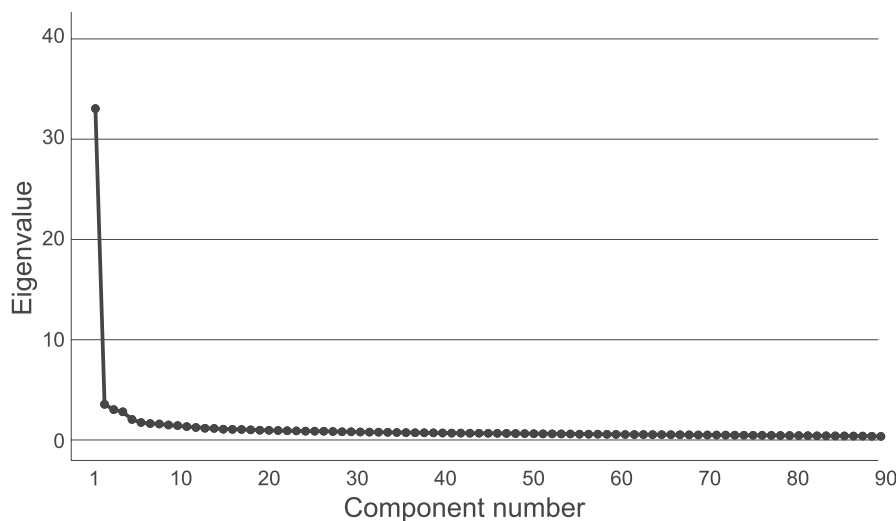


FIGURE 3 Scree plot of Principal Component Analysis

of the SCL-5, they first determined the factor structure of the SCL-25 and stated that a short form should reflect the same factor structure. They found two components, depression and anxiety. The authors selected the items based

on principal component and regression analyses. For the SCL-8, Fink et al.⁸ applied first a factor analysis for excluding items that did not fit in a 1-factor model. Next, a latent trait model was analysed with two parameters for each

TABLE 4 Relative precision and correlations

| | Total | | r | ω | α | Outpatients | | Day care | | Inpatients | | Juvenile | | Addicted | |
|---|-----------|--------------------|-------|----------|----------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|----------|-------|
| | n items | Relative precision | | | | Relative precision | Relative precision | Relative precision | Relative precision | Relative precision | Relative precision | Relative precision | Relative precision | | |
| SCL-90-R classical sumscore | 90 | 100.0 | 1.00 | 0.99 | 0.99 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |
| SCL-5 | 5 | 105.2 | 0.903 | 0.91 | 0.91 | 108.8 | 102.5 | 106.0 | 110.0 | 106.0 | 110.0 | 110.0 | 110.0 | 97.3 | 97.3 |
| SCL-8 | 8 | 105.4 | 0.934 | 0.93 | 0.94 | 106.7 | 103.7 | 106.9 | 107.5 | 106.9 | 107.5 | 107.5 | 107.5 | 93.4 | 93.4 |
| SCL-9 | 9 | 104.1 | 0.947 | 0.90 | 0.90 | 102.9 | 103.9 | 104.4 | 103.7 | 104.4 | 103.7 | 103.7 | 103.7 | 107.2 | 107.2 |
| SCL-10 | 10 | 104.2 | 0.949 | 0.93 | 0.93 | 103.3 | 102.8 | 105.5 | 107.1 | 105.5 | 107.1 | 107.1 | 107.1 | 79.0 | 79.0 |
| GPCM score general solution | 65 | 110.5 | 0.900 | 0.98 | 0.98 | 123.3 | 116.0 | 115.1 | 108.0 | 115.1 | 108.0 | 108.0 | 108.0 | 34.3 | 34.3 |
| SCL-3007 general solution ^a | 3/7 | 110.9 | 0.822 | 0.92 | 0.92 | 122.7 | 111.8 | 108.3 | 103.8 | 108.3 | 103.8 | 103.8 | 103.8 | 23.2 | 23.2 |
| SCL-3007 individual solution ^b | 3/7 | | | | | 114.9 | 95.0 | 100.0 | 96.2 | 100.0 | 96.2 | 96.2 | 96.2 | 26.8 | 26.8 |

Abbreviations: GPCM, generalised Partial Credit Model; r , Pearson's correlation with the SCL-90-R classical sumscore; α , Cronbach's alpha; ω , McDonald's omega.

