

# Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR)

Julie Williams, Mary Leamy, Francesca Pesola, Victoria Bird, Clair Le Boutillier and Mike Slade

## Background

Supporting recovery is the aim of national mental health policy in many countries. However, only one measure of recovery has been developed in England: the Questionnaire about the Process of Recovery (QPR), which measures recovery from the perspective of adult mental health service users with a psychosis diagnosis.

## Aims

To independently evaluate the psychometric properties of the 15- and 22-item versions of the QPR.

## Method

Two samples were used: data-set 1 ( $n=88$ ) involved assessment of the QPR at baseline, 2 weeks and 3 months. Data-set 2 ( $n=399$ ; trial registration: ISRCTN02507940) involved assessment of the QPR at baseline and 1 year.

## Results

For the 15-item version, internal consistency was 0.89,

convergent validity was 0.73, test–retest reliability was 0.74 and sensitivity to change was 0.40. Confirmatory factor analysis showed the 15-item version offered a good fit. For the 22-item version, the interpersonal subscale was found to underperform and the intrapersonal subscale overlaps substantially with the 15-item version.

## Conclusions

Both the 15-item and the intrapersonal subscale of the 22-item versions of the QPR demonstrated satisfactory psychometric properties. The 15-item version is slightly more robust and also less burdensome, so it can be recommended for use in research and clinical practice.

## Declaration of interest

None.

## Copyright and usage

© The Royal College of Psychiatrists 2015.

Personal recovery has been defined as ‘a deeply personal, unique process of changing one’s attitudes, values, feelings, goals, skills, and/or roles. It is a way of living a satisfying, hopeful, and contributing life even with limitations caused by illness.’<sup>1</sup> Supporting personal recovery has become a mental health policy goal in many countries.<sup>2</sup> Research in recent years has helped to further define recovery<sup>3</sup> and propose a conceptual framework of recovery-oriented practice.<sup>4</sup> Measures are required to evaluate and assess how staff and services can support recovery, from the viewpoint of people who use these services. Outcome measures have been developed to assess personal recovery.<sup>5,6</sup> Only one measure of recovery has been developed in England – the Questionnaire about the Process of Recovery (QPR).<sup>7</sup> Initial psychometric evaluation of the QPR by its developers showed good internal consistency, construct validity and test–retest reliability. An exploratory factor analysis identified two factors that were labelled ‘intrapersonal’ and ‘interpersonal’. A subsequent re-evaluation of the QPR, conducted by its developers, found similar results and a 15-item one-factor solution.<sup>8</sup> The current study was undertaken to investigate whether these findings could be replicated in a different sample, and whether the original 22-item version or the 15-item version can be recommended. The aims were to independently evaluate the internal consistency, convergent validity, 2-week test–retest reliability, 3-month and 12-month sensitivity to change of both the 22- and 15-item versions of the QPR, and to explore the factor structure of both versions.

## Method

### Design

Data from two studies were used. Data-set 1 came from a psychometric evaluation study in South London, with data collected between March 2011 and May 2012. Data-set 2 is pooled

baseline and 1-year follow-up data from a cluster randomised controlled trial of a pro-recovery intervention<sup>9</sup> (trial registration: ISRCTN02507940), with data from each team (cluster) collected between April 2011 and December 2013. Ethical approval was obtained for both studies.

### Participants

Participants were recruited from adult community mental health teams. Inclusion criteria were (a) aged between 18 and 65, (b) well enough to participate and able to complete measures in the view of staff, (c) able to give informed consent, and (d) able to speak and understand English well enough to complete the measures. Additional inclusion criteria for data-set 1 was any mental disorder and for data-set 2 was a diagnosis of psychosis.

### Procedures

Participants were recruited via their care coordinator from community adult mental health teams in South London (data-sets 1 and 2) and Gloucestershire (data-set 2). All participants received payment of £10 for each round of data collection.

Data-set 1 comprised a convenience sample with care coordinators identifying people on their case-load who matched the inclusion criteria and who would be willing to participate. Willing participants were then contacted by a researcher who explained the study and answered any questions, before taking informed consent. Participants completed measures with a researcher at three time points: baseline, 2 weeks and 3 months. At baseline, participants completed a battery of measures including the QPR (22-item version), the Recovery Assessment Scale (RAS)<sup>10</sup> and the Warwick-Edinburgh Mental Well-being Scale (WEMWBS).<sup>11</sup> Two weeks later, they completed the QPR. Three months after baseline they completed the QPR and WEMWBS. If participants did not complete the measures at the

2-week point, they were still invited to complete the measures at the 3-month point.

For data-set 2, the case-load for each team was randomly ordered. Researchers then contacted the care coordinator of each randomised person in sequence until the required 15 participants per team were recruited. The recruitment procedure was as per data-set 1. Participants completed an extensive assessment battery including the QPR (22-item version) and WEMWBS at baseline. One year later participants completed the same assessment battery including the QPR and WEMWBS.

