Psychological Medicine, 2007, **37**, 85–95. © 2007 Cambridge University Press
doi:10.1017/S0033291706009299 Printed in the United Kingdom

A new scale to assess the therapeutic relationship in community mental health care: STAR

REBECCA McGUIRE-SNIECKUS[1], ROSEMARIE McCABE[1], JOCELYN CATTY[2], LARS HANSSON[3] AND STEFAN PRIEBE[1]*

[1] *Unit for Social and Community Psychiatry, Barts and the London School of Medicine, Queen Mary, University of London, UK*; [2] *Division of Mental Health, St George's, University of London, UK*; [3] *Department of Health Sciences, Lund University, Sweden*

ABSTRACT

Background. No instrument has been developed specifically for assessing the clinician–patient therapeutic relationship (TR) in community psychiatry. This study aimed to develop a measure of the TR with clinician and patient versions using psychometric principles for test construction.

Method. A four-stage prospective study was undertaken, comprising qualitative semi-structured interviews about TRs with clinicians and patients and their assessment of nine established scales for their applicability to community care, administering an amalgamated scale of more than 100 items, followed by Principal Components Analysis (PCA) of these ratings for preliminary scale construction, test–retest reliability of the scale and administering the scale in a new sample to confirm its factorial structure. The sample consisted of patients with severe mental illness and a designated key worker in the care of 17 community mental health teams in England and Sweden.

Results. New items not covered by established scales were identified, including clinician helpfulness in accessing services, patient aggression and family interference. The new patient (STAR-P) and clinician scales (STAR-C) each have 12 items comprising three subscales: positive collaboration and positive clinician input in both versions, non-supportive clinician input in the patient version, and emotional difficulties in the clinician version. Test–retest reliability was $r = 0{\cdot}76$ for STAR-P and $r = 0{\cdot}68$ for STAR-C. The factorial structure of the new scale was confirmed with a good fit.

Conclusions. STAR is a specifically developed, brief scale to assess TRs in community psychiatry with good psychometric properties and is suitable for use in research and routine care.

BACKGROUND

The therapeutic relationship (TR) between a patient and a clinician is at the centre of care delivery in community mental health services. The quality of the TR has been found to predict treatment adherence and outcome across a range of patient diagnoses and treatment settings (Oliver-Martin, 1986; Frank & Gunderson, 1990; Priebe & Gruyters, 1993; Bröker *et al.* 1995; Gaston *et al.* 1998; Martin *et al.* 2000; McCabe & Priebe, 2004) and may even be considered a curative agent in its own right (Catty, 2004). In community psychiatry, community mental health teams provide comprehensive care programmes for people with severe mental illness. Although there is a shared caseload in assertive community treatment (Mueser *et al.* 1998), one named person is usually responsible for keeping in close contact with the patient and coordinating care.

Despite the clinical importance of the TR, no scale has been specifically developed for its assessment in this setting. Although scales have been applied to psychiatric care, they have either been designed for psychotherapy [e.g. Allen's Collaboration Scale (Allen *et al.* 1984); the Barrett-Lennard Relationship Inventory

* Address for correspondence: Professor Stefan Priebe, Unit for Social and Community Psychiatry, Newham Centre for Mental Health, London E13 8SP, UK.
(Email: s.priebe@qmul.ac.uk)

(Barrett-Lennard, 1962); the Psychotherapy Status Report (Frank & Gunderson, 1990); the Working Alliance Inventory (WAI; Horvath & Greenberg, 1986)] or developed on an *ad hoc* basis [e.g. Clarkin's Scale to Assess the Therapeutic Alliance (Clarkin *et al*. 1987); the Helping Alliance Scale (HAS; Priebe & Gruyters, 1993)]. One scale, the Therapist–Patient Scales with Schizophrenic Patients (TPPS; Stark *et al*. 1992), was developed for people with schizophrenia but is limited to capturing aspects of expressed emotion.

Simply importing scales developed for conventional psychotherapy is problematic as the TR in the treatment of severe mental illness is different in many ways from that in conventional psychotherapy. In psychiatric settings, the therapist typically practices in a variable organizational setting including in-patient wards, out-patient clinics, community mental health centres, and the patient's home. The duration of treatment is not fixed and can often last a lifetime. The professional tasks are heterogeneous, spanning treatment, rehabilitation, prevention of relapse and accessing services. Finally, there are statutory responsibilities for care, sometimes requiring compulsory treatment.

Hence, there is a need for an empirically developed and tested scale to measure the TR in this setting (Mueser *et al*. 1998). To this end, we aimed to develop a *S*cale *T*o *A*ssess the Therapeutic *R*elationship in community mental health care (STAR) that is specifically designed for community mental health care, short and easy to administer, has versions for patients and clinicians, good psychometric properties, and, if possible, captures distinct factors.

METHOD

The scale was developed and tested in four stages (see Fig. 1). In stage 1, an item pool was generated. In stage 2, the items from stage 1 were administered to patients and clinicians and then reduced through Principal Components Analysis (PCA). In stage 3, the test–retest reliability of the reduced item pool was tested and the items selected for the final scales. In stage 4, the factorial structure of the scale was tested in a new sample of clinicians and patients. Ethical approval was granted by the relevant research ethics committees.

