Chapter 4

Scalability, Reliability and Validity of the Benzodiazepine Dependence Self-Report Questionnaire in Outpatient Benzodiazepine Users

**C.C. Kan[1], M.H.M. Breteler[2], E.A.Y. Timmermans[1], A.H.G.S. van der Ven[3]
and F.G. Zitman[1]**

[1]Department of Psychiatry*, University of Nijmegen,

PO Box 9101, 6500 HB Nijmegen, The Netherlands

[2]Department of Clinical Psychology*, University of Nijmegen,

PO Box 9104, 6500 HE Nijmegen, The Netherlands

[3]Department of Mathematical Psychology, University of Nijmegen,

PO Box 9104, 6500 HE Nijmegen, The Netherlands

*University of Nijmegen Research Group on Addictive Behaviours (UNRAB)

ABSTRACT

As there is no multidimensional instrument available that reflects the severity of benzodiazepine (BZD) dependence comprehensively, the *Benzodiazepine Dependence Self-Report Questionnaire (Bendep-SRQ)* was developed and investigated.

The Bendep-SRQ, Symptom Checlist-90(SCL-90), Schedules for Clinical Assessments in Neuropsychiatry (SCAN) and Addiction Severity Index-Revised (ASI-R) were administered to115 general practice patients, 124 psychiatric outpatients and 33 self-help patients who were using BZDs. Factor and Rasch analyses were applied to construct scales. Reliability assessments were made in terms of subject discriminability, item discriminability and test-retest stability. To support the construct validity of the scales, theoretical rationales were required to explain the specific item order provided by the Rasch scale values. To assess the concurrent and discriminant validity, a matrix consisting of the above-mentioned measures was factor-analysed.

Four Rasch-homogeneous scales were delineated: 'Problematic Use', 'Preoccupation', 'Lack of Compliance' and 'Withdrawal'. Nearly all subject discriminability, item discriminability, and and test-retest results indicated good reliability. A BZD dependence factor was extracted with high loadings for the Bendep-SRQ scales and the concurrent measures. The discriminant measures had high loadings on other factors.

The scalability, reliability and validity of the Bendep-SRQ scales appeared to be good. The Bendep-SRQ shows great promise as a useful and easily manageable instrument for the assessment of the severity of Benzodiazepine Dependence in clinical practice and scientific research.

INTRODUCTION

Since their introduction in the early sixties, the dependence liability of benzodiazepines (BZDs) has been debated in the medical literature.[1-4] The main problem in this debate appears to be the lack of consensus about the definition of benzodiazepine (BZD) dependence.[5]

Although the WHO expert committee on dependence-producing drugs defined psychological dependence in addition to physical dependence in 1964,[6] withdrawal and tolerance remained to be considered as the cardinal elements of BZD dependence in most reports and the DSM-III classification.[7] Consequently, at least three Self-Report Questionnaires (SRQs) have been developed with the aim to reflect the severity of BZD withdrawal.[8-10] However, the validity of these questionnaires was only assessed on an item level.[9,10]

Since 1981, the WHO has been propagating a psycho-physiological-social model for dependence on all psychoactive substances, called the drug dependence syndrome.[11,12] Both the DSM-III-R and DSM-IV[13,14] and the ICD-10[15] are primarily based on this general syndrome, which was originally proposed for alcohol.[16] The substance dependence criteria of these classifications have been operationalized in structured diagnostic interviews, such as the *Composite International Diagnostic Interview (CIDI)*[17,18] and the *Schedules for Clinical Assessments in Neuropsychiatry (SCAN)*.[18,19] Using the SCAN, the DSM-III-R and ICD-10 BZD dependence constructs were recently assessed epidemiologically and psychometrically.[20,21] The Rasch-homogeneous BZD dependence scales which were derived from the SCAN data,[21] consisting of subsets of the DSM-III-R and ICD-10 criteria, can be considered to probe BZD dependence at its general core. They do not reflect BZD dependence comprehensively, because all the criteria related to BZD withdrawal had to be

67

omitted, potentially reflecting a separate dimension that could not be analysed further.

Recently, the *Benzodiazepine Dependence Questionnaire (BDEPQ)* was developed by Baillie and Mattick (1996)[22] with the aim to reflect the severity of BZD dependence comprehensively. Nonetheless, the BDEPQ total sumscore was considered to be a proper scale, although the three factors yielded by factor analysis - labelled 'general dependence', 'pleasant effects' and 'perceived need'- suggested a multidimensional structure. Furthermore, a withdrawal scale was not included and the validity results were modest.

