

# Development, internal consistency and reliability of the Verona Service Satisfaction Scale – European Version

## EPSILON Study 7

MIRELLA RUGGERI, ANTONIO LASALVIA, ROSA DALL'AGNOLA, BOB VAN WIJNGAARDEN, HELLE CHARLOTTE KNUDSEN, MORVEN LEESE, LUIS GAITE, MICHELE TANSELLA and the EPSILON STUDY GROUP

**Background** Satisfaction with mental health services is an important quality and outcome variable. The Verona Service Satisfaction Scale (VSSS) is a well-established method for measuring service satisfaction.

**Aim** To report the development and reliability study of the European Version of the VSSS (VSSS–EU).

**Method** A sample of people with schizophrenia on the case-load of local mental health services in the five European participating countries was assessed. The VSSS–EU was administered at one site in each country at two points in time. Internal consistency and test–retest reliability were assessed and compared between the five sites.

**Results** The  $\alpha$  coefficient for the VSSS–EU total score in the pooled sample was 0.96 (95% CI 0.94–0.97) and ranged from 0.92 (95% CI 0.60–1.00) to 0.96 (95% CI 0.93–0.98) across the sites. Test–retest reliability for VSSS–EU total score, pooled over sites, was 0.82 (95% CI 0.78–0.85) and ranged from 0.73 (95% CI 0.6–0.86) to 0.93 (95% CI 0.89–0.97) across the sites.

**Conclusion** VSSS–EU is a reliable instrument for measuring service satisfaction in people with schizophrenia, for use in comparative cross-national research projects and in routine clinical practice in mental health services across Europe.

**Declaration of interest** No conflict of interest. Funding detailed in Acknowledgements.

Patient satisfaction is an important variable in the evaluation of psychiatric services and it complements the measurements of other outcome variables. It has been suggested (Kalman, 1983) that satisfaction is strictly linked to the effectiveness of the care provided, whereas dissatisfaction is frequently the reason behind patients' discontinuing psychiatric care (Ware *et al*, 1978; Hansen *et al*, 1992). Ensuring high levels of patient satisfaction is, therefore, an essential aim for any mental health service, and its measurement constitutes a valid and important aspect of service planning and evaluation (Donabedian, 1966), to the extent that in many countries providers of health care are increasingly required to monitor levels of satisfaction among patients.

Unfortunately, there is a lack of knowledge about the extent to which patients with psychosis are satisfied with services. This seems to be due both to prejudice against such patients, believing them to be incapable of meaningfully judging the care they receive, and to methodological problems of measurement, mainly related to the difficulty of providing instruments which are acceptable to these patients.

Research in the area of satisfaction with psychiatric services has been hampered by the widespread use of many non-standardised methods, so that direct comparison between studies is usually impossible. Most studies have used instruments with few or no data regarding their validity or reliability, and investigators have frequently designed their own instruments for specific studies. As a result, findings are not generalisable (Ruggeri, 1994). In addition, although satisfaction has been demonstrated to be a multi-dimensional concept (Ware *et al*, 1978), instruments have often been limited to a few broad items which only enquire about one or two dimensions of mental health care. Thus not only may they fail to detect any dissatisfaction – they

are inherently unable to detect the reasons for dissatisfaction.

The Verona Service Satisfaction Scale (VSSS) (Ruggeri & Dall'Agnola, 1993) is a questionnaire meant to fill this gap: it is a validated, multi-dimensional scale which measures the satisfaction of patients with mental health services. First, an 82-item version was developed: it consisted of a set of 37 items cross-setting for health services and a set of 45 items specific for mental health services. The former group of items involves aspects meant a priori to be relevant across a broad array of both medical and psychiatric settings, and was derived from the Service Satisfaction Scale (SSS–30) (Greenfield & Attkinson, 1989; Attkinson & Greenfield, 1994; Ruggeri & Greenfield, 1995). The latter group of items involves aspects which are specifically relevant in mental health settings, particularly in community-based services, such as social skills and types of intervention (e.g. admissions, psychotherapy, rehabilitation) and has been developed for the purpose by the authors of the VSSS. VSSS–82, in its versions for patients and relatives, has then been tested for acceptability, content validity, sensitivity and test–retest reliability (Ruggeri & Dall'Agnola, 1993; Ruggeri *et al*, 1994) in 75 patients and 75 relatives. Finally, a factor analysis was performed (Ruggeri *et al*, 1996). The combination of results obtained in the validation study and factor analysis gave rise to the Intermediate (VSSS–54) and the Short (VSSS–32) versions, two reliable instruments that can be easily used in everyday clinical settings. So far, the VSSS has been translated into various languages and used in studies performed in many sites in the world (Cozza *et al*, 1997; Parkman *et al*, 1997; Leese *et al*, 1998; Boardman *et al*, 1999; Clarkson *et al*, 1999; Henderson *et al*, 1999; Merinder *et al*, 1999).

The aim of this paper is to describe the development, translation, cultural validation and reliability of a new European Version of the VSSS (VSSS–EU), for use in multi-site international comparative studies.

## METHOD

### EPSILON Study

The present study is part of a larger research project, the EPSILON (European Psychiatric Services: Inputs Linked to Outcome Domains and Needs) Study, a comparative cross-national cross-sectional

study of the characteristics, needs, quality of life, patterns of care, associated costs and satisfaction levels of people with schizophrenia in five European countries (Becker *et al*, 1999) which aims to:

- (a) produce standardised versions of five instruments in key areas of mental health service research in five European languages (Danish, Dutch, English, Italian and Spanish), each of which was converted from the original into the other four languages;
- (b) obtain and compare data from five regions in different European countries about social and clinical variables, characteristics of mental health care and its costs; and
- (c) test both instrument and cross-instrument hypotheses.