## Measures

The original QPR is a 22-item, service user-rated measure of personal recovery developed in the UK.<sup>7</sup> The measure was developed from a qualitative study led by service user–researchers. Each item is scored using a 5-point Likert scale ranging from 0 (disagree strongly) to 4 (agree strongly). The initial version comprised two subscales: QPR intrapersonal (17 items) (range 0–68) and QPR interpersonal (5 items) (range 0–20), with higher scores indicating increased recovery on both subscales. Adequate internal consistency (intrapersonal  $\alpha=0.94$ , interpersonal  $\alpha=0.77$ ), construct validity and test–retest reliability (intrapersonal  $r=0.87$ , interpersonal  $r=0.76$ ) were shown. A subsequent evaluation by the developers of the psychometric properties using a new data-set found a 15-item (range 0–60) one-factor solution called QPR total, which demonstrated adequate internal consistency ( $\alpha=0.93$ ) and test–retest reliability using Pearson's correlation ( $r=0.70$ ).<sup>8</sup> In our study both data-sets were collected using the 22-item QPR, with the 15-item QPR total score being extrapolated. In this analysis we refer to the two subscales of the 22-item QPR as QPR intrapersonal and QPR interpersonal, and the 15-item QPR as QPR total.

The RAS is a 41-item service user-rated measure assessing five domains of recovery: personal confidence and hope, willingness to ask for help, goal and success orientation, reliance on others and no domination by symptoms.<sup>10</sup> Each item is scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree) with the total score ranging from 41 (low recovery) to 205. Good internal consistency of  $\alpha=0.93$  and test–retest reliability using Pearson's correlation of  $r=0.88$  have been demonstrated.<sup>10</sup>

The WEMWBS is a 14-item self-report measure assessing well-being.<sup>11</sup> Respondents rate their experience regarding each statement over the past 2 weeks. Each item is scored using a 5-point Likert scale ranging from 1 (none of the time) to 5 (all of the time) with the total score ranging from 14 (low well-being) to 70. In the initial validation study, good content validity, internal consistency of  $\alpha=0.89$ , and test–retest reliability using intraclass correlation coefficients (ICC) of  $r=0.83$  were demonstrated.<sup>11</sup> The measure has also been validated with adolescent and minority ethnic groups.<sup>12</sup>

## Analysis

Using data-set 1, convergent validity was assessed using Pearson's correlation between RAS and QPR intrapersonal, QPR interpersonal and QPR total at baseline. Test–retest reliability was assessed by exploring agreement at the individual level using two-way random effects ICCs between QPR intrapersonal, QPR interpersonal and QPR total at baseline and 2 weeks. Sensitivity to change was assessed using the correlation between QPR intrapersonal, QPR interpersonal and QPR total and WEMWBS change scores from baseline to 3-month follow-up.

Using data-set 2, internal consistency was assessed using Cronbach's alpha. Sensitivity to change was assessed using the association between QPR intrapersonal, QPR interpersonal and

QPR total and WEMWBS change scores between baseline and 12-month follow-up. This was achieved by regressing each QPR scale change score onto the WEMWBS while accounting for clustering at the team level by using the 'xtmixed' command with maximum likelihood estimation in Stata version 11. Site and study arm were entered as covariates in the model to reflect the study design. Prior to conducting the regression analysis, change scores were standardised (mean 0, s.d.=1) to obtain a standardised regression coefficient that is equivalent to a regression coefficient. All analyses were conducted on complete cases using Stata version 11.

Two separate confirmatory factor analyses (CAF) were conducted to assess the fit of baseline data (a) to QPR total (using the one-factor solution previously identified<sup>8</sup>), and (b) to QPR interpersonal and QPR intrapersonal (using the 2-factor solution previously identified<sup>7</sup>). The CAF were conducted in Mplus 7.2 using the weighted least-squares mean variance (WLSMV) estimator,<sup>13</sup> taking into account clustering at the team level and adjusting the model for National Health Service (NHS) trust as this captures the study design. A goodness of fit was assessed using several fit indices:  $\chi^2$  ( $P>0.05$ ), root mean square error of approximation (RMSEA  $<0.06$ ), Tucker–Lewis Index (TLI  $>0.95$ ) and comparative fit index (CFI  $>0.95$ ).

## Results

### Participants

Demographics and QPR scores for both samples are shown in Table 1. The two samples did not differ in sociodemographic characteristics, other than data-set 1 had a higher number of participants from a Black ethnic background ( $\chi^2_{(2)}=10.7$ ,  $P=0.005$ ).