^aScore constructed within all groups.

^bScore constructed within subgroups.

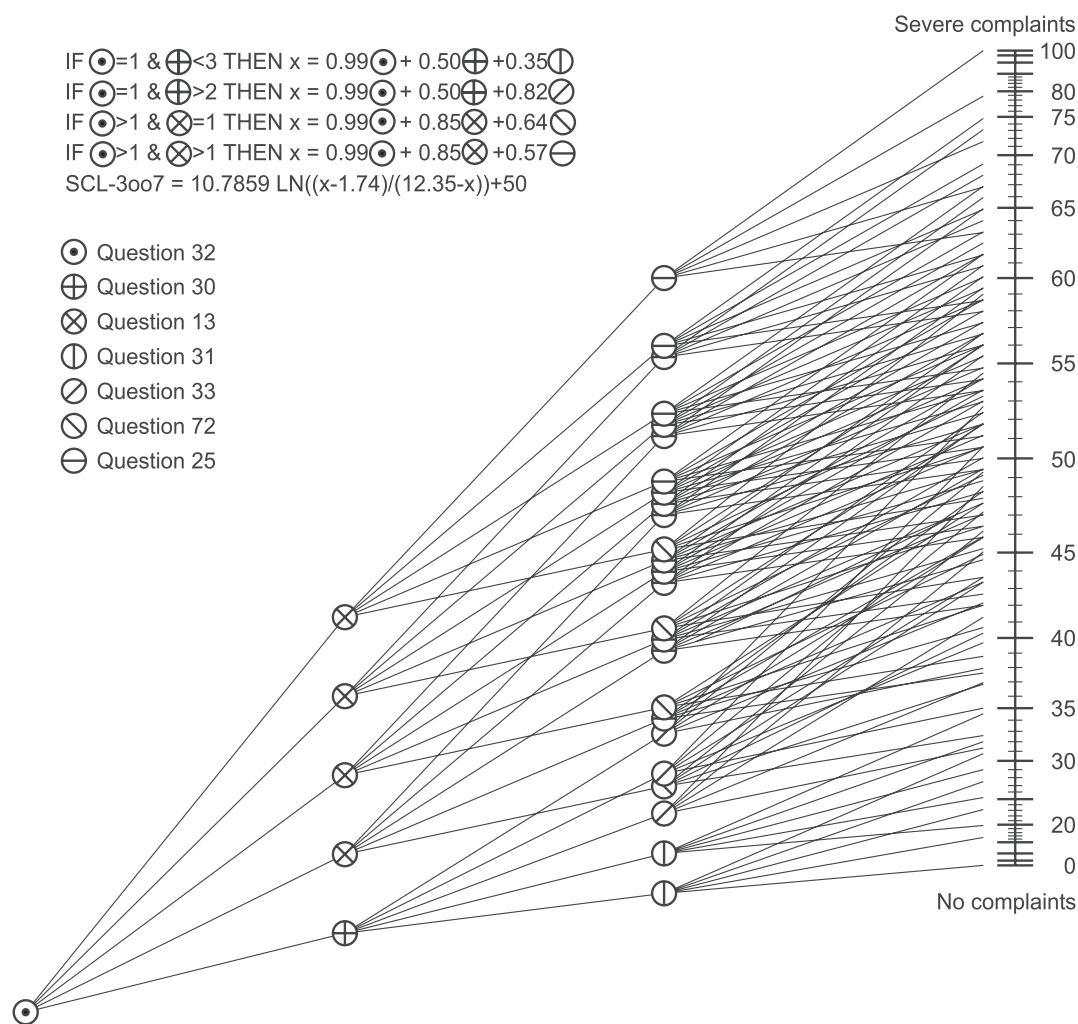


FIGURE 4 Exact presentation of SCL-3/7 calculation. Starting point is the lower left symbol. Then, 0.99 times the response on the first item (32, original SCL-90-R numbering) is added. When this first response is 1, the next question is 30, and 0.50 times the response on item 30 is added. The same procedure holds for the third question. When the response on the first question is larger than 1, the second question is 13, and subsequently 0.85 times the response on this question is added. Then, the third question follows in the same way. Note that the scale is not linear, but a logit scale. The routing is based on medians, therefore question 25 is most frequent as a third question in this figure