Five instruments were converted for use from their original language into the other study languages by: (a) accurate translation and back-translation into/from the other languages, (b) checks of cross-cultural applicability using focus groups, and (c) assessment of instrument reliability. Fuller details on the translation and cross-cultural adaptation process are given in this supplement by Knudsen *et al* (2000). The five instruments were the Camberwell Assessment of Need (CAN) (Phelan *et al*, 1995), Client Service Receipt Inventory (CSRI) (Beecham & Knapp, 1992), Involvement Evaluation Questionnaire (IEQ) (Schene *et al*, 1998), Lancashire Quality of Life Profile (LQoLP) (Oliver, 1991; Oliver *et al*, 1996) and Verona Service Satisfaction Scale (VSSS) (Ruggeri & Dall'Agnola, 1993).

The VSSS reliability results are presented in this paper, while the results for the other scales are given in other papers in this supplement (Chisholm *et al*, 2000; Gaité *et al*, 2000; McCrone *et al*, 2000; van Wijngaarden *et al*, 2000).

### Study sites

Six partners in five European countries (The Netherlands, Denmark, England, Spain and Italy) joined forces. The teams are located in Amsterdam, Copenhagen, London (Centre for Economics in Mental Health, and Section of Community Psychiatry), Santander and Verona. The criteria used to identify study centres and the key summary characteristics of each study site are given in Becker *et al* (1999).

### Case identification

Cases included in the study were adults aged 18–65 inclusive with an ICD–10 diag-

nosis of any of F20 to F25. Administrative prevalence samples of people with these diagnoses were identified either from Psychiatric Case Registers (in Copenhagen and in Verona) or from the case-loads of local specialist mental health services (in-patients, out-patients and community). Patients needed to have been in contact with mental health services during the 3-month period preceding the start of the study. Thus, an administrative prevalence sample of people with schizophrenia in contact with mental health services was used in each site as the sampling frame. Cases identified were diagnosed using the item group checklist (IGC) of the Schedule for Clinical Assessment in Neuropsychiatry (SCAN) (World Health Organization, 1992). Only patients with an ICD–10 F20 research diagnosis were included in the study.

The exclusion criteria were: current residence in prison, secure residential services or hostels for long-term patients; co-existing learning disability (mental retardation), primary dementia or other severe organic disorder; and extended inpatient treatment episodes longer than one year. The numbers of patients finally included in the study varied from 52 to 107 between the five sites, with a total of 404.

### Verona Service Satisfaction Scale – European Version (VSSS–EU)

The VSSS–EU was developed from the Italian VSSS–54 version, which was first translated into Danish, Dutch, English and Spanish by professional translators. The resulting four translations were then back-translated into Italian. The back-translations were checked by the authors of the VSSS and compared with the original Italian version. Small discrepancies were examined and alterations made to the Italian version in order to preserve the precise meaning of each question, while still producing an understandable and acceptable translation into the various languages. Specific items were changed to adapt them to the context of each country's mental health system. When those changes occurred, the local researchers made a list of modifications. The next step was the focus group process, which discussed the content and the language of the translated instruments. In the light of the comments and recommendations, the instrument was then revised both in its original Italian version and in each of the four translations. On the basis of the focus groups' discussions, minor

changes were made to the wording of the items, and a further change was made by deciding to assess separately the skills and behaviour of psychiatrists, psychologists, nurses and social workers, who in the previous version had been assessed in pairs (psychiatrists/psychologists, nurses/social workers) (see Appendix).

Conceptually, the items in the VSSS–EU cover seven dimensions: Overall Satisfaction, Professionals' Skills and Behaviour, Information, Access, Efficacy, Types of Intervention and Relative's Involvement. Items in the first five dimensions cover all areas belonging to Ware's taxonomy of satisfaction (Ware *et al*, 1983). The last two dimensions, on the other hand, examine domains that have not been assessed systematically in previous studies and have been specifically developed for the VSSS.

Each conceptual dimension of the VSSS–EU consists of a certain number of items that cover various aspects of satisfaction with services (see Appendix for a detailed description of each item):

- (a) the Overall Satisfaction dimension consists of three items which cover general aspects of satisfaction with psychiatric services;
- (b) the Professionals' Skills and Behaviour dimension consists of 24 items which cover various aspects of satisfaction with the professionals' behaviour, such as technical skills, interpersonal skills, cooperation between service providers, respect of patients' rights, etc.; psychiatrists, psychologists, nurses and social workers are assessed in separate items;
- (c) the Information dimension consists of three items which cover aspects related to satisfaction with information on services, disorders and therapies;
- (d) the Access dimension consists of two items which cover aspects related to satisfaction with service location, physical layout and costs;
- (e) the Efficacy dimension consists of eight items which cover aspects related to satisfaction with overall efficacy of the service, and service efficacy on specific aspects such as symptoms, social skills and family relationships;
- (f) the Types of Intervention dimension consists of 17 items which cover various aspects of satisfaction with care, such as drugs prescription, response to emergency, psychotherapy, rehabilitation, domiciliary care, admissions, housing, recreational activities, work, benefits, etc.