The present study describes the development of the *Benzodiazepine Dependence Self-Report Questionnaire (Bendep-SRQ)* and the assessment of its scalability, reliability and validity in a representative patient sample consisting of general practice (GP) patients, psychiatric outpatients and self-help patients. In contrast to the above-mentioned studies, the approach was multidimensional and included a withdrawal dimension. Rasch modelling was applied to refine the dimensions into proper scales.

METHOD

Settings and Subjects

This study was conducted at four general practices, three psychiatric outpatient departments and two self-help groups concerned with addictive medication use. To participate in the investigation, the subjects had to meet the following inclusion criteria:
1) actual BZD use; 2) average frequency of BZD use of at least once a week; 3) age between 17 and 70 years; 4) ability to speak and read Dutch.
The patients who visited the general practices, psychiatric outpatient departments or self-help

meetings during the period of investigation, or who had individual contact with a self-help team member, were screened according to these inclusion criteria. Eligible patients were asked to participate by a representative of the treatment or self-help team. Informed consent was obtained from the majority of selected subjects: 67% (115 out of the 172) from general practice, 70% (124 out of the 178) from psychiatric outpatient departments and 70% (33 out of the 47) from self-help groups. The total sample of participants consisted of 272 subjects.

Study Design

This study formed part of a larger project conducted by the University of Nijmegen Research Group on Addictive Behaviours (UNRAB) in The Netherlands on the detection and diagnosis of BZD dependence. The study population participated in two interviews, separated by three weeks. During the first interview, sociodemographic data were gathered and the following questionnaires were administered: Bendep-SRQ, L-scale of the Minnesota Multiphasic Personality Inventory (MMPI-2), Benzodiazepine Dependence Structured Diagnostic Interview (Bendep-SDI) and the Schedules for Clinical Assessments in Neuropsychiatry (SCAN).[23] The Bendep-SRQ and Bendep-SDI have been constructed by our research group. The second interview, which was conducted by the same interviewer as the first, consisted of a second administration of the Bendep-SRQ followed by the Symptom Checklist-90 (SCL-90)[24] and the Addiction Severity Index-Revised (ASI-R).[25]

Bendep-SRQ

Item Formulation

Specific criteria for BZD Dependence, summarized by Linsen et al.,[5] reflecting psychological, physiological and social aspects of BZD dependence, were formulated at the Department of Psychiatry of the University Hospital of Nijmegen in The Netherlands. This set of criteria served as the main source of Bendep-SRQ items. To ensure a good face-validity, these criteria were judged by a number of international experts and adapted to meet most of their critical comments. A minority of the Bendep-SRQ items were added more intuitively on the basis of clinical experience and judgment. From the premises of the directives of Cook and Campbell,[26] 10 independent psychiatrists, 10 general practitioners and 10 former self-help patients were subsequently asked to score the initial set of Bendep-SRQ items on comprehensibility, unambiguity and recognition, using four-point scales. A number of items were subsequently altered or removed. The remaining items were used to construct the Bendep-SRQ.

Format

In the first brief section of the Bendep-SRQ, the person who administers the questionnaire notes the sex and age of the respondent, the average BZD dose(s) used and the duration of BZD use. In the second and third sections, five-point items are rated by the respondent according to the degree in which they apply to him or her.

Scale Construction

Principal component analysis and principal axis factor analysis with varimax rotation were used to explore the data. Interpretable factors were analysed separately at a 5% significance level by Maximum Likelihood Factor Analysis to test the goodness of fit of a single factor model in order to confirm or reject unidimensionality. This yielded four unidimensional factors, which formed the basis of the present scales of the Bendep-SRQ. As it appeared that a normal distribution could not be assumed for all the items of the scales, it was decided to dichotomize the five-point items to perform Rasch analysis, which does not require normality. All items on each of the scales were dichotomized between 'this is not true for me' (option 2) and 'this is partly true, partly false for me' (option 3).

Scalability

In using the Bendep-SRQ scales, i.e. the sum scores of the dichotomized item responses, certain assumptions are implicitly made, which are specified in the Rasch scaling model. To justify the use of the sum scores, these assumptions must be tested, which implies that the Rasch model should hold true. The assumptions from which the Rasch model can be derived are given below.

According to Fischer,[27] these assumptions are:

(1) *Unidimensionality.* All items are functionally dependent upon only one underlying continuum, u.

(2) *Monotonicity.* All item characteristic functions are strictly monotonic in the latent trait, u. The item characteristic function describes the probability of a predefined response as a function of the latent trait, u.

(3) *Local stochastic independence.* Every person has a certain probability of giving a

71

predefined response to each item and this probability is independent of the answers given to the preceding items.

(4) *Sufficiency of a simple sum statistic.* The number of predefined responses is a sufficient statistic for the latent parameter u.

(5) *Dichotomy of the items.* For each item there are only two different responses, for example, positive and negative.