item: a threshold figure and a slope. The difference with the gPCM analyses is that the items were dichotomised. For the SCL-9, Petrovsky et al.⁹ selected from each of the 9 dimensions one item with the highest correlation with the total SCL-90 score. They tested their list with Classical Test Theory, confirmatory factor analysis, Cronbach's alpha and correlations with the larger list (SCL-27). Almost all correlations of the SCL-5, SCL-8, SCL-9 and SCL-10 with SCL-90-R total score were well over 0.90 (Table 4). The correlations of the SCL-3/7 were generally lower than 0.90. All McDonald's omegas and Cronbach's alphas were 0.90 or higher.

Within the total group of participants, all relative precisions for the SCL-5 to SCL-10 versions were around 105%. The generalised Partial Credit Model score that was based on 65 items had a higher precision, namely 110.5%. The newly constructed SCL-3/7 had the highest precision

of the SCL-90-R short forms (110.9%). In the outpatient group, the SCL-3/7 had by far the highest precision (122.7%): the other four forms had precisions of 103% to 109%. In day care patients, the SCL-3/7 also had the highest precision (111.8%), as well as in the inpatient group (108.3%). For juvenile patients, the SCL-5 was clearly superior (110.0%), and for addicted patients, the SCL-9 was best (107.2%).

4 | DISCUSSION

This study showed that the SCL-3/7 is successful for routine outcome monitoring in adult patient groups: outpatients, day care patients and inpatients. So, the number of items of the SCL-90-R can be reduced from 90 to three items out of seven to be administered. For juvenile

patients, the SCL-5 is to be preferred. For addicted patients, the SCL-9 is to be preferred.

Three items had too many missing values and 22 items showed component loadings that were not high enough (<0.50) and were therefore excluded. Using IRT analysis, the remaining 65 items were reduced to seven items. A routing reduced the number of questions to be answered to three. The weighted sum of these three items correlated 0.842 with the IRT score of the 65 relevant items. The correlation of the SCL-3/7 with the classical SCL-90-R score was lower than the correlations of the other short forms. This is not surprising because the correlation of the generalised Partial Credit Model and SCL-90-R score was 0.90. Since the SCL-3/7 was based on maximising the correlation with the generalised Partial Credit Model score, the correlation with the SCL-90-R was logically below 0.90.

The relative precision of the SCL-3/7 was higher than the original SCL-90-R sum score based on 90 items for outpatients, day care patients and inpatients. We claim that we have not only reduced the number of items successfully, but also that we have extracted the items with the most relevant context, a characteristic that is of importance for reducing response burden.⁶ However, the relative precision was lower for juvenile patients and much lower for addicted patients. The SCL-3/7 is not recommended for use in these patient groups.

4.1 | Principal findings in relation to the existing literature

In some situations, routine outcome monitoring, or providing feedback on the course of the treatment to therapists, is thought to reduce the number of therapy sessions.³² In this line of reasoning, routine outcome monitoring can be used as an aid for therapists to discontinue treatment when a patient falls below a certain level. However, it must be emphasised that ending a therapy should be the decision of the therapists in agreement with the patient. The course of SCL-3/7 scores may only serve as an indication for ending treatment, or for showing whether therapy passes off successfully.

4.2 | Latent trait

The latent trait underlying the SCL-3/7 is only a part of the complete SCL-90-R. The mild region is covered by three depression items—original items 31, 30 and 32 (Figure 1). The intermediate position is formed by two anxiety items; 33 and 72. The severe problems are determined by two phobic anxiety (agoraphobic) items: 13 and 25. Interestingly, these are the two symptoms that emerged from the recent

factor analyses by Schmalbach et al.³³ Thus, these two dimensions are generally found to be of some importance. The obsessive-compulsive, hostility, interpersonal sensitivity, psychoticism, paranoid ideation and somatisation dimensions are not addressed by the SCL-3/7. Notably, the SCL-5 and SCL-8 items also considered depression and anxiety, but did not include items from the severe, phobic part of the latent trait (Table 3).