**Table 1** The Verona Service Satisfaction Scale – European Version (VSSS–EU) dimensions and total means in the pooled sample and by site

Dimension	Pooled n=399		Amsterdam n=58		Copenhagen n=51		London n=83		Santander n=100		Verona n=107		Test of equality of means (P-value)	Test of equality of s.d. (P-value)
	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.	mean	s.d.		
Overall satisfaction	3.83	0.79	3.90	0.80	4.04	0.79	3.45	0.67	3.79	0.84	4.01	0.72	0.05	0.27
Professionals' skills and behaviour	3.88	0.57	3.97	0.51	4.13	0.56	3.46	0.40	3.94	0.53	4.00	0.60	<0.01	0.03
Information	3.39	0.93	3.66	0.75	3.69	0.86	3.26	0.65	2.93	1.09	3.64	0.88	<0.01	<0.01
Access	3.83	0.73	3.63	0.73	4.19	0.71	3.97	0.67	3.84	0.67	3.64	0.79	<0.01	0.35
Efficacy	3.56	0.74	3.69	0.69	3.80	0.76	3.22	0.55	3.41	0.77	3.81	0.74	<0.01	0.03
Types of intervention	3.64	0.42	3.65	0.47	3.72	0.42	3.68	0.24	3.42	0.41	3.75	0.46	<0.01	<0.01
Relative's involvement	3.39	0.96	3.57	0.92	3.32	1.21	2.91	0.67	3.39	0.96	3.75	0.91	<0.01	<0.01
Total score	3.70	0.50	3.79	0.46	3.89	0.48	3.45	0.34	3.59	0.51	3.86	0.54	<0.01	0.01

In all sites, sample sizes for each dimension varied slightly from the maximum: the number of cases dropped in each dimension was less than two, except in Relative's Involvement, for which the sample size was 359 in the pooled sample (49 in Amsterdam, 45 in Copenhagen, 80 in London, 96 in Santander and 89 in Verona).

(g) the Relative's Involvement dimension consists of six items which cover various aspects of the patient's satisfaction with help given to his/her closest relative, such as listening, understanding, advice, information, help coping with the patient's problems, etc.

The VSSS–EU is designed for self-administration and can be completed without prior training. In cases of cognitive deficit, severe psychopathology or low level of literacy, a research worker may assist the patient and/or the relative by reading through the items with them. Special care must be taken to guarantee confidentiality and anonymity and, in the case of assisted administration, to stress the independence of the research worker from the clinical team. Administering the questionnaire takes 20–30 minutes.

In the VSSS–EU, subjects are asked to express their overall feeling about their experience of the mental health service they have been using in the past year.

For items 1–40, satisfaction ratings are on a 5-point Likert scale (1=terrible, 2=mostly dissatisfactory, 3=mixed, 4=mostly satisfactory, 5=excellent), presented with alternate directionality to reduce stereotypic response. Items 41–54 consist of three questions: first, the subject is asked if he/she has received the specific intervention (*Question A*: "Did you receive the intervention x in the last year?"). If the answer is 'yes', he/she is asked his/her satisfaction on a 5-point Likert scale (*Question B*). If the answer is 'no', he/she is asked *Question C*: "Do you think you would have liked to receive intervention x?" (6=no, 7=don't

know, 8=yes). These questions permit the estimation of the subject's degree of satisfaction both with the interventions provided and with the professional's decision not to provide an intervention (if that was the case). The latter may be considered a measure of underprovision of care, from the patient's point of view.

### Data management

SPSS (Version 6.1 and above) was used for data entry. Data consistency and homogeneity were ensured by the coordinating centre (London) preparing the SPSS templates for use at all participating sites; this ensured consistent data structures across sites. Subsequent cross-instrument merging and analysis was possible as the same patient/carer identifiers were used throughout the study.

### Reliability assessment procedures

Reliability testing in the EPSILON project has been conducted on several levels, depending on the nature of the instruments involved and the way they were administered (interviews v. questionnaires). Details on the general methodology of the reliability study are given elsewhere in this supplement (Schene *et al.*, 2000).

As far as the VSSS–EU is concerned, the reliability tests have been performed: (a) on the VSSS–EU total mean score; (b) on the mean scores of each dimension; and (c) item by item, both for the pooled sample and across the five sites. Three kinds of reliability tests have been used: Cronbach's  $\alpha$  (Cronbach, 1951), to check the internal

consistency of the whole questionnaire and the different dimensions; the intraclass correlation coefficient (ICC) (Bartko & Carpenter, 1976), to evaluate test–retest reliability of the VSSS–EU total mean score and dimension mean scores; and Cohen's weighted  $\kappa$  (Cohen, 1968), to evaluate test–retest reliability of single VSSS–EU items. Additional statistics estimated were the standard errors of measurement, which were obtained from the analysis of variance used to estimate the intraclass correlations.

Statistical analysis was performed using SPSS for Windows, release 7.5 (Norusis, 1997), the Amsterdam  $\alpha$ -testing program, ALPHA.EXE based on Feldt *et al.* (1987) and EXCEL for tests of the homogeneity of ICCs.

## RESULTS

A total sample of 289 subjects (49 in Amsterdam, 43 in Copenhagen, 81 in London, 50 in Santander and 66 in Verona) completed both test and retest VSSS–EU administration. Retesting was done after a mean of 10.4 days (s.d. 6.03) in the pooled sample, 10.7 days (s.d. 7.45) in Amsterdam, 9.3 days (s.d. 3.57) in Copenhagen, 9.4 days (s.d. 3.93) in London, 6.1 days (s.d. 2.82) in Santander, 15.4 days (s.d. 6.70) in Verona. The shortest and longest test–retest intervals were in Santander and Verona, respectively (Kruskal–Wallis test,  $P < 0.001$ ).

### Internal consistency

Table 1 shows the mean scores for each VSSS–EU dimension in the various EPSILON sites and the tests of homogeneity of

variance (Levene test) and of means. The latter tests are not discussed in the present study, but will be investigated in a future paper. Assuming that the standard error of measurement is the same in all countries, the reliability coefficient depends on the variance in each sample. Conversely, where standard errors of measurement do differ, sample variables will be affected, since they are composed both of variance in true scores and variation due to measurement error. Therefore, when computing reliability coefficients and testing differences between samples, differences between sites in terms of both sample variance and standard errors of measurement should be considered.