The Rasch model requires that an additive structure underlies the observed data. This additive structure applies to the logit of p_{ij} , where p_{ij} is the probability that subject i will give a predefined response to item j, being the sum of a subject scale value u_i and an item scale value v_j ,

i.e.: $ln(p_{ij}/1- p_{ij}) = u_i + v_j$

While the item responses depend on the respective underlying probabilities in a random way, the response probabilities themselves depend in a deterministic way on the subject and item scale values.

The above-mentioned additive structure, which is required by the Rasch model, implies that both subjects and items can be arrayed on a common unidimensional scale and that the items have equal discriminative power (i.e. the property of equi-discriminability). Glas[28] has developed two statistical tests for the dichotomous Rasch model, which are known as R1 and R2. The statistic R1 is especially sensitive to equi-discriminability, while the statistic R2 is sensitive to unidimensionality and local stochastic independence. If R1 is not significant at a 1% significance level ($P > 0.01$) the null hypothesis that all the items have equal discriminative power cannot be rejected and equi-discriminability can be assumed. Similarly, unidimensionality and local stochastic independence hold true when R2 is not significant ($P > 0.01$). Rasch-homogeneity is demonstrated if both statistics hold true, meaning that the sum

72

score across items is a sufficient statistic for the subject scale and that the sum score across subjects is a sufficient statistic for the underlying item scale. To compute R1 and R2 for our Bendep-SRQ data the Rasch Scaling Program (RSP)[29,30] was used.

Reliability

To estimate the reliability of the Bendep-SRQ scales, subject discriminability, item discriminability and test stability were assessed.

Subject discriminability (Internal Consistency). Subject discriminability implies that the subjects should differ systematically, i.e., the variation between subjects should be larger than the variation due to random error. The subject discriminability of the Bendep-SRQ scales was evaluated by the KR-20 coefficient. The size of KR-20 reflects the reliability of the scale, as the error variance of the estimator decreases if KR-20 increases.

Item discriminability. This should not be confused with the above-mentioned term equi-discriminability. It implies that the items should differ systematically, i.e., the variation between items should be larger than the variation due to random error. This was tested by Cochran's Q test.[31] If the test result is significant, items can be considered to occupy distinct points on the scale. Additionally, analogous to the concept of reliability, which is a measure of intersubject discriminability, a measure of interitem discriminability was developed, the item discriminability coefficient (IDC). As this measure has not yet been described in literature, its mathematical foundation is given below.

If the subjects-by-items data matrix is regarded as a two-way factorial design for analysis of variance without replications, then according to Hoyt,[32] the reliability coefficient $\Psi(T_j,T_j')$ of the subject sum score may then be defined as: $\Psi(T_j,T_j') = [\ ^2_S - [\ ^2_{Res} / [\ ^2_S$ Eq 1

In the case of binary data, Hoyt's formula results in the KR-20. Analogously to Eq 1, one

73

might define the reliability coefficient $\Psi(T_i, T_i')$ of the item sum score (or equivalently, the

item *mean* score) as: $\qquad \Psi(T_i, T_i') = [\,\sigma^2_I - \sigma^2_{Res} \,/\, \sigma^2_I$ $\hspace{2cm}$ Eq

$$2$$

In these equations σ^2_S refers to the mean square for the variation among subjects, σ^2_I refers to

the mean square for the variation among items and σ^2_{Res} refers to the mean square for the

residual variation.

The term 'item discriminability' is used here instead of 'the reliability of the item sum score

or item mean score' to avoid this complex terminology. The IDC shows to what extent the

differences between items are systematic. The higher the IDC, the more powerful the

predictions about the item scale. Naturally, the formulas for subject and item discriminability

may only be applied when the underlying item response model holds true.

Stability. To assess the test-retest reliability of the Bendep-SRQ scales, Pearson Product-

Moment correlation coefficients were computed from the Bendep-SRQ data obtained from

the first and the second interviews. The subjects who discontinued their BZD use in the

period between the interview sessions were excluded from the analysis**.**

Validity

The validity of the Bendep-SRQ scales was assessed in terms of construct, concurrent and

discriminant validity.

Construct Validity. So far, the above-methodology does not specify the constructs of the

scales. If all the scalability and reliability requirements are fulfilled, a possible procedure to

establish the construct validity of a scale is to develop a theoretical rationale. The latter

should explain the specific item-order of a Rasch homogeneous set of items, which is

provided by the Rasch scale values of the items and reflects an increasing severity of the

measured construct. This theoretical rationale should determine the actual responses given by the subjects. Therefore, it should only match with one specific item-order in contrast to any other order, to provide a more profound theoretical understanding of the underlying aspect of BZD dependence. The quality of this match of the rationales described in this study reflects our theoretical understanding of the processes that determine the severity of BZD dependence.