4.3 | Limitations

Firstly, we emphasise that the SCL-3/7 and other short forms are not meant for diagnostic and screening purposes. The SCL-3/7 is particularly meant to be used in routine outcome monitoring. It may be administered with paper and pencil (we supply PDFs for that purpose in Appendices S3 and S4), even though many monitoring programs are nowadays web- or computer-based. These programs can easily perform the routing and generate progression graphs, which can be presented to the therapist at the start of the session, immediately after administration of the SCL-3/7.

For addicted and juvenile patients, the relative precisions of the SCL-3/7 were too low. For these groups, the SCL-3/7 is not recommended for application in routine outcome monitoring. For juvenile patients, the SCL-5 performs much better with a relative precision of 110.0%. With a relative precision of 107.2%, the SCL-9 performs the best for addicted patients. A potential reason for this limitation may lie in the fact that the SCL-3/7 does not cover all a priori dimensions of the total SCL-90-R. Juvenile and in particular addicted patients obviously suffer more from symptoms that are not covered by the SCL-3/7.

Differential item functioning was observed for a number of items in the outpatient and inpatient group. This implies that the SCL-3/7 would not be suited to compare the effectiveness of treatments between these groups. For individuals within these treatments, DIF does not represent a serious problem when monitoring the progress of treatment. Considering the fact that the SCL-3/7 version that was based on the overall sample performed better than the versions developed within the various patient groups, there would be no need to use separate versions. It cannot be ruled out that analyses based on the larger number of cases used for obtaining the general (final) version had greater power, and thus produced better results that were also to a lesser extent influenced by sub-sample specific characteristics.

The SCL-3/7 contains items that were originally embedded within the full scale. Berndt³⁴ demonstrated that when an instrument targeted at assessing depressive symptoms was administered under two conditions,

namely items (of a short form) embedded within the full scale versus only the items of the short form, identical methods of factor analysis produced different factor patterns across conditions and that the short form measured different dimensions than the original, long form (i.e. dimensional shift).

4.4 | Future perspective

The present findings were based on Dutch data. Cross-national studies are needed to examine the cross-language generalisability of the present findings. Future studies could also address the effectiveness of the SCL-3/7 in other patient populations, its test-retest reliability, its ability to predict clinically significant change related to diagnosis or impairment, its capability of predicting change on separate independent criterion measures and the issue referred to above of eventual dimensional shift by having taken items out of their original context (i.e. tackle the question as to whether the 7 items of the SCL-3/7 still hang together when administered outside the context of the original 90-item measure).

To conclude, the goal of the present study was firstly to reduce the number of SCL-90-R items to 7 items of which 3 are to be administered by the patient while maximising the distinctive capacity in order to make it suitable for routine patient reported outcome measurement in clinical practice. The maximisation of the distinctive capacity was tested with the relative precision. The second goal was to compare the various short forms of the SCL-90-R. The relative precision of over 100% in the three adult non-addict groups indicates that the SCL-3/7 is potentially more sensitive to change than the complete and other very short forms of the SCL-90-R.

The SCL-3/7 is ideal for computerised systems, but we also provide a PDF and an Excel calculation sheet for paper and pencil administration (Appendix S5). The relevant instrument in English and Dutch is shown in Supplementary Materials.

CONFLICT OF INTEREST

The authors report no conflict of interests.

ACKNOWLEDGEMENTS

We are grateful to the participants in the STEP project, who were treated in Altrecht (Zeist), De Viersprong (Halsteren), Overwaal (Lent), Mentrum (Amsterdam), De Gelderse Roos (Luntenen), GGZ-E (Eindhoven), Symfona (Amersfoort), Mediant (Enschede), CSB Friesland (Leeuwarden) and Triversum (Alkmaar). We are also indebted to the participants who were treated at Novadic-Kentron in Roosendaal and Bergen op Zoom. Special

gratitude is expressed to the late Wim Trijsburg, who was the driving force behind the collection of the STEP data. The support by Dr. Điệp Ngô-Xuân, the Dean of the Faculty of Psychology of Vietnam National University in HCMC, Vietnam, too is gratefully acknowledged.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/acps.13396>.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in SCL-3/7 data at <https://doi.org/10.17026/dans-zt4-hmac>.