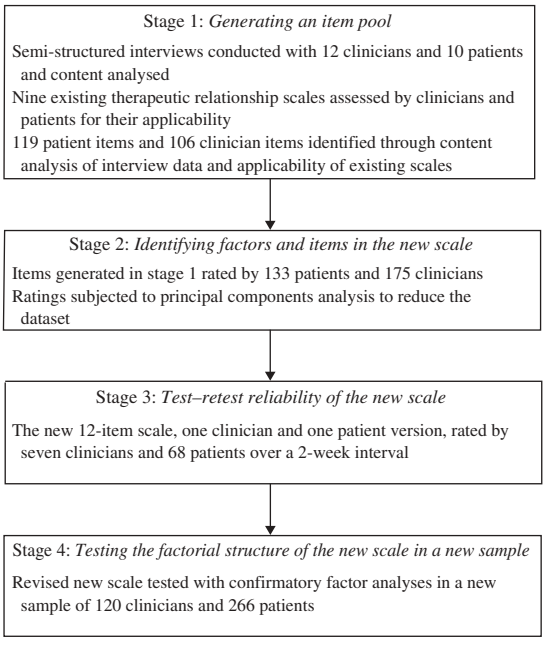

FIG. 1. Flowchart of the four stages in the development of STAR.

Stage 1: Generating an item pool

Two approaches were used to generate an initial item pool. First, semi-structured interviews were conducted with clinicians and patients to explore ideas about the TR in this setting from both perspectives. Ten open-ended questions were asked of the participants. The questions were hypothetical and did not address any specific relationship (e.g. Imagine an ideal/difficult clinician–patient relationship. How would you describe it? What elements would make it ideal/difficult? What actions can a clinician/patient take to help facilitate the development of a good relationship with a patient/clinician?).

The study was presented at three community mental health team (CMHT) meetings and clinicians were invited to participate. The clinician sample was purposive and consisted of 12 clinicians (five social workers, four community psychiatric nurses, two psychologists and one occupational therapist; seven females, five males; six white British, four African-Caribbean, two other ethnic origins). Ten patients (five females, five males; six African-Caribbean, three white British, one Asian) were randomly selected on the basis that they were

aged 18–65, in the care of a CMHT, had severe mental illness, and a designated clinician with lead responsibility for care coordination and delivery. The inclusion criteria for patients and clinicians were identical in stages 1–3 of the study.

Data analysis

The audiotaped and transcribed interviews with clinicians and patients were content analysed using the software package Nudist (N4 Classic). The transcriptions were analysed according to five stages, in alignment with the 'framework' approach to data analysis (Pope *et al.* 2000). A second researcher coded 18% of the transcripts (two patient and two clinician interviews). The inter-rater reliability was good ($\kappa = 0{\cdot}80$, $p < 0{\cdot}01$).

Secondly, after the semi-structured interview, participants were presented with nine existing TR scales to evaluate their applicability to community care. The scales were selected on the basis that they (*a*) have been used previously in at least one empirical study in a psychiatric context and (*b*) do not need to be completed by an expert rater, as this would not be practical in community care settings under routine conditions. The nine scales were: the California Psychotherapy Alliance Scales (CALPAS; Gaston & Marmar, 1991); Allen's five-point professional-rated collaboration scale (Allen *et al.* 1984); Clarkin's six-point scale to assess the therapeutic alliance (Clarkin *et al.* 1987); the Psychotherapy Status Report (Frank & Gunderson, 1990); the Barrett-Lennard Relationship Inventory (Barrett-Lennard, 1962); the WAI (Horvath & Greenberg, 1986); the TPPS (Stark *et al.* 1992), the HAS (Priebe & Gruyters, 1993); and the Engagement and Acceptance Scale (EAS; Park *et al.* 2002).

Stage 2: Identifying factors and items

In stage 2, items generated from stage 1 were amalgamated into one scale for clinicians and one for patients. These scales were administered to a new sample of 26 clinicians and 133 of their patients. Clinicians rated their relationship with named patients (participating in the study) and patients rated their relationship with that particular clinician. Clinicians were recruited from five CMHTs in London. There were 16 community psychiatric nurses, eight social workers, one occupational therapist and one psychologist with an average age of 41 years (54% female; 46% White, 27% Black Caribbean, 15% Black African, 12% Pakistani/Chinese/other ethnic origin).

All clinicians were asked to provide the researcher with the contact details of each person with severe mental illness on their caseload. A total of 481 patients were suggested by clinicians: six were regarded as unsuitable, three were withdrawn from the clinicians' caseload and 13 were listed as residents of hospitals. The contact details of eight patients were incorrect. In total, 451 letters were sent to patients with information about the study and an invitation to participate. Three patients replied by post, and all agreed to participate in the study. Of the possible total, 287 were without a contact telephone number and six listed telephone numbers were incorrect, leaving 155 contactable patients. Of those who were contacted by telephone, 28 refused to participate in the study (non-consent rate of 18%). The total number of people interviewed, all in face-to-face interviews, represented 27% of the total pool.

The average patient age was 40 years; 53% were female; 50% were White, 14% African-Caribbean, 12% Black African, 4% Black other, 7% Indian, 4% Pakistani, 2% Bangladeshi, and 7% other ethnic origins; 48% lived alone. Most patients were diagnosed with schizophrenia (59%) or mood disorder (36%). The mean onset of illness was 21 years before the interview, with an average number of five hospitalizations, and an average of 9 months spent in hospital in total.