Concurrent and Discriminant Validity. To investigate the concurrent and discriminant validity of the Bendep-SRQ we conducted principal axis factor analyses on the data matrix of the subjects who completed both interviews (n = 265). The data matrix consisted of the sum scores of the Bendep-SRQ scales except for 'withdrawal' (to avoid the selection of patients with withdrawal experience only), the SCL-90 subscales, the L-scale of the MMPI-2, the ASI-R problem severity scores, and either the Rasch-homogeneous ICD-10 or DSM-III-R BZD Dependence scale. The latter two scales consist of core sets of substance dependence items of the SCAN, as described in a separate report.[21] The concurrent validity of Bendep-SRQ would be supported if the Bendep-SRQ scales, the ICD-10 or DSM-III-R BZD dependence scale and the ASI problem severity score for drug use load substantially on one common factor, which could be interpreted as a BZD dependence factor. If the sum scores of the SCL-90 subscales, the L-scale of the MMPI and the remaining ASI problem severity scores are found to load substantially on different factors, this would support the discriminant validity of the Bendep-SRQ scales.

RESULTS

Sociodemographic features and pattern of BZD use

The sex ratios and the mean values for age, BZD dose and duration of BZD use of the total sample and subsamples are shown in Table 1. Women were overrepresented in all subsamples, but especially in the self-help sample. The latter was partly caused by the fact that one of the self-help groups was restricted to women only. The sociodemographic features were described in more detail in a separate report.[20]

Scalability

The test results for Rasch analyses are shown in Table 2. The R1 and R2 statistics were non-significant for all Bendep-SRQ scales. These results demonstrate that all four Bendep-SRQ scales, which appeared to reflect 'Problematic Use', 'Preoccupation', 'Lack of Compliance' and 'Withdrawal', convincingly met the requirements of the Rasch model. Therefore, they can be considered proper Rasch-homogeneous scales.

Table 1. Sex ratios and mean values for age, BZD dose and duration of BZD use

Variables	GP* patients (n=115)	Psychiatric Out-patients (n=124)	Self-Help Patients (n=33)	Total Sample (n=272
Sex				
male (%)	30	42	15	33
female (%)	70	58	85	67
Mean age (years)"sd	50±13	47±12	44±11	48±12
MDD/DDD˜	0.9	1.2	1.3	1.1
Quartiles	.3 - .5 - 1.0	.5 - .9 - 1.5	.5 - 1.0 - 2.0	.4 - .7 - 1.2
Duration of BZD# use (months)				
Mean	88	40	103	68
Quartiles	9 - 48 -120	6 - 13 - 42	20 - 90 - 152	8 - 30 - 96

*GP   : General Practice
˜MDD/DDD : Mean Daily BZD Dose/Defined Daily BZD Dose
#BZD   : if more than 1 BZD was being used, the duration was based on the BZD which had been used the longest

Table 2: Test results of Rasch analysis on Bendep-SRQ scales by RSP

Bendep-SRQ Scale	i	R1	df	p	g	R2	df	p	n
Problematic Use	5	12.18	8	.14	3	20.07	8	.01	183
Preoccupation	5	5.93	8	.66	3	16.82	8	.03	199
Lack of Compliance	5	1.67	4	.80	2	8.11	8	.42	109
Withdrawal	5	1.23	4	.87	2	6.40	8	.60	113

RSP: Rasch Scaling Program
R1 : test statistic of Rasch analysis with regard to equi-discriminability
R2 : test statistic of Rasch analysis with regard to unidimensionality and local stochastic independence
i   : number of items in the scale
df : degrees of freedom
p   : p-value
g   : number of subgroups formed by Rasch analysis
n   : number of subjects left in the analysis

Reliability

   Subject Discriminability (Internal Consistency). All reliability measures of the Bendep-SRQ

scales are shown in Table 3. Given the dichotomous nature of the items and the limited length

of the scales (5 items), the KR-20 values were within an acceptable range (0.60-0.88) in all

groups. The total sample yielded values of greater than or equal to 0.70 on all four scales. For

all of the groups, the reliability of the scales was comparable; only the 'Withdrawal' scale,

which was not completed by patients who had never reduced or discontinued their BZD use,

generally yielded high KR-20 values and therefore appeared to be the most reliable scale. In

terms of discriminability, the KR-20 results show that the Bendep-SRQ scales have

substantial differentiating power between subjects.