The variability of the scores differed across the sites, both in the VSSS-EU total mean score and in all VSSS-EU dimensions, with the exception of Overall Satisfaction and Access; therefore the lack of homogeneity across the sites for the total mean scores and most of the dimension scores may account to some extent for the degree of variability in the reliability coefficients across the sites.

Table 2 shows the  $\alpha$  coefficients and the test for the equality of  $\alpha$  across the sites. Alpha coefficients indicate the degree to which items exhibit a positive correlation (internal consistency above 0.7 is considered adequate; Bech *et al*, 1993).

It should first be noted that the high degree of variability in the number of cases is usually not due to missing values but to the fact that some items (i.e. items on the performances of the different professionals assessed in the VSSS, items on the relatives, items on response of the service to emergencies during the night or at weekends) are not applicable to all patients. For this reason,  $\alpha$  values in the VSSS-EU total mean scores have been computed for a small number of cases only. However, they were always over 0.90 and were similar across the sites, with the exception of Relative's Involvement. In the VSSS-EU dimensions,  $\alpha$  coefficients ranged from 0.60 (Information dimension, in Copenhagen) to 0.93 (Relative's Involvement dimension, in Copenhagen) and did not differ significantly across the sites. The dimension Access, which consists of just two items (Costs of Service and Physical Layout) measuring different constructs, is a special case. In this dimension,  $\alpha$  varied greatly, ranging from 0.08 (London) to 0.96 (Copenhagen), and differed significantly across the sites. Dimensions constituted by a higher number of items are expected to have higher  $\alpha$  values. This was true for the dimension Professionals' Skills and Behaviour, but not for the dimension Types of Intervention, which is not expected to have high

internal consistency, due to the wide range of different interventions explored by the questionnaire.

On the whole,  $\alpha$  values of the pooled sample were good, with the above-mentioned exception of the Access dimension, and ranged from 0.72 (Information) to 0.91 (Professionals' Skills and Behaviour).

### Stability

The test-retest reliability was studied by considering the ratings both in each dimension (ICC) and item by item (weighted  $\kappa$ ).

Table 3 shows the test-retest reliability for the VSSS-EU total mean score and VSSS-EU dimension scores, both in the pooled sample and at each EPSILON site. ICCs and the standard errors of measurements are reported.

The VSSS-EU total mean score in the pooled sample shows a high degree of reliability (0.82; 95% CI 0.78–0.85), although there are some differences between the sites, with all sites above 0.70 and three sites (Copenhagen, London and Santander) above 0.80. Each of the VSSS-EU dimension mean scores in the pooled sample had a degree of reliability ranging from 0.56 to 0.78. The reliability of the VSSS-EU dimension mean scores across the sites was over 0.50 in all cases, with the exception of the Access dimension in Santander (0.43) and

**Table 2** Internal consistency of the Verona Service Satisfaction Scale – European Version (VSSS-EU):  $\alpha$  coefficients (95% CI) in the pooled sample and by site

	No. of items	Pooled		Amsterdam		Copenhagen		London		Santander		Verona		Test of equality of $\alpha$ (P-value)
		n	$\alpha$	n	$\alpha$	n	$\alpha$	n	$\alpha$	n	$\alpha$	n	$\alpha$	
Overall satisfaction	3	384	0.80 (0.77–0.83)	57	0.80 (0.71–0.87)	45	0.82 (0.73–0.88)	83	0.77 (0.69–0.83)	97	0.83 (0.78–0.88)	102	0.73 (0.65–0.80)	0.67
Professionals' skills and behaviour <sup>1</sup>	16	275	0.91 (0.89–0.92)	38	0.90 (0.89–0.95)	24	0.89 (0.82–0.95)	75	0.85 (0.80–0.89)	72	0.90 (0.86–0.93)	66	0.91 (0.88–0.94)	0.58
Information	3	342	0.72 (0.68–0.76)	51	0.67 (0.51–0.78)	40	0.60 (0.37–0.75)	82	0.63 (0.49–0.73)	72	0.79 (0.71–0.85)	97	0.72 (0.63–0.79)	0.39
Access	2	386	0.06 (–0.11 to 0.20)	53	0.16 (–0.33 to 0.47)	49	0.96 (0.93–0.97)	83	0.08 (–0.33 to 0.36)	99	0.77 (0.68–0.83)	102	0.29 (0.02–0.49)	<0.01
Efficacy	8	254	0.87 (0.84–0.87)	37	0.83 (0.75–0.89)	28	0.82 (0.73–0.90)	58	0.77 (0.68–0.84)	69	0.89 (0.85–0.92)	62	0.89 (0.85–0.92)	0.11
Types of intervention	17	115	0.73 (0.67–0.79)	19	0.81 (0.69–0.90)	8	0.61 (0.17–0.88)	52	0.62 (0.48–0.74)	4	– <sup>2</sup>	32	0.77 (0.66–0.86)	0.71
Relative's involvement <sup>1</sup>	5	298	0.89 (0.87–0.91)	40	0.85 (0.77–0.91)	26	0.93 (0.88–0.97)	67	0.81 (0.72–0.87)	92	0.91 (0.87–0.94)	73	0.88 (0.82–0.91)	0.04
Total score	54	74	0.96 (0.94–0.97)	16	0.95 (0.91–0.98)	4	– <sup>2</sup>	33	0.93 (0.90–0.96)	2	– <sup>2</sup>	19	0.96 (0.93–0.98)	1.00

1. For psychometric reasons the pairs of items analysing the same characteristic of different professionals (either psychiatrists and psychologists, or nurses and social workers) have been collapsed into one item. This process involved the following items (see Appendix): 3a and 3b; 5a and 5b; 6a and 6b; 16a and 16b; 22a and 22b; 25a and 25b; 30a and 30b; 35a and 35b; 37a and 37b.