Clinicians were asked to complete the corresponding TR questionnaire. If the number of participating patients per clinician was less than 10, the clinician was asked to fill in a corresponding number of questionnaires for randomly selected patients (other patients on their caseload who were not participating in the study but fulfilled the inclusion criteria) that would amount to 10. In total, 133 patient versions and 175 clinician versions were completed.

A researcher who was not involved in treatment interviewed patients and assessed symptoms on the Brief Psychiatric Rating Scale (BPRS; Overall & Gorham, 1962).

Statistical analysis

Patient and clinician ratings on the scales were subjected to a PCA with Varimax rotation to reduce the large data set to a smaller set (Ferguson & Takane, 1989). The number of factors with eigenvalues greater than 3 were noted and those with factor loadings of 0·5 or greater were retained. For subscale construction, factors with three or more items were retained, as well as factors with an acceptable internal consistency (i.e. Cronbach's $\alpha > 0{\cdot}65$). Items for subscales were selected on the basis of internal consistency and predictive validity for the subscale score (regression analysis).

Stage 3: Test–retest reliability and finalizing the scale

The reduced item pool was administered twice to 68 patients with an average age of 43 years (41 female, 27 male; four White, one Black Caribbean, one Chinese, one other ethnic origin), randomly selected from the original sample, and their clinicians ($n = 7$), with a 2-week interval. The mean duration of illness for this sample was 18 years, with an average number of six hospitalizations, five of which were involuntary, and an average of 11 months spent in hospital in total. Most patients were diagnosed with schizophrenia (64%) or mood disorder (35%). The clinician sample included three social workers, three community psychiatric nurses and one psychologist (five female, two male). The mean caseload for this sample was 21.

Statistical analysis

A regression analysis was conducted to test how much of the variance of the previous reduced item set was explained by the remaining items after the final reduction.

Stage 4: Testing the scale in a new sample

The final scale was administered to a new sample of 266 community mental health care patients and their 120 clinicians in London (UK) and Lund (Sweden). Clinician (STAR-P) and patient scales (STAR-C) were translated into Swedish and then back-translated. Patients were aged 18–65, had a psychotic illness, had been ill for at least 2 years and on the caseload of a CMHT for at least 3 months. Their mean age was 42·4 years (S.D. = 11·6; 61% male; 66·5% White British/White European, 22·9% Black British/Black African/Black Caribbean, 7·2% Asian British/Asian, 3·4% other origins). Most patients were diagnosed with schizophrenia (67%) or mood disorder (13%). The average length of illness was 18·3 years (S.D. = 11·4). The mean age of the clinician sample was 45·8 years (S.D. = 9·8; 63% female; 59% White British/White European, 24% Black British/Black African/Black Caribbean, 15% Asian British/Asian, 2% other origins). Most were community psychiatric nurses (68%), followed by social workers (17%), occupational therapists (8%), psychologists (3%) and psychiatrists (1%).

Statistical analysis

Confirmatory factor analysis models were fitted to both the new clinician and new patient data, using the sem package and *R* version 2.2.0 (Ihaka & Gentleman, 1996) to test the hypothesis that the new data fit the stage 2 theoretical model. In each case, a model with three factors derived from stage 2 was fitted.

RESULTS

Stage 1

The content analysis of the interview transcripts (clinician and patient transcripts were analysed separately) revealed a number of concepts that may be relevant to any relationship between a helping professional and a patient (e.g. trust, respect, openness, commitment). Clinicians and patients also reported aspects that may be specific to psychiatric or community care settings, some of which were not covered by existing scales. For example, both clinicians and patients reported the clinician's helpfulness in accessing other services and benefits, and patient aggression. Clinicians emphasized family interference, patient's trust, patient's willingness to engage, and shared realistic expectations of progress. Patients stressed clinician's reliability, support, open communication, and their own willingness to accept treatment. Thirty-two items were mentioned by at least four participants and a further nine items were mentioned by nine participants or more. These were retained as 'new' items to be administered in stage 2.

Based on clinicians' and patients' assessments of their applicability to their relationships, five

of the nine existing scales were also retained for further testing in stage 2: the HAS (six items for patients, five for clinicians), WAI (12 items each for patients and clinicians), TPPS (30 items for patients, 16 for clinicians), CALPAS (24 items each), and EAS (four items for clinicians). Four of the nine scales were deemed unsuitable; the reasons are available from the authors on request.

A global assessment of the TR was added for patients and clinicians and, given the relevance of systemic relationships in this setting, a two-part question on the relationship between clinicians and significant others for patients only (Priebe, 1989). Thus, the final item pool generated in stage 1 consisted of 119 items for patients (one global rating, 72 items from existing scales, two systemic items, 41 new items and three open-ended questions) and 106 for clinicians (one global rating, 61 items from existing scales, 41 new items and three open-ended questions). The 119 items were put into one amalgamated scale to be administered to patients in the next stage and the 106 items were put into one amalgamated scale to be administered to clinicians.