   Item discriminability. Good item discriminability was indicated for all of the subscales in

the GP patients, the psychiatric outpatients and the total sample by the statistically significant

results of Cochran's Q test and high IDC values (Table 3). In contrast to the other two scales,

non-significant Cochran's Q test results and low IDC values (0.23 and 0.48) were encountered

for 'Preoccupation' and 'Withdrawal' in the group of self-help patients. These findings

appeared to be due not to the small sample size (n = 33), but to an extreme response set in

this highly selected subgroup. Therefore, the item discriminability of the 'Preoccupation' and

'Withdrawal' scales must be considered insufficient in this specific self-help sample.

However, in general, good item discriminability can be assumed for all of the Bendep-SRQ

scales.

   Stability (Test-Retest Reliability). In all groups, we found acceptable test-retest correlations,

ranging from 0.63 to 0.88 (Table 3), which agreed well with the KR-20 reliability values and

provide further support for the reliability of the Bendep-SRQ scales.

Table 3. Reliability of the Bendep-SRQ scales in terms of subject discriminability, item discriminability and test stability

Parameter	I. Problematic Use				II. Preoccupation				III. Lack of Compliance				IV. Withdrawal[#]			
	GP	PO	SH	TS	GP	PO	SH	TS	GP	PO	SH	TS	GP	PO	SH	TS
KR 20	.66	.66	.63	.70	.62	.68	.79	.68	.61	.69	.78	.70	.88	.83	.60	.85
CQ	73	29	18	113	130	142	5	262	30	30	30	83	12	32	7	46
p	<.001	<.001	.001	<.001	<.001	<.001	.281	<.001	<.001	<.001	<.001	<.001	.018	<.001	.116	<.001
IDC	.95	.87	.80	.97	.98	.98	.23	.99	.88	.87	.89	.95	.67	.86	.48	.91
TRT[$]	.74	.64	.68	.71	.63	.71	.86	.71	.66	.70	.88	.74	.69	.76	.78	.75
p	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001	<.001

GP	: General Practice patients (n = 115)
PO	: Psychiatric Outpatients (n = 124)
SH	: Self-Help patients (n = 33)
TS	: Total Sample (n = 272)
[#]	: Respondents who had never reduced or discontinued BZD use were excluded; GP: n = 87, PO: n = 101, SH: n = 30, TS: n = 218
KR-20	: Kuder-Richardson-20 coefficient of internal consistency
CQ	: Cochran's Q (in rounded figures)
IDC	: Item Discriminability Coefficient
TRT	: Test-Retest Correlation
[$]	: Subjects who had discontinued BZD use before the retest session were excluded; GP: n = 102, PO: n = 114, SH: n = 33, TS: n = 249

Validity

Construct Validity. The Rasch-homogeneous Bendep-SRQ scales are shown in Table 4. At the end of each scale, a *theoretical rationale* is given that reflects the item-order. This item-order is determined by the increasing scale values of the Rasch analyses (in the downward direction of Table 4) reflecting increasing severity of the concept measured by the scale. For example, in the first scale, which we named 'problematic use', all items reflect a certain 'degree of awareness of problematic BZD use'. This degree is low if there is only a positive response on 'I have been thinking about giving up the medication' and gradually increases to its highest level if there is a positive response to all the items of the scale, including 'the medication is getting me into trouble'. While the scales 'Preoccupation' and 'Withdrawal' comprise the more classical aspects of dependence, the 'Lack of Compliance' scale is novel because it is entirely based on the medical context of BZD use, which is absent in the case of alcohol and illicit drug use.

Concurrent Validity. Results of the principal axis factor analysis with varimax rotation are shown in Table 5. A four-factor solution was recommended by the scree plot, showing a substantial decrease and a gradual further decline in the eigenvalues of additional factors. Of course, the below-mentioned interpretation of this factor solution is not necessarily the most appropriate, but it appeared to be the most plausible. The highest loadings of the Bendep-SRQ scales, the Rasch homogeneous ICD-10 or DSM-III-R BZD dependence scale and the ASI-R problem severity score for drug use were observed on the second factor that was extracted, which therefore appeared to reflect a dimension of BZD dependence. Moreover, these scales did not show any loadings of greater than .3 on any other factor. These findings suggest good concurrent validity of the Bendep-SRQ scales.