2. Insufficient data to estimate  $\alpha$ .

**Table 3** Test–retest reliability of Verona Service Satisfaction Scale – European Version (VSSS–EU) dimensions and total means in the pooled sample and by site

Sub-scale	Pooled n=289		Amsterdam n=49		Copenhagen n=43		London n=81		Santander n=50		Verona n=66		Test of equality of ICCs (P-value)
	ICC	(s.e.) <sub>m</sub>	ICC	(s.e.) <sub>m</sub>	ICC	(s.e.) <sub>m</sub>	ICC	(s.e.) <sub>m</sub>	ICC	(s.e.) <sub>m</sub>	ICC	(s.e.) <sub>m</sub>	
Overall satisfaction	0.66	0.44	0.50	0.51	0.67	0.46	0.76	0.33	0.80	0.37	0.52	0.52	0.01
Professional's skills and behaviour	0.76	0.25	0.66	0.29	0.80	0.24	0.78	0.19	0.86	0.19	0.72	0.31	0.06
Information	0.75	0.14	0.49	0.52	0.75	0.42	0.66	0.37	0.84	0.44	0.76	0.41	<0.01
Access	0.56	0.47	0.56	0.48	0.51	0.49	0.73	0.33	0.43	0.51	0.51	0.55	0.05
Efficacy	0.75	0.34	0.70	0.36	0.76	0.35	0.77	0.26	0.90	0.23	0.59	0.45	<0.01
Types of intervention	0.69	0.22	0.56	0.32	0.64	0.25	0.82	0.09	0.90	0.13	0.66	0.26	<0.01
Relative's involvement	0.78	0.43	0.72	0.47	0.82	0.53	0.70	0.38	0.89	0.32	0.67	0.48	<0.01
Total score	0.82	0.20	0.73	0.24	0.85	0.19	0.82	0.15	0.93	0.13	0.76	0.25	<0.01

ICC, Intraclass correlation coefficient; (s.e.)<sub>m</sub>, standard error of measurement (square root of error component of variance).

**Table 4** Test–retest reliability of individual items of the Verona Service Satisfaction Scale – European Version (VSSS–EU) for the pooled sample and in each study site. Number (percentage) of items in Cohen's weighted  $\kappa$  bands suggested by Landis & Koch (1977)

Weighted $\kappa$	Strength of agreement	Pooled n=289	Amsterdam n=49	Copenhagen n=43	London n=81	Santander n=50	Verona n=66
		n (%)	n (%)	n (%)	n (%)	n (%)	n (%)
0.81–1.0	Almost perfect			2 (3)	14 (22)	10 (16)	
0.61–0.80	Substantial	31 (49)	8 (13)	23 (36)	36 (57)	33 (52)	8 (12.7)
0.41–0.60	Moderate	32 (51)	37 (59)	25 (40)	11 (17)	16 (25)	26 (41.2)
0.21–0.40	Fair		18 (29)	10 (16)	2 (3)	4 (6)	22 (34.9)
0–0.20	Slight			3 (5)			7 (11.1)

the Information dimension in Amsterdam (0.49). Overall, there was some degree of variability in the ICC coefficients across the sites. This could be due either to differences in measurement errors between different sites, or to lack of homogeneity in the samples. The lower reliabilities tend to be associated with higher standard errors of measurement, and therefore lack of homogeneity in the samples is unlikely to be the explanation for differences in reliability; rather, the performance of the instrument itself differs.

Table 4 shows Cohen's weighted  $\kappa$  for both the pooled sample and the five EPSILON sites, calculated in each VSSS–EU item by taking the disagreements' weight to be equal to the square of the distances. According to Landis & Koch (1977), a  $\kappa$  coefficient of 0.2–0.4 indicates fair agreement, 0.4–0.6 indicates moderate agreement, 0.6–0.8 indicates a substantial agreement and 0.8–1.0 indicates almost perfect agreement, although alternative schemes are possible (see Schene *et al*, 2000, this supplement).

In the pooled sample, all items generated  $\kappa$  coefficients from 0.4 to 0.8, indicating a

moderate–substantial agreement. In the various EPSILON sites the percentage of items with  $\kappa$  coefficients exceeding 0.4 (generally accepted as the minimum value) ranged from 54% (Verona) to 97% (London).

A paired sample *t*-test on the difference between test and retest for the VSSS–EU total mean scores and dimension sub-scores revealed no significant differences, both pooled across sites and at individual sites (with the exception of Access in London and Santander, where the retest values were respectively higher and lower than the time 1 values,  $P < 0.001$ ), thus showing no overall tendency for patients to respond more or less favourably after an interval of time ranging from 1 to 2 weeks.

## DISCUSSION

This paper describes the development of the VSSS–EU, a modified version of the VSSS–54, for measuring satisfaction with mental health services in people with schizophrenia and intended to be used in international collaborative studies across

Europe. It also describes the assessment of the psychometric properties of the new instrument, with regard to its internal consistency and stability. The investigation of the psychometric properties of the measures of satisfaction is an outstanding issue in research into satisfaction with mental health services. Although there is little doubt that patient satisfaction is an important aspect in the assessment of the quality and outcome of community-based mental health programmes, research in the field suffers from various methodological limitations regarding study design, the construction of the instrument and the lack of attention to its psychometric properties. According to many authors, the lack of attention to methodological aspects and the lack of confidentiality strongly influence the levels of reported satisfaction (Larsen *et al*, 1979; Weinstein, 1979; Keppler-Seid *et al*, 1980; Ruggeri, 1994).