Stage 2

For the clinician version, the PCA revealed six factors with an eigenvalue $\geqslant 3$ explaining 53·2% of the variance. For subscale construction, factors with three or more items were retained, as were factors with an acceptable internal consistency ($\alpha > 0{\cdot}65$), in accordance with Ferguson & Takane (1989) and Nunnally & Bernstein (1994). One factor was dropped because it had only two items and another was dropped because of low internal consistency ($\alpha = 0{\cdot}17$). The summative scales formed from selected items on each component were correlated to check if there was an overlap. The summative scales of the first and fourth factor were correlated (subscale correlation: $r = 0{\cdot}20$, $p < 0{\cdot}01$). Therefore, factor 4 was also dropped. Table 1 shows the three retained factors and the item loadings.

The first factor comprised items relating to a 'positive collaborative' TR. Six items accounted for 83% of the variance in factor 1 (adjusted R^2) with an α coefficient of 0·94. The second factor consisted of items relating to 'emotional difficulties' of the clinician. Five items accounted for 82% of the variance of factor 2 with $\alpha = 0{\cdot}88$. The third factor captured aspects of 'positive clinician' input. Three items accounted for 65% of the variance in factor 3 with $\alpha = 0{\cdot}73$.

Patient data were analysed using the above-mentioned criteria. The PCA also showed six factors with an eigenvalue $\geqslant 3$ explaining 57·3% of the variance. Applying the same criteria as in the analysis of clinician data, one factor was dropped because it had only two items, and another was dropped because of low internal consistency ($\alpha = 0{\cdot}58$). The summative scales of the second and fifth factors were correlated so factor 5 was also dropped. Table 2 shows the three retained factors and the item loadings.

Again, the first subscale consisted of items reflecting a 'positive collaborative' TR. Six items accounted for 88% of the variance in factor 1 (adjusted R^2) with $\alpha = 0{\cdot}91$. The second subscale comprised items relating to 'positive clinician input'. Five items explained 62% of the variance with $\alpha = 0{\cdot}86$. The third subscale consisted of items relating to 'non-supportive clinician input'. Five items accounted for 71% of the variance in factor 3 with $\alpha = 0{\cdot}76$.

Stage 3

The test–retest reliability for all items of the reduced item pool ranged from $r = 0{\cdot}44$ to 0·73 ($p < 0{\cdot}05$) for STAR-C items and from $r = 0{\cdot}44$ to 0·80 ($p < 0{\cdot}05$) for STAR-P items. The reliability of items retained in the scale ranged from $r = 0{\cdot}46$ to 0·73 for STAR-C and from $r = 0{\cdot}52$ to 0·80 for STAR-P. Aiming to develop a brief scale and considering the minimum number of three items for a subscale (Burnett *et al.* 1997), the number of items in each version was reduced to 12, that is six for the first factor of 'positive collaboration' and three for each for the other two factors, and items with the lowest test–retest reliability on each factor were dropped.

The correlation (Pearson's r) between the total and subscale scores of the two versions revealed a significant, negative association between clinician ratings of their emotional difficulties and patient total ratings ($r = -0{\cdot}33$, $p < 0{\cdot}05$), patient-rated 'positive collaboration' ($r = -0{\cdot}34$, $p < 0{\cdot}05$), and patient-rated 'positive clinician input' ($r = -0{\cdot}34$, $p < 0{\cdot}05$). All other correlations failed to reach statistical

Table 1. *Clinician Principal Component Analysis*

Item	Factor loading	α	Item	Factor loading	α
Factor 1: 'Positive collaboration'			**Factor 2: 'Emotional difficulty'**		
24·67% of variance, eigenvalue = 25·41			9·29% of variance, eigenvalue = 9·57		
Trust (NEW)	0·86		Inferiority to P (TPPS)	0·83	
Global assessment	0·85		Not give instructions P understands (CALPAS)	0·75	
P likes me (WAI)	0·82		How P feels about clinician (HEAS)	0·75	
Gets along with P (HAS)	0·82		Cannot empathize (TPPS)	0·67	
Rapport (NEW)	0·81		C criticism (NEW)	0·66	
Trust (NEW)	0·8		Not feel accepted by P (TPPS)	0·66	
C empathy (NEW)	0·8		P found medication difficult (CALPAS)	0·65	
Open communication (NEW)	0·79		P attitude towards help (HEAS)	0·65	
P respect for C ability (NEW)	0·78		C irritated, annoyed, disappointed (CALPAS)	0·65	
Looks forward to seeing P (HAS)	0·78		Would prefer to transfer P (TPPS)	0·63	
Actively involved (HAS)	0·77		P goals differ from C (CALPAS	0·62	
Built mutual trust (WAI)	0·76		P difficulty ask questions re medication (CALPAS)	0·6	
C right one for P (NEW)	0·76		C dislike of P (TPPS)	0·54	0·71
Respect for P (NEW)	0·75		**Factor 3: 'Positive clinician factor'**		
P openness (NEW)	0·75		3·25% of variance, eigenvalue = 3·35		
Can help (HAS)	0·73		C takes perspective of P (NEW-Y/N)	0·71	
C takes P perspective (NEW)	0·72		C listens to P (NEW-Y/N)	0·67	
C reliability (NEW)	0·72		C is supportive (NEW-Y/N)	0·66	0·75
Ability to help P (NEW)	0·7				
P commitment (NEW)	0·69				
Confidence can help (WAI)	0·68				
P willing to work with C (NEW)	0·68				
C accessibility (NEW)	0·67				
P agency (NEW)	0·67				
C flexibility (NEW)	0·66				
C patience (NEW)	0·64				
P disclosure to C (NEW)	0·63				
Free will of P (NEW)	0·63				
C listens to P (NEW)	0·62				
Shared expectations (NEW)	0·62				
Way working correct (WAI)	0·62				
Understand changes needed (WAI)	0·62				
Agree how to improve (WAI)	0·61				
Desire to understand P (CALPAS)	0·61				
Work on same goals (WAI)	0·59				
C frequency of contact (NEW)	0·58				
Appreciate P as person (WAI)	0·57				
P trust (NEW)	0·57				
Confidence to help (CALPAS)	0·56				
Agree what to work on (WAI)	0·55				
Degree P engaged (HEAS)	0·51				
Help P see difficulties differently (CALPAS)	0·51	0·95			