Table 4. Rasch-homogeneous Bendep-SRQ scales and rationales which reflect the item-order provided by the Rasch scale values

Problematic Use (B)[SE(B)]	Preoccupation (B)[SE(B)]	Lack of Compliance (B)[SE(B)]	Withdrawal (B)
I have been thinking about giving up the medication* (-1.41)[.15]	I take another dose of medication on time, because otherwise I would suffer complaints (-1.63)[.18]	I take more medication than is written on the label (-1.12)[.19]	restlessness (-.88)[.19]
At present, the medication is less effective than it used to be (.01)[.14]	I feel safe when I have my medication with me (-.97)[.16]	My medication is gone too quickly (-.72)[.19]	feeling depressed (-.29)[.18]
Other people have urged me to use less medication (.11)[.14]	I get nervous if my medication is out of reach (-.40)[.15]	I go and get a new prescription before the appointed time (-.49)[.19]	tiredness (-.21)[.18]
I think the medication is destroying my life (.64)[.15]	Just before I take my medication, that is the only thing I can think about (1.47)[.16]	I take a lot of medication in one go (.18)[.20]	irritability (0.38)[.18]
The medication is getting me into trouble (.66)[.15]	I spend a great deal of time thinking about medication (1.53)[.16]	I alter what is written on the prescription (2.15)[.33]	shaking (1.0)[.18]

Rationale:

Degree of awareness of problematic BZD use	*Degree of preoccupation/obsession with respect to the availability of BZDs*	*Degree of lack of compliance with the therapeutic BZD regimen*	*Degree of unambiguity of experienced BZD withdrawal*

(B)   : Rasch scale value estimate
[SE(B)]   : Standard Error of the Rasch scale value estimate
*   : Respondents were instructed to substitute their specific BZD(s) for 'medication'

Table 5. Principal Axis Factor Analyses with Varimax Rotation on a matrix (n = 265) consisting of scale scores

Scale	Factor I	II	III	IV
Bendep-SRQ				
Problematic Use		.66 (.70)		
Preoccupation		.62 (.56)		
Lack of Compliance		.66 (.72)		
SCAN				
ICD-10* (DSM-III-R*) past year BZD dependence		.68 (.71)		
SCL-90				
Anxiety	.83 (.83)			
Agoraphobia	.65 (.66)			
Distrust and Interpersonal Sensitivity	.79 (.79)		.37 (.35)	
Depression	.80 (.79)		.34 (.34)	
Insufficiency in thinking and acting	.75 (.75)			
Hostility	.63 (.62)		.30(<.30)	
Sleeping problems	.48 (.48)			.33 (.34)
Somatization	.71 (.71)			.46 (.46)
Remaining items	.81 (.81)		.31 (.31)	
ASI-R problem severity areas				
Physical				.55 (.57)
Professional			.45 (.44)	
Alcohol			.36 (.36)	
Drugs		.44 (.41)		
Social			.62 (.62)	.30(<.30)
Psychiatric	.34 (.33)		.61 (.62)	.37 (.36)
MMPI-2				
L-scale			-.36 (-.34)	

NOTE. Loadings between parentheses refer to the analysis using the DSM-III-R instead of the ICD-10 scale. Factor loadings of smaller than .3 are not shown.
*Rasch-homogeneous subset of BZD dependence criteria (Kan et al.[21])

Discriminant Validity. The distribution of the factor loadings of the remaining scales supported good discriminant validity of the Bendep-SRQ scales. The first factor appeared to be a dimension of the psychological status, as all the SCL-90 subscales showed the highest loadings on this factor. The highest loadings of the ASI-R severity scores on the areas of alcohol, professional, social and psychiatric problems were found on the third factor, while the problem severity for physical health showed the highest loading on the fourth factor. The third and fourth factors were therefore interpreted as addiction-related dimensions of psychosocial problem severity and physical problem severity, respectively. The L-scale of the MMPI-2 showed the highest negative loading on the third factor, indicating that this psychosocial area may be sensitive to socially desirable responses. Insofar as the secondary loadings were greater than .3 (Table 5), these all agreed well within the above-mentioned interpretation of the factors.

DISCUSSION

In this study, four Rasch-homogeneous scales were extracted from the Bendep-SRQ, reflecting the BZD dependence dimensions 'Problematic Use', 'Preoccupation', 'Lack of Compliance' and 'Withdrawal'. These dimensions appeared to be meaningful aspects related to difficulties in controlling, reducing or ceasing BZD use. The 'Lack of Compliance' scale reflected a new type of dimension based on the medical context of BZD use.

In contrast with most of the former studies on BZD dependence measures, the scalability of the Bendep-SRQ scales was clearly assessed. All four Bendep-SRQ scales were shown to meet the strict scalability requirements of the Rasch model, which implies that the sum scores

are sufficient statistics of the underlying dimensions. So far, this method to assess scalability was only applied to the DSM-III-R and ICD-10 BZD dependence constructs.[21] However, to obtain Rasch-homogeneous DSM-III-R and ICD-10 scales, the BZD withdrawal criteria had to be excluded. In the studies on the BZD Withdrawal questionnaires,[8-10] no attention was given to the scalability of the presented scales. Baillie and Mattick[22] also neglected this issue by using the total sum score of the BDEPQ in their analyses, although principal components factor analysis suggested three potential scales.