ORCID

Reinier Timman  <https://orcid.org/0000-0002-7772-0394>

Willem A. Arrindell  <https://orcid.org/0000-0003-1004-7411>

REFERENCES

1. Derogatis LR. SCL-90: administration, Scoring and Procedures Manual-I for the R(evised) Version. Johns Hopkins University School of Medicine, Clinical Psychometrics Research Unit; 1977.
2. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. *BMJ Qual Saf.* 2014;23:508-518. doi:10.1136/bmjqs-2013-002524
3. Macdonald A, Fugard AJ. Routine mental health outcome measurement in the UK. *Int Rev Psychiatry.* 2015;27(4):306-319. doi:10.3109/09540261.2015.1015505
4. Arrindell WA, Ettema JHM. Symptom Checklist, SCL-90. Handleiding bij een Multidimensionale Psychopathologie-Indicator. Harcourt Test Publishers. 2003:120.
5. Hatfield DR, Ogles BM. The use of outcome measures by psychologists in clinical practice. *Prof Psychol Res Pract.* 2004;35(5):485-491. doi:10.1037/0735-7028.35.5.485
6. Rolstad S, Adler J, Rydén A. Response burden and questionnaire length: is shorter better? a review and meta-analysis. *Value Health.* 2011;14:1101-1108. doi:10.1016/j.jval.2011.06.003
7. Tambs K, Moum T. How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatr Scand.* 1993;87:364-367. doi:10.1111/j.1600-0447.1993.tb03388.x
8. Fink P, Jensen J, Borgquist L, et al. Psychiatric morbidity in primary public health care. A Nordic multicenter investigation: Part I. method and prevalence of psychiatric morbidity. *Acta Psychiatr Scand.* 1995;92(6):409-418. doi:10.1111/j.1600-0447.1995.tb09605.x
9. Petrowski K, Schmalbach B, Kliem S, Hinze A, Elmar B. Symptom-Checklist-K-9: norm values and factorial structure in a representative German sample. *PLoS One.* 2019;14(4):e0213490. doi:10.1371/journal.pone.0213490
10. Strand BH, Dalgard OS, Tambs K, Rognerud M. Measuring the mental health status of the Norwegian population: a comparison