P, Patient; C, clinician; WAI, Working Alliance Inventory; HAS, Helping Alliance Scale; CALPAS, California Psychotherapy Alliance Scales; TPPS, Therapist Patient Scales with Schizophrenic Patients; HEAS, Homelessness Engagement and Acceptance Scale; NEW, new items; NEW-Y/N, new yes/no items.

significance. The test–retest reliability for the final scales revealed positive correlations for the items and mean total and subscale scores (Table 3).

A regression analysis showed that the items of the final scale explained 87% of the variance of the previous reduced item pool in the clinician version, and 94% in the patient version (adjusted R^2).

The clinician version (STAR-C) and the patient version (STAR-P) are presented in the Appendix. Completing the scale usually takes 5 minutes or less. Scores can be obtained for the total scale and subscales. In the sample in which the scale was developed, the mean sum score of STAR-C was 31·5 (S.D. = 6·9). On the subscales, mean sum scores were 15·3 (4·0) for 'positive collaboration', 7·4 (2·7) on 'emotional difficulties, and 8·9 (1·6) on 'positive clinician input'. On STAR-P, the mean sum score was 38·4 (12·0) for the total scale, 19·9 (6·7) for 'positive collaboration', 9·3 (3·0) for 'positive clinician

Table 2. *Patient Principal Component Analysis*

Item	Factor loading	α
Factor 1: 'Positive collaboration'		
28·51% of variance, eigenvalue = 33·10		
Confidence in C ability to help (WAI)	0·88	
Understanding of changes needed (WAI)	0·86	
Open communication (NEW)	0·86	
Feel appreciated (WAI)	0·85	
Agree what to work on (WAI)	0·85	
Feel supported by C (NEW)	0·84	
Honesty (NEW)	0·84	
P trust (NEW)	0·83	
P openness (NEW)	0·83	
Working towards mutual goals (WAI)	0·82	
C helpful (NEW)	0·82	
C patience (NEW)	0·8	
Trust in C's competence (HAS)	0·79	
C listens (NEW)	0·79	
Willing to work with C (NEW)	0·79	
Global assessment	0·79	
Mutual trust (WAI)	0·79	
C frequency of contact (NEW)	0·78	
Feels respected by C (NEW)	0·78	
C likes me (WAI)	0·77	
P commitment (NEW)	0·77	
C reliable (NEW)	0·76	
C takes perspective (NEW)	0·75	
C right one for P (NEW)	0·75	
New ways of looking at problem (WAI)	0·75	
Agree what to do (WAI)	0·74	
C empathy (NEW)	0·73	
C availability (NEW)	0·73	
P feels understood (HAS)	0·73	
C sensitivity to cultural background (NEW)	0·72	
P feels respected by C (HAS)	0·71	
C approachable (NEW)	0·7	
Rapport (NEW)	0·69	
C desire to understand P (CALPAS)	0·68	
Way working on problem correct (WAI)	0·66	
P feels free to express worries (CALPAS)	0·65	
C listens to P (NEW-Y/N)	0·6	
C supportive (NEW-Y/N)	0·6	
C gives satisfactory answers (CALPAS)	0·6	
C understands what P wants (CALPAS)	0·58	
Last appointment important (CALPAS)	0·56	
Free will of P (NEW)	0·56	
Trust (NEW-Y/N)	0·55	
Different goals (CALPAS)	0·52	
C positive feedback (TPPS)	0·52	
How P feels after seeing C (HAS)	0·52	
P respect for C professional ability (NEW)	0·52	
C takes perspective (NEW-Y/N)	0·51	0·98

Item	Factor loading	α
Factor 2: 'Positive clinician input'		
8·56% of variance, eigenvalue = 9·23		
C encouragement (TPPS)	0·7	
C helpful (TPPS)	0·64	
C regard (TPPS)	0·64	
C understanding (TPPS)	0·63	
C discuss P goals (TPPS)	0·63	
C allows open conversation (TPPS)	0·62	
Trust (TPPS)	0·62	
C understanding (TPPS)	0·6	
C perceptiveness (TPPS)	0·58	
C positive regard (TPPS)	0·55	
P would prefer another C (TPPS)	0·55	0·91
Factor 3: 'Non-supportive clinician input'		
6·37% of variance, eigenvalue = 7·39		
C withholds truth (TPPS)	0·66	
C overwhelms (TPPS)	0·61	
C not understand what P wants (WAI)	0·56	
C empathy (TPPS)	0·54	
C impatience (TPPS)	0·53	
C pressure (TPPS)	0·51	
C authoritarianism (TPPS)	0·51	0·79

C, Clinician; P, patient; WAI, Working Alliance Inventory; HAS, Helping Alliance Scale; CALPAS, California Psychotherapy Alliance Scales; TPPS, Therapist Patient Scales with Schizophrenic Patients; HEAS, Homelessness Engagement and Acceptance Scale; NEW, new items; NEW-Y/N), new yes/no items.

input', and 9·3 (3·3) for 'non-supportive clinician input'.