The reliability results for the four Bendep-SRQ scales proved to be generally good (Table 3). Considering the limited length of the scales (five items), some improvement in subject discriminability might be accomplished in the future by adding new items to the scales. Such items should be formulated in line with the presented theoretical rationales of the Bendep-SRQ scales (Table 4). Reliability assessment in terms of subject and item discriminability was also applied to the revised Rasch-homogeneous DSM-III-R and ICD-10 BZD dependence scales with satisfactory results.[21] So far, no reliability assessments have been made on the BZD withdrawal questionnaires.[8-10] With respect to the BDEPQ, very good reliability results were reported by Baillie and Mattick[22]. However, these should be interpreted with caution, for a reliability assessment implicitly presumes a proper scale. The meaning of the high values of Cronbach's alpha and test-retest correlation coefficient of the total BDEPQ sum score remains unclear as long as scalability analyses of all potential BDEPQ scales are lacking.

The concurrent and discriminant validity of the Bendep-SRQ scales were supported by the results of principal axis factor analysis (Table 5). On an item level, limited validity assessment was performed on two BZD withdrawal scales; Busto et al.[9] assessed the withdrawal items in relation to the greatest relative daily reduction in plasma BZD concentrations and Tyrer et al.[10] determined the sensitivity to change of their withdrawal

items in patients who were previously unable to cease BZD use because of apparent withdrawal symptoms. On a scale level, the validity of the BDEPQ was assessed by comparison with convergent and divergent measures.[22] However, the BDEPQ total sum score, correlated only moderately with the DSM-III-R and ICD-10 diagnoses made using the CIDI and even more strongly with scores of supposed divergent measures.

The Bendep-SRQ, the first multidimensional psychometric instrument for assessment of the severity of BZD dependence, has been assessed more thoroughly in representative samples of BZD users with regard to scalability, reliability and validity than the mentioned BZD withdrawal questionnaires and the BDEPQ. It shows promise as a new instrument to assess BZD use in clinical practice and scientific research. Its administration does not require much time, preliminary training or special medical skills. The course of BZD dependence can be monitored in BZD users by repeated administrations of the Bendep-SRQ. This course is reflected by changes in the BZD dependence severity profile, constituted by the scores on the four Bendep-SRQ scales. If the 'Withdrawal' scale is not applicable to a respondent, because he has never made an attempt to reduce or cease his BZD use, a complete severity profile can only be obtained after a subsequent attempt to reduce BZD dose. The course of BZD dependence as reflected by the Bendep-SRQ severity profile might be of benefit to clinical decision making. Signs of increasing scale scores could initiate interventions, such as patient education or an inpatient or outpatient dose-reduction program. The effects of these interventions could be evaluated by a follow-up administration of the Bendep-SRQ.

The present study on the Bendep-SRQ aimed to collect representative outpatient samples to yield meaningful results for clinical practice. However, a single study has limitations, and thus additional research is required. The present results need to be reinforced by cross-validation of the Bendep-SRQ in similar outpatient samples from other settings. In addition,

the present research into the Bendep-SRQ should be repeated on inpatient and outpatient BZD users who are primarily treated for dependence on other substances to assess the secondary nature of BZD dependence in such subjects. Conclusions with respect to the predictive validity of the Bendep-SRQ scale scores can be made on the basis of longitudinal research data, which can be obtained if the Bendep-SRQ is administered repeatedly to monitor BZD users. Although such additional studies are obligatory, the present results indicate that the Bendep-SRQ is potentially a useful instrument in a policy aimed at the reduction of nonindicated chronic BZD use in outpatients.

NOTE

The Bendep-SRQ can be obtained from the authors (C.Kan@czzopsy.azn.nl) and is also available for on-line administration on site http://baserv.uci.kun.nl/~fzitman/Bendep-SRQ.html.

ACKNOWLEDGMENT

REFERENCES

1. Tyrer PJ. Benzodiazepines on trial [editorial]. Br Med J Clin Res Ed 1984;288:1101-1102.

2. Marks J. The Benzodiazepines. Use, overuse, misuse, abuse. Lancaster, MTP press; 1985.

3. Clinthorne JK, Cisin IH, Balter MB, Mellinger GD, Uhlenhuth EH. Changes in popular attitudes and beliefs about tranquilizers. 1970-1979. Arch Gen Psychiatry 1986;43:527-532.

4. Uhlenhuth EH, DeWit H, Balter MB, Johanson CE, Mellinger GD. Risks and benefits of long-term benzodiazepine use. J Clin Psychopharmacol 1988;8:161-167.

5. Linsen SM, Breteler MHM, Zitman FG. Defining benzodiazepine dependence: the confusion persists. European Psychiatry 1995;10:306-311.