In the past few years, the need for a measure which could assess satisfaction with mental health services in a manner which would be correct from a methodological point of view led to the

development of the VSSS (Ruggeri & Dall'Agnola, 1993). One of the most important aims during the development of the original version of the VSSS was the avoidance of methodological biases: data previously obtained on content validity, test-retest reliability and factor analysis showed that the original version measures satisfaction in a sensitive, valid and reliable way (Ruggeri & Dall'Agnola, 1993; Ruggeri *et al*, 1994, 1996). The present paper demonstrates that the newly developed VSSS-EU also has good psychometric properties.

The scale has an excellent overall Cronbach's  $\alpha$  (0.96), which confirms its adequacy from the standpoint of internal consistency if it is used as a global satisfaction measure. All dimensions except one (Access) showed good internal consistency, as indicated by coefficients of Cronbach's  $\alpha$  ranging from 0.72 for Information to 0.91 for Professionals' Skills and Behaviour. The dimension with the lowest consistency (Access), made up of two items (Cost of Service and Physical Layout), was not expected to show high interrelatedness, because theoretically these two items are quite independent of one another.

The test-retest data provided encouraging results. The stability of the scale was satisfactory both at dimensional level and item by item. Although there was some evidence of differing reliability between sites, which was associated with differing standard errors of measurement rather than differing homogeneity of the samples, all the individual site reliability coefficients for the total score were above 0.7. Interestingly, the lowest test-retest reliability was found in Verona, thus indicating that the translation of the scale did not affect its acceptability to patients. The overall pooled test-retest reliability for the total score was excellent at 0.82 (95% CI 0.78–0.85). Furthermore, there was no evidence of any tendency for the measurements to change systematically over time, either in the pooled sample or across the sites (with the above-mentioned exception of Access).

In conclusion, the analysis of the data presented here demonstrates that the VSSS-EU has good psychometric properties and suggests that it is a reliable instrument for measuring satisfaction with mental health services in people with schizophrenia, for use in comparative cross-national research projects and in routine clinical practice in mental health services across Europe.

## ACKNOWLEDGEMENTS

The following colleagues contributed to the EPSILON Study. Amsterdam: Dr Maarten Koeter, Karin Meijer, Dr Marcel Monden, Professor Aart Schene, Madelon Sijenaar, Bob van Wijngaarden; Copenhagen: Dr Helle Charlotte Knudsen, Dr Anni Larsen, Dr Klaus Martiny, Dr Carsten Schou, Dr Birgitte Welcher; London: Professor Thomas Becker, Dr Jennifer Beecham, Liz Brooks, Daniel Chisholm, Gwyn Griffiths, Julie Grove, Professor Martin Knapp, Dr Morven Leese, Paul McCrone, Sarah Padfield, Professor Graham Thornicroft, Ian R. White; Santander: Andrés Arriaga Arrizabalaga, Sara Herrera Castaneda, Dr Luis Gaité, Andrés Herran, Modesto Perez Retuerto, Professor José Luis Vázquez-Barquero, Elena Vázquez-Bourgon; Verona: Dr Francesco Amaddeo, Dr Giulia Bisoffi, Dr Doriana Cristofalo, Dr Rosa Dall'Agnola, Dr Antonio Lasalvia, Dr Mirella Ruggeri, Professor Michele Tansella.

This study was supported by the European Commission BIOMED-2 Programme (Contract BMH4-CT95-1151). We would also like to acknowledge the sustained and valuable assistance of the users, carers and the clinical staff of the services in the five study sites. In Amsterdam, the EPSILON Study was partly supported by a grant from the Nationaal Fonds Geestelijke Volksgezondheid and a grant from the Netherlands Organization for Scientific Research (940-32-007). In Santander the EPSILON Study was partially supported by the Spanish Institute of Health (FIS) (FIS Exp. No. 97/1240). In Verona additional funding for studying patterns of care and costs of a cohort of patients with schizophrenia were provided by the Regione del Veneto, Giunta Regionale, Ricerca Sanitaria Finalizzata, Venezia, Italia (Grant No. 723/01/96 to Professor M. Tansella).

## APPENDIX

(The number of items in each dimension is shown in parentheses.)

### Overall Satisfaction (3)

11. Amount of help received
20. Kind of services
21. Overall satisfaction

### Professionals' Skills and Behaviour (16)

- 3a. Professionalism and competence of the psychiatrists
- 3b. Professionalism and competence of the psychologists
- 16a. Thoroughness of psychiatrists
- 16b. Thoroughness of psychologists
- 22a. Professional competence of the nursing staff
- 22b. Professional competence of social workers
- 35a. Thoroughness of nurses
- 35b. Thoroughness of social workers
- 6a. Psychiatrists' manner
- 6b. Psychologists' manner

- 5a. Ability of psychiatrists to listen and understand problems
- 5b. Ability of psychologists to listen and understand problems
- 25a. Nurses' manner
- 25b. Social workers' manner
28. Nurses' knowledge of patient's medical history
- 37a. Ability of nurses to listen and to understand problems
- 37b. Ability of social workers to listen and to understand problems
2. Behaviour and manners of reception or secretarial staff
33. Instruction on what to do between visits
18. Cooperation between service providers
17. Referring to general practitioner or other specialists
40. Opportunity of being followed up by the same professionals
10. Confidentiality and respect for patient's rights
7. Punctuality of the professionals when patient comes for an appointment

### Information (3)

12. Explanation of specific procedures and approaches used
29. Information on diagnosis and prognosis
19. Publicity or information on mental health services which are offered