- of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). *Nord J Psychiatry*. 2003;57(2):113-118. doi:10.1080/08039480310000932
11. Timman R. Standaard Evaluatie Project II: 2000-2008. Erasmus MC. Updated 01-03-2009. <http://repub.eur.nl/pub/16960>. Accessed 29 January 2022.
 12. Derogatis LR. SCL-90-R: Administration Scoring and Procedures Manual II. Clinical Psychometrics Research Unit; 1983.
 13. Derogatis LR, Savitz KL. The SCL-90-R and the Brief Symptom Inventory (BSI) in primary care. In: Maruish ME, ed. *Handbook of Psychological assessment in Primary Care Settings*. Routledge; 2000:297-334. doi:10.4324/9781315827346
 14. Visser MS, Timman R, Nijmeijer KJ, Lemij HG, Kiliç E, Busschbach JJ. A very short version of the Visual Function Questionnaire (VFQ-3007) for use as a routinely applied Patient Reported Outcome Measure. *Acta Ophthalmol*. 2020;98(6):618-626. doi:10.1111/aos.14378
 15. Arrindell WA, Urbán R, Carrozzino D, Bech P, Demetrovics Z, Roozen HG. SCL-90-R emotional distress ratings in substance use and impulse control disorders: one-factor, oblique first-order, higher-order, and bi-factor models compared. *Psychiatr Res*. 2017;255:173-185. doi:10.1016/j.psychres.2017.05.019
 16. Chen I-H, Lin C-Y, Zheng X, Griffiths MD. Assessing mental health for China's police: psychometric features of the Self-Rating Depression Scale and Symptom Checklist 90-Revised. *Int J Environ Res Public Health*. 2020;17:2737. doi:10.3390/ijerph17082737
 17. Preti A, Carta MG, Petretto DR. Factor structure models of the SCL-90-R: replicability across community samples of adolescents. *Psychiatr Res*. 2019;272:491-498. doi:10.1016/j.psychres.2018.12.146
 18. Cohen J. A power primer. *Psychol Bull*. 1992;112(1):155-159. doi:10.1037//0033-2909.112.1.155
 19. Embretson SE, Reise SP. *Item Response Theory for Psychologists*. Erlbaum Assoc; 2000:384. doi:10.4324/9781410605269
 20. Levine SZ, Rabinowitz J, Rizopoulos D. Recommendations to improve the Positive and Negative Syndrome Scale (PANSS) based on item response theory. *Psychiatry Res*. 2011;188(3):446-452.
 21. Wilson JE, Niu K, Nicolson SE, Levine SZ, Heckers S. The diagnostic criteria and structure of catatonia. *Schizophr Res*. 2015;164(1-3):256-262.
 22. Velthorst E, Levine SZ, Henquet C, et al. To cut a short test even shorter: Reliability and validity of a brief assessment of intellectual ability in Schizophrenia—a control-case family study. *Cognitive Neuropsychiatry*. 2013;18(6):574-593. doi:10.1080/13546805.2012.731390
 23. Levine SZ, Leucht S. Psychometric analysis in support of shortening the scale for the assessment of negative symptoms. *Eur Neuropsychopharmacol*. 2013;23(9):1051-1056.
 24. Mallinson T. Rasch analysis of repeated measures. *Rasch Meas Trans*. 2011;25:1317.
 25. Wright BD, Linacre JM, Gustafson J, Martin-Löf P. Reasonable meansquare fit values. <https://rasch.org/rmt/rmt83b.htm>. Accessed 29 January 2022.
 26. Tristan A. An adjustment for sample size in DIF analysis. <https://www.rasch.org/rmt/rmt203e.htm>. Accessed 29 January 2022.
 27. McDonald RP. *Theory: A Unified Treatment*. Taylor & Francis; 1999. doi:10.4324/9781410601087
 28. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334. doi:10.1007/bf02310555
 29. McHorney CA, Ware JE, Rogers W, Raczek AE, Lu JFR. The validity and relative precision of MOS short-, and long- form health status scales and dartmouth COOP Charts. *Med Care*. 1992;30(Supplement):MS253-MS265. doi:10.1097/00005650-199205001-00025
 30. McHorney CA, Haley SM, Ware JEJ. Evaluation of the MOS SF-36 Physical Functioning Scale (PF-10): II. Comparison of relative precision using Likert and Rasch scoring methods. *J Clin Epidemiol*. 1997;50(4):451-461. doi:10.1016/s0895-4356(96)00424-6
 31. Psych: procedures for psychological, psychometric, and personality research. <https://rdrr.io/cran/psych/>. Accessed 13 December, 2021;32.
 32. Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: are outcomes enhanced? *Psychother Res*. 2001;11(1):49-67. doi:10.1080/713663852
 33. Schmalbach B, Zenger M, Tibubos AN, Kliem S, Petrowski K, Brähler E. Psychometric properties of two brief versions of the hopkins symptom checklist: HSCL-5 and HSCL-10. *Assessment*. 2021;28(2):617-631. doi:10.1177/1073191119860910
 34. Berndt DJ. Taking items out of context: dimensional shifts with the short form of the Beck Depression Inventory. *Psychol Rep*. 1979;45(2):569-570. doi:10.2466/pr0.1979.45.2.569

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Timman R, Arrindell WA. A very short Symptom Checklist-90-R version for routine outcome monitoring in psychotherapy; The SCL-3/7. *Acta Psychiatr Scand*. 2022;145:397–411. doi:[10.1111/acps.13396](https://doi.org/10.1111/acps.13396)