6. World Health Organization. WHO expert committee on addiction-producing drugs; thirteenth report. World Health Organ Tech Rep Ser 1964; no. 273.

7. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders third edition. Washington, DC: American Psychiatric Association; 1980.

8. Merz WA, Ballmer U. Symptoms of the barbiturate/benzodiazepine withdrawal syndrome in healthy volunteers: standardized assessment by a newly developed self-rating scale. J Psychoactive Drugs 1983;15:71-84.

9. Busto UE, Sykora K, Sellers EM. A clinical scale to assess benzodiazepine withdrawal. J Clin Psychopharmacol 1989;9:412-416.

10. Tyrer P, Murphy S, Riley P. The Benzodiazepine Withdrawal Symptom Questionnaire. J Affect Disord 1990;19:53-61.

11. Edwards G, Arif A, Hadgson R. Nomenclature and classification of drug- and alcohol-related problems: a WHO Memorandum. Bull World Health Organ 1981;59:225-242.

12. World Health Organization. WHO Expert Committee on drug dependence; twenty-eighth report. World Health Organ Tech Rep Ser 1993;836:1-44.

13. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, Third edition., Revised. Washington, DC: American Psychiatric Association, 1987.

14. American Psychiatric Association. Diagnostic and Statistical Manual for Mental Disorders, Fourth edition. Washington, DC: American Psychiatric Association, 1994.

15. World Health Organization. The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research. Geneva: World Health Organization, 1992.

16. Edwards G, Gross MM. Alcohol dependence: provisional description of a clinical syndrome. Br Med J 1976;1:1058-1061.

17. World Health Organization. The Composite International Diagnostic Interview (CIDI). Geneva: World Health Organization, 1990.

18. Janca A, Ustun TB, Sartorius N. New versions of World Health Organization instruments for the assessment of mental disorders. Acta Psychiatr Scand 1994;90:73-83.

19. Wing JK, Babor T, Brugha T, Burke J, Cooper JE, Giel R et al. SCAN, Schedules for clinical assessment in

neuropsychiatry. Arch Gen Psychiatry 1990;47:589-593.

20. Kan CC, Breteler MHM, Zitman FG. High prevalence of benzodiazepine dependence in outpatient users, based on the DSM-III-R and ICD10 criteria. Acta Psychiatr Scand 1997;96:85-93.

21. Kan CC, Breteler MHM, Van der Ven AHGS, Zitman FG. An evaluation of the DSM-III-R and ICD-10 benzodiazepine dependence criteria using Rasch modelling. Addiction 1998;93(3):349-359.

22. Baillie AJ, Mattick RP. The benzodiazepine dependence questionnaire: development, reliability and validity. Br J Psychiatry 1996;169:276-281.

23. Giel R, Nienhuis FJ. [SCAN. Schedules for Clinical Assessment in Neuropsychiatry] SCAN. Vragenschema's voor de klinische beoordeling in de neuropsychiatrie. Geneve: World Health Organization; [division for Mental Health] divisie voor de Geestelijke Gezondheid, 1992.

24. Arrindell WA, Ettema H. [Dimensional structure, reliability and validity of the Dutch version of the Symptom Checlist (SCL-90); data based on a phobic and a 'normal' population] Dimensionele structuur, betrouwbaarheid en validiteit van de Nederlandse bewerking van de Symptom Checklist (SCL-90); gegevens gebaseerd op een fobische en een 'normale' populatie. Ned T v Psychologie 1981;36:77-108.

25. Hendriks VM, Kaplan CD, Limbeek Jv, Geerlings P. The addiction severity index: reliability and validity in a Dutch addict population. J Subst Abuse Treat 1989;6:133-141.

26. Cook TD, Campbell DT. Quasi-experimentation; Design and Analysis Issues for Field Settings. Chicago: Rand McNally College Publishing Company, 1979.

27. Fischer GH. Derivations of the Rasch Model. In: Fischer GH, Molenaar IW (eds): Rasch models: Foundations, Recent Developments and Applications. New York, NY: Springer Verlag, 1995:15-38.

28. Glas CAW. RSP: Rasch Scaling Program. Groningen, The Netherlands: IEC ProGAMMA, 1993.

29. Glas CAW. RSP: Rasch Scaling Program User's Manual. Groningen, The Netherlands: IEC ProGAMMA, 1993.

30. Glas CAW. RSP: Rasch Scaling Program. (user's manual), (Groningen, The Netherlands, IEC ProGAMMA) 1993.

31. Siegel S. The case of k related samples. In: Non-parametric Statistics for the Behavioral Scienses. New York: McGraw-Hill, inc., 1956:161-165.

32. Hoyt C. Test reliability by analysis of variance. Psychometrika 1941;6:135-160.