### Access (2)

4. The appearance, comfort level and physical layout
8. Costs of the service

### Efficacy (8)

9. Effectiveness of the service in attaining well-being and preventing relapses
1. Effectiveness of the service in helping patient deal with problems
24. Effectiveness of the service in helping patient improve knowledge and understanding of his/her problems
13. Effectiveness of the service in helping patient to relieve symptoms
26. Effectiveness of the service in improving the relationship between patient and relative
34. Effectiveness of the service in helping patient improve capacity to look after her/himself
31. Effectiveness of the service in helping patient establish good relationships outside family environment
38. Effectiveness of the service in helping patient improve abilities to work

### Types of Intervention (17)

14. Response of the service to crisis or urgent needs during office hours

15. Response of the service to emergencies during the night, weekends and Bank Holidays
39. Help received for unexpected outcomes, discomfort or side-effects of medication
41. Medication prescription
42. Individual rehabilitation
43. Individual psychotherapy
44. Compulsory treatment in hospital
45. Family therapy
46. Living in sheltered accommodation
47. Participation in the recreational activities organised by mental health services
48. Group psychotherapy
49. Sheltered work
50. Informal admission to hospital
51. Practical help by the service at home
52. Help in obtaining welfare benefits or exemptions
53. Help to find open employment
54. Help from the service to join in recreational activities separate from the mental health services

#### Relative's Involvement (5)

- 30a. Ability of psychiatrists to listen and understand the concerns and the opinions relative may have about patient
- 30b. Ability of psychologists to listen and understand the concerns and the opinions relative may have about patient
23. Recommendations about how relative could help
32. Information to relative about diagnosis and prognosis
36. Effectiveness of the service in helping relative to deal better with patient's problems
27. Effectiveness of the service in helping relative improve his/her understanding of patient's problems

## REFERENCES

**Attkinson, C. C. & Greenfield, T. K. (1994)** The Client Satisfaction Questionnaire-8 and the Service Satisfaction Questionnaire-30. In *Psychological Testing: Treatment Planning and Outcome Assessment* (ed. M. Maruish), pp. 404–420. San Francisco: Erlbaum.

**Bartko, J. J. & Carpenter, W. T. (1976)** On the methods and theory of reliability. *Journal of Nervous and Mental Disease*, **163**, 307–317.

**Bech, P., Mault, U. F., Dencker, S. J., et al (1993)** Scales for assessment of diagnosis and severity of mental disorders. *Acta Psychiatrica Scandinavica*, **87** (suppl. 372), 1–87.

**Becker, T., Knapp, M., Knudsen, H. C., et al (1999)** The EPSILON study of schizophrenia in five European countries: design and methodology for standardising outcome measures and comparing patterns of care and service costs. *British Journal of Psychiatry*, **175**, 514–521.

**Beecham, J. & Knapp, M. (1992)** Costing psychiatric interventions. In *Measuring Mental Health Needs* (eds G.

MIRELLA RUGGERI, MD, ANTONIO LASALVIA, MD, ROSA DALLAGNOLA, PhD, MICHELE TANSELLA, MD, Department of Medicine and Public Health, Section of Psychiatry, University of Verona, Italy; BOB VAN WIJNGAARDEN, MA, Department of Psychiatry, Academic Medical Centre, Amsterdam, The Netherlands; HELLE CHARLOTTE KNUDSEN, MD, Institute of Preventive Medicine, Copenhagen University Hospital, Denmark; MORVEN LEESE, PhD, Section of Community Psychiatry (PRiSM), Institute of Psychiatry, King's College, London, UK; LUIS GAITE, MD, Clinical and Social Psychiatry Research Unit, University of Cantabria, Santander, Spain

Dr Mirella Ruggeri, Dipartimento di Medicina e Sanità Pubblica, Sezione di Psichiatria, Università di Verona, Ospedale Policlinico, via delle Menegone, 10, 37134 Verona, Italy. Tel: +39 045 8074441; fax: +39 045 585871; e-mail: mruggeri@borgoroma.univr.it

Thornicroft, C. R. Brewin & J. Wing), pp. 163–183. London: Gaskell.

**Boardman, A. P., Hodgson, R. E., Lewis, M., et al (1999)** North Staffordshire Community Beds Study: longitudinal evaluation of psychiatric in-patient units attached to community mental health centres. I: Methods, outcome and patient satisfaction. *British Journal of Psychiatry*, **175**, 70–78.

**Chisholm, D., Knapp, M. R. J., Knudsen, H. C., et al (2000)** The Client Socio-Demographic and Service Receipt Inventory: development of an instrument for international research. EPSILON Study 5. *British Journal of Psychiatry*, **177** (suppl. 39), s28–s33.

**Clarkson, P., McCrone, P., Sutherby, K., et al (1999)** Outcomes and costs of a community support worker service for the severely mentally ill. *Acta Psychiatrica Scandinavica*, **99**, 196–206.

**Cohen, J. (1968)** Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, **70**, 213–220.

**Cozza, M., Amara, M., Butera, N., et al (1997)** La soddisfazione dei pazienti e dei familiari nei confronti di un Dipartimento di Salute Mentale di Roma. *Epidemiologia e Psichiatria Sociale*, **6**, 173–183.

**Cronbach, L. J. (1951)** Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

**Donabedian, A. (1966)** Evaluating the quality of medical care. *Millbank Memorial Fund Quarterly*, **44**, 166–203.

**Feldt, L. S., Woodruff, D. J. & Salih, F. A. (1987)** Statistical inference for coefficient alpha. *Applied Psychological Measurement*, **11**, 93–103.

**Gaite, L., Vázquez-Barquero, J. L., Arriaga Arrizabalaga, A., et al (2000)** Quality of life in schizophrenia: development, reliability and internal consistency of the Lancashire Quality of Life Profile – European Version. EPSILON Study 8. *British Journal of Psychiatry*, **177** (suppl. 39), s49–s54.

**Greenfield, T. K. & Attkinson, C. C. (1989)** Steps toward a multifactorial satisfaction scale for primary care and mental health services. *Evaluation and Program Planning*, **12**, 271–278.