Stage 4

Fitting the stage 2 three-factor model to the new patient data resulted in a $\chi^2 = 153{\cdot}87$, df = 51, $p < 0{\cdot}001$. The goodness-of-fit index (GFI) was 0·91. The root-mean-square error of approximation (RMSEA) index was 0·09 [90% confidence interval (CI) 0·08–0·11]. A GFI of 0·90, or above, and RMSEA index of 0·08, or below, are generally regarded as satisfactory. Fitting the stage 2 three-factor model to the new clinician data resulted in a $\chi^2 = 107{\cdot}92$, df = 51, $p < 0{\cdot}001$.

Table 3. *Test–retest reliability of the sum and subscale scores for STAR-C and STAR-P*

	Test–retest correlation coefficient, R (all significant at $p<0{\cdot}05$)
STAR-C	
Sum score	0·68
Positive collaboration subscale	0·72
Clinician emotional difficulties subscale	0·58
Positive clinician input subscale	0·73
STAR-P	
Sum score	0·76
Positive collaboration subscale	0·78
Positive clinician input subscale	0·81
Non-supportive clinician input subscale	0·68

The GIF was 0·88 and the RMSEA index was 0·07 (90% CI 0·05–0·09).

Final scales association with sociodemographic and clinical characteristics

Total and subscale scores of both STAR versions were investigated in relation to sociodemographic characteristics of clinicians and patients, and patient symptom levels. As a single clinician was rated by a number of patients, there is a cluster effect that could lead to overestimation of correlations. Hence, aggregate means were calculated for the total STAR-P and subscale scores (i.e. an average STAR score per clinician based on their patients' ratings) to investigate their relationship with clinician characteristics.

Clinicians' (i.e. STAR-C) ratings were neither significantly associated with their own sociodemographic characteristics nor with those of their patients with the exception of patient age; clinicians rated less 'emotional difficulties' ($r=0{\cdot}16$, $p<0{\cdot}01$) and more 'positive clinician input' ($r=0{\cdot}15$, $p<0{\cdot}05$) with older patients. There were less favourable ratings of the TR with patients who had higher BPRS scores on the subscales 'emotional difficulties' ($r=-0{\cdot}2$, $p<0{\cdot}005$) and 'positive clinician input' ($r=-0{\cdot}15$, $p<0{\cdot}02$).

Patients' (i.e. STAR-P) ratings were related to patient sex, with female patients rating 'positive clinician input' more highly ($t=-2{\cdot}1$, $p<0{\cdot}05$). Patients with higher symptom levels had lower total STAR-P scores ($r=-0{\cdot}14$, $p<0{\cdot}05$) and higher (i.e. worse) scores on the 'non-supportive clinician input' subscale ($r=-0{\cdot}14$, $p<0{\cdot}05$).

DISCUSSION

STAR has been specifically developed to assess the relationship between multidisciplinary clinicians and patients with severe mental illness in community care settings. In line with psychometric theory of psychological test construction, a rigorous and systematic development process was conducted, through all stages from item generation in open interviews, preliminary test construction, assessment of test–retest reliability, to testing the factorial structure of the scale in a new sample. The scale had to be brief and easy to administer so that it can be realistically applied in community mental health care practice.

The new scale's psychometric properties such as internal consistency and test–retest reliability are acceptable and the original factorial structure was confirmed when the scale was tested in a new sample that included an international subsample. The fact that only a few of the clinician and patient ratings of their relationship were weakly correlated is in line with the extensive research in psychotherapy settings and new research in psychiatry (Couture *et al.* 2006), which also shows that patients and clinicians may perceive their relationship differently, with only weak to moderate associations between the two ratings (Horvath & Greenberg, 1986; Marmar *et al.* 1986). The predictive validity of the new scale, that is its association with factors such as treatment adherence, admissions and symptom severity, remains to be explored in future research.

Neither clinicians' nor patients' ratings were strongly associated with sociodemographic characteristics of clinicians and patients, although some weak correlations were found for patients' assessments. With respect to clinical characteristics, McCabe & Priebe (2003) reported that psychopathology may account for 3–28% of the variance in patient ratings of the relationship depending on the specific sample and treatment setting. In this sample, patients' ratings of STAR were weakly related to symptom levels, and clinicians tended to assess relationships less favourably when patients were more symptomatic. The latter finding has also

been reported by Frank & Gunderson (1990), Neale & Rosenheck, (1995), Klinkenberg *et al.* (1998) and Couture *et al.* (2006). Although symptoms consistently influence the TR, it may be concluded that STAR scores are not dominated by the degree of patient symptom levels.

Clinicians and patients distinguished between different aspects of the TR, and the factors are somewhat different than those found in relationships in psychotherapy. STAR captures three similar factors in each version. The first subscale, 'positive collaboration', reflects a good rapport, a shared understanding of goals and the experience of mutual openness and trust. For clinicians and patients alike, this factor explained most of the variance of the original item pool and might capture the general quality of the relationship, the 'chemistry' between the two people and the overall degree to which the relationship works. As such, it might be difficult to influence directly through skills training. 'Positive clinician input' reflects to what extent clinicians (is perceived by the patient to) encourage, regard, support, listen to and understand the patient. This factor is characterized by more behavioural aspects, which might be easier to modify through training and supervision of clinicians. Finally, 'emotional difficulties' in the clinicians' rating and 'non-supportive clinician input' in the patients' assessment reflect problems in the relationship such as the clinician's feeling that they cannot empathize with and are not accepted by the patient, and the patient's perception that the clinician withholds the truth and is impatient and authoritarian. While such feelings are clearly not helpful in establishing or maintaining a positive relationship, they are important to identify and could be addressed in ongoing clinical supervision. Further research might identify the extent to which each of these aspects can be affected through specific clinical interventions or, possibly, changing the clinician in the case of a very unfavourable TR.

Some limitations of this study should be considered. First, the scale was necessarily developed and tested within a selective sample, that is those patients who agreed to take part in the research. The most difficult to engage patients are unlikely to participate in such research, and STAR might not adequately capture the views of that patient group. Second, a larger sample size would have been preferable for the PSA in the development of the scale. However, the factorial structure of the scale was confirmed in the new sample across different settings. Third, the scale was developed in a deprived multi-ethnic inner city area and although it was validated in different areas (with respect to socio-economic and urban–rural conditions), it remains to be tested outside the context of Western/Northern European health-care systems. Finally, the new scale's responsiveness to change also remains to be tested in other studies and health-care systems.

Although STAR measures three distinct aspects of the TR in community mental health care in a clinician and a patient version, it is brief and simple to use. The versions have been developed separately, but capture similar aspects from different perspectives. The scale can be used in research and routine clinical practice. In research it may be applied to assess the quality of the TR as an outcome criterion in its own right or as a mediating factor explaining variance in surveys and trials. In practical care, identifying the quality of a TR may have implications for clinical decisions as well as for professional education and training. This, in turn, might impact on the patients' experience of TR in mental health care, which is viewed by patients as the most crucial factor in good psychiatric care (Johansson & Eklund, 2003).

ACKNOWLEDGEMENTS

Funding was provided by the Barts Joint Research Board, the UK National Coordinating Centre for Service Delivery and Organization and the Research Directorate of the European Commission within the Framework Programme 5 (QLG5-CT-2002-01938).

DECLARATION OF INTEREST

None.

REFERENCES

Allen, J. G., Newsom, G. E., Gabbard, G. O. & Coyne, L. (1984). Scales to assess the therapeutic alliance from a psychoanalytic perspective. *Bulletin of the Menninger Clinic* **48**, 383–400.

Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs* **76**, 1–35.

Bröker, M., Röhricht, F. & Priebe, S. (1995). Initial assessment of hospital treatment by patients with paranoid schizophrenia: a predictor of outcome. *Psychiatry Research* **58**, 77–81.

Burnett, P., Middleton, W., Raphael, B. & Martinek, N. (1997). Measuring core bereavement phenomena. *Psychological Medicine* **27**, 49–57.

Catty, J. (2004). 'The vehicle of success': theoretical and empirical perspectives on the therapeutic alliance in psychotherapy and psychiatry. *Psychology and Psychotherapy* **77**, 255–272.

Clarkin, J. F., Hurt, S. W. & Crilly, J. L. (1987). Therapeutic alliance and hospital treatment outcome. *Hospital and Community Psychiatry* **38**, 871–875.

Couture, S. M., Roberts, D. L., Penn, D. L., Cather, C., Otto, M. W. & Goff, D. (2006). Do baseline client characteristics predict the therapeutic alliance in the treatment of schizophrenia? *Journal of Nervous and Mental Disease* **194**, 10–14.

Ferguson, G. A. & Takane, Y. (1989). *Statistical Analysis in Psychology and Education*. McGraw-Hill: New York.

Frank, A. F. & Gunderson, J. G. (1990). The role of the therapeutic alliance in the treatment of schizophrenia. *Archives of General Psychiatry* **47**, 228–235.

Gaston, L. & Marmar, C. R. (1991). *Manual for the California Psychotherapy Alliance Scales – CALPAS*. Unpublished manuscript. Department of Psychiatry, McGill University: Montreal, Canada.

Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L. G. & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioural, cognitive, and brief dynamic therapy. *Psychotherapy Research* **8**, 190–209.

Horvath, A. O. & Greenberg, L. (1986). The development of the Working Alliance Inventory. In *Psychotherapeutic Processes: A Research Handbook* (ed. L. Greenberg and W. Pinsoff), pp. 529–556. Guilford Press: New York.

Ihaka, R. & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299–314.

Johansson, H. & Eklund, M. (2003). Patients' opinion on what constitutes good psychiatric care. *Scandinavian Journal of Caring Sciences* **17**, 339–346.

Klinkenberg, W. D., Calsyn, R. J. & Morse, G. A. (1998). The helping alliance in case management for homeless persons with severe mental illness. *Community Mental Health Journal* **34**, 569–578.

Marmar, C. R., Horowitz, M. J., Weiss, D. S. & Marziali, E. (1986). The development of the Therapeutic Alliance Rating System. In *Psychotherapeutic Processes: A Research Handbook* (ed. L. Greenberg and W. Pinsoff), pp. 285–324. Guilford Press: New York.