**Hansen, A. M. D., Hoogduin, C. A. L., Schaap, C., et al (1992)** Do dropouts differ from successfully treated obsessive–compulsives? *Behavioural Research and Therapy*, **30**, 547–550.

**Henderson, C., Phelan, M., Loftus, L., et al (1999)** Comparison of patient satisfaction with community-based vs. hospital psychiatric services. *Acta Psychiatrica Scandinavica*, **99**, 188–195.

**Kalman, T. P. (1983)** An overview of patient satisfaction with psychiatric treatment. *Hospital and Community Psychiatry*, **34**, 48–54.

**Keppeler-Seid, H., Windle, C. & Woy, J. (1980)** Performance measures for mental health programs. *Community Mental Health Journal*, **16**, 217–234.

**Knudsen, H. C., Vázquez-Barquero, J. L., Welcher, B., et al (2000)** Translation and cross-cultural adaptation of outcome measurements for schizophrenia. EPSILON Study 2. *British Journal of Psychiatry*, **177** (suppl. 39), s8–s14.

**Landis, J. & Koch, G. (1977)** The measurement of the observer agreement for categorical data. *Biometrics*, **33**, 159–174.

**Larsen, D. L., Attkinson, C. C., Hargreaves, W. A., et al (1979)** Assessment of client/patients' satisfaction: development of a general scale. *Evaluation and Program Planning*, **2**, 197–207.

**Leese, M., Johnson, S., Slade, M., et al (1998)** User perspective on needs and satisfaction with mental health services. PRiSM Psychosis Study 8. *British Journal of Psychiatry*, **173**, 409–415.

**McCrone, P., Leese, M., Thornicroft, G., et al (2000)** Reliability of the Camberwell Assessment of Need – European Version. EPSILON Study 6. *British Journal of Psychiatry*, **177** (suppl. 39), s34–s40.

**Merinder, L. B., Viuff, A. G., Laugesen, H. D., et al (1999)** Patient and relative education in community psychiatry: a randomized controlled trial regarding its effectiveness. *Social Psychiatry and Psychiatric Epidemiology*, **34**, 287–294.

**Norusis, G. (1997)** *Statistical Package for Social Sciences (SPSS). Release 7.5*. Chicago, IL: SPSS Inc.

**Oliver, J. (1991)** The Social Care Directive: development of a quality of life profile for use in community services for the mentally ill. *Social Work and Social Sciences Review*, **3**, 5–45.

—, **Huxley, P., Bridges, K., et al (1996)** *Quality of Life and Mental Health Services*. London: Routledge.

**Parkman, S., Davies S., Leese, M., et al (1997)** Ethnic differences in satisfaction with mental health services among representative people with psychosis in South London: PRiSM Study 4. *British Journal of Psychiatry*, **171**, 260–264.

**Phelan, M., Slade, M., Thornicroft, G., et al (1995)** The Camberwell Assessment of Need: the validity and reliability of an instrument to assess the needs of people with severe mental illness. *British Journal of Psychiatry*, **167**, 589–595.

**Ruggeri, M. (1994)** Patients' and relatives' satisfaction with psychiatric services: the state of the art of its measurement. *Social Psychiatry and Psychiatric Epidemiology*, **28**, 212–227.

— & **Dall'Agnola, R. (1993)** The development and use of the Verona Expectations for Care Scale (VECS) and the Verona Service Satisfaction Scale (VSSS) for measuring expectations and satisfaction with community-based psychiatric services in patients,

relatives and professionals. *Psychological Medicine*, **23**, 511–523.

—, —, **Agostini, C., et al (1994)** Acceptability, sensitivity and content validity of VECS and VSSS in measuring expectations and satisfaction in psychiatric patients and their relatives. *Social Psychiatry and Psychiatric Epidemiology*, **29**, 265–276.

— & **Greenfield, T. (1995)** The Italian version of the Service Satisfaction Scale (SSS–30) adapted for community-based psychiatric services: development, factor analysis and application. *Evaluation and Program Planning*, **18**, 191–202.

—, **Dall'Agnola, R., Bisoffi, G., et al (1996)** Factor analysis of the Verona Service Satisfaction Scale-82 and development of reduced versions. *International Journal of Methods in Psychiatric Research*, **6**, 23–38.

**Schene, A. H., van Wijngaarden, B. & Koeter, M. W. J. (1998)** Family caregiving in schizophrenia: domains and distress. *Schizophrenia Bulletin*, **24**, 609–618.

—, **Koeter, M., van Wijngaarden, B., et al (2000)** Methodology of a multi-site reliability study. EPSILON Study 3. *British Journal of Psychiatry*, **177** (suppl. 39), s15–s20.

**van Wijngaarden, B., Schene, A. H., Koeter, M., et al (2000)** Caregiving in schizophrenia: development, internal consistency and reliability of the Involvement Evaluation Questionnaire – European Version. EPSILON Study 4. *British Journal of Psychiatry*, **177** (suppl. 39), s21–s27.

**Ware, J. E., Davies-Avery, A. & Stewart, A. (1978)** The measurement and meaning of patient satisfaction: a

review of the recent literature. *Health and Medical Care Services Review*, **1**, 1–15.

—, **Snyder, M. K., Wright, W. R., et al (1983)** Defining and measuring patient satisfaction with medical care. *Evaluation and Program Planning*, **6**, 247–263.

**Weinstein, R. (1979)** Patient attitudes toward mental hospitals. *Journal of Health and Social Behaviour*, **20**, 237–258.

**World Health Organization (1992)** *Schedules for Clinical Assessment in Neuropsychiatry* (ed.-in-chief J. K. Wing). Geneva: WHO.