Martin, D. J., Garske, J. P. & Davis, M. K. (2000). Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *Journal of Consulting and Clinical Psychology* **68**, 438–450.

McCabe, R. & Priebe, S. (2003). Are therapeutic relationships in psychiatry explained by patients' symptoms? Factors influencing patient ratings. *European Psychiatry* **18**, 220–225.

McCabe, R. & Priebe, S. (2004). The therapeutic relationship in the treatment of severe mental illness: a review of methods and findings. *International Journal of Social Psychiatry* **50**, 115–128.

Mueser, K. T., Bond, G. R., Drake, R. E. & Resnick, S. G. (1998). Models of community care for severe mental illness: a review of research on case management. *Schizophrenia Bulletin* **24**, 37–74.

Neale, M. S. & Rosenheck, R. A. (1995). Therapeutic alliance and outcome in a VA intensive case management programme. *Psychiatric Services* **46**, 719–721.

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory*. McGraw-Hill: New York.

Park, M. J., Tyrer, P., Elsworth, E., Fox, J., Ukoumunne, O. C. & MacDonald, A. (2002). The measurement of engagement in the homeless mentally ill: the Homeless Engagement and Acceptance Scale – HEAS. *Psychological Medicine* **32**, 855–861.

Pope, C., Ziebland, S. & Mays, N. (2000). Qualitative research in health care: analysing qualitative data. *British Medical Journal* **320**, 114–116.

Priebe, S. (1989). Can patients' views of a therapeutic system predict outcome? An empirical study with depressive patients. *Family Process* **28**, 349–355.

Priebe, S. & Gruyters, T. (1993). The role of the helping alliance in psychiatric community care: a prospective study. *Journal of Nervous and Mental Disease* **181**, 552–557.

Olivier-Martin, R. (1986). Psychological factors, observation and resistance to antidepressive treatments [in French]. *L'Encephale* **7**, 197–203.

Overall, J. & Gorham, D. R. (1962). The Brief Psychiatric Rating Scale. *Psychological Reports* **10**, 799–812.

Stark, F. M., Lewandowski, L. & Buchkremer, G. (1992). Therapist–patient relationship as a predictor of the course of schizophrenic illness. *European Psychiatry* **7**, 101–169.

APPENDIX. *S*cale *T*o *A*ssess Therapeutic *R*elationships in Community Mental Health Care (STAR)

STAR-C: Clinician Version[a]	STAR-P: Patient Version[a]
1. I get along well with my patient.	1. My clinician speaks with me about my personal goals and thoughts about treatment.
2. My patient and I share a good rapport.	2. My clinician and I are open with one another.
3. I listen to my patient.	3. My clinician and I share a trusting relationship.
4. I feel that my patient rejects me as a clinician.	4. I believe my clinician withholds the truth from me.
5. I believe my patient and I share a good relationship.	5. My clinician and I share an honest relationship.
6. I feel inferior to my patient.	6. My clinician and I work towards mutually agreed upon goals.
7. My patient and I share similar expectations regarding his/her progress in treatment.	7. My clinician is stern with me when I speak about things that are important to me and my situation.
8. I feel that I am supportive of my patient.	8. My clinician and I have established an understanding of the kind of changes that would be good for me.
9. It is difficult for me to empathize with or relate to my patient's problems.	9. My clinician is impatient with me.
10. My patient and I are open with one another.	10. My clinician seems to like me regardless of what I do or say.
11. I am able to take my patient's perspective when working with him/her.	11. We agree on what is important for me to work on.
12. My patient and I share a trusting relationship.	12. I believe my clinician has an understanding of what my experiences have meant to me.

[a] Rate each item on the following scale:

Never	Rarely	Sometimes	Often	Always
0	1	2	3	4

Scoring protocol

A total STAR-C score and three subscale scores can be obtained. Before scoring, scores for the Emotional Difficulties subscale are reversed. Subtract each of the item ratings in this subscale from 4: a rating of 0 becomes 4 (4 − 0); a rating of 1 becomes 3 (4 − 1); a rating of 2 remains 2 (4 − 2); a rating of 3 becomes 1 (4 − 3); and a rating of 4 becomes 0 (4 − 4). After reversing items for this subscale, the total STAR-C score is obtained by adding the scores for each of the 12 items (range 0–48). The three subscale scores are each obtained by summing the relevant subscale items as follows:

Positive Collaboration: 1, 2, 5, 7, 10, 12
Emotional Difficulties: 4, 6, 9
Positive Clinician Input: 3, 8, 11

A total STAR-P score and three subscale scores can be obtained. Before scoring, scores for the Non-Supportive Clinician Input subscale are reversed. Subtract each of the item ratings in this subscale from 4: therefore, a rating of 0 becomes 4 (4 − 0); a rating of 1 becomes 3 (4 − 1); a rating of 2 remains 2 (4 − 2); a rating of 3 becomes 1 (4 − 3); and a rating of 4 becomes 0 (4 − 4). After reversing, the total STAR-P score is obtained by adding the scores for each of the 12 items (range 0–48). The three subscale scores are obtained by summing the relevant subscale items as follows:

Positive Collaboration: 2, 3, 5, 6, 8, 11
Positive Clinician Input: 1, 10, 12
Non-Supportive Clinician Input: 4, 7, 9