### Convergent validity

In data-set 1 ( $n=76$ ), the baseline RAS score was positively correlated with baseline QPR interpersonal ( $r=0.46$ , 95% CI

**Table 1** Baseline characteristics

	Data-set 1 ( $n=88$ )	Data-set 2 ( $n=399$ )
Gender, <sup>a</sup> $n$ (%)		
Men	62 (70.4)	256 (64.2)
Women	26 (29.6)	142 (35.6)
Age, years: mean (s.d.)	42.3 (10.5)	43.8 (10.9)
Diagnosis, $n$ (%)		
Schizophrenia	43 (48.9)	234 (58.6)
Bipolar disorder	21 (23.9)	50 (12.5)
Schizoaffective disorder	0	46 (11.5)
Depression	2 (2.3)	0
Personality disorder	2 (2.3)	0
Not known	3 (3.4)	0
Other	11 (12.5)	69 (17.3)
More than one	6 (6.8)	0
Ethnicity, <sup>b</sup> $n$ (%)		
White	37 (42.0)	228 (57.1)
Black	40 (45.5)	109 (27.3)
Other/mixed	10 (11.3)	59 (14.8)
Unemployed, $n$ (%)	63 (71.6)	291 (73.1)
Single, $n$ (%)	63 (71.6)	306 (76.7)
QPR, mean (s.d.)		
Intrapersonal	46.33 (9.6)	48.83 (10.1)
Interpersonal	13.73 (2.8)	13.27 (2.6)
Total	41.17 (8.6)	38.72 (9.1)
QPR, Questionnaire about the Process of Recovery.		
a. For data-set 1, missing $n=1$ .		
b. For data-set 1, missing $n=1$ ; for data-set 2, missing $n=0.8$ .		

0.26–0.67,  $P < 0.001$ ), QPR intrapersonal ( $r = 0.75$ , 95% CI 0.63–0.83,  $P < 0.001$ ) and QPR total ( $r = 0.73$ , 95% CI 0.61–0.82,  $P < 0.001$ ), indicating adequate convergent validity for each scale.

### Test-retest reliability

The ICCs in data-set 1 ( $n = 88$ ) between baseline and 2 weeks were ‘good’ for QPR interpersonal (ICC = 0.66, 95% CI 0.53–0.77) and QPR intrapersonal (ICC = 0.75, 95% CI 0.64–0.83), and ‘fair to good’ for QPR total (ICC = 0.74, 95% CI 0.63–0.82).

### Sensitivity to change

The sensitivity to change of QPR was tested using WEMWBS as a comparator, as shown in Table 2. Sensitivity to change in both data-sets was moderate for QPR intrapersonal and QPR total, and low for QPR interpersonal.

### Internal consistency

Cronbach’s alpha coefficients for baseline scores in data-set 2 ( $n = 399$ ) indicated excellent internal consistency for QPR total ( $\alpha = 0.89$ ) and the QPR intrapersonal subscale ( $\alpha = 0.90$ ). However, internal consistency for QPR interpersonal was poor ( $\alpha = 0.49$ ).

### Factor structure

As a first step, we fitted a one-factor model in data-set 2 ( $n = 399$ ) for QPR total, finding an adequate fit ( $\chi^2_{(90)} = 233.2$ ,  $P < 0.001$ , RMSEA = 0.063, 90% CI 0.05–0.07, CFI = 0.97, TLI = 0.97). Table 3 shows the standardised loadings for all 15 items showing that all items load onto the factor.

We then fitted a two-factor model for the two QPR subscales (intrapersonal and interpersonal), also shown in Table 2. This model also offered a good fit ( $\chi^2_{(208)} = 407.5$ ,  $P < 0.001$ , RMSEA = 0.049, 90% CI 0.04–0.06, CFI = 0.96, TLI = 0.96),

**Table 2** Questionnaire about the Process of Recovery (QPR) sensitivity to change

	Warwick-Edinburgh Mental Well-being Scale, $r$ (95% CI)	
	Data-set 1 ( $n = 57$ )	Data-set 2 ( $n = 267$ )
QPR intrapersonal	0.50 (0.28 to 0.67)	0.39 (0.27 to 0.49)
QPR interpersonal	0.18 (–0.09 to 0.42)	0.18 (0.07 to 0.31)
QPR total	0.47 (0.24 to 0.64)	0.40 (0.27 to 0.49)

although items 20 and 22 had low factor loadings, indicating they are weakly associated with the latent construct.

## Discussion

The study used two samples to evaluate the psychometric properties of the two subscales of the 22-item version (QPR intrapersonal and QPR interpersonal) and the 15-item version (QPR total). Both QPR intrapersonal and QPR total demonstrated adequate psychometric properties, whereas QPR interpersonal did not demonstrate psychometric adequacy.

The QPR intrapersonal subscale demonstrated excellent psychometric properties in all areas tested, apart from moderate sensitivity to change. All items had a loading above 0.5 in the CFA, indicating they capture the intrapersonal scale. By contrast, the QPR interpersonal subscale had poor internal consistency and sensitivity to change and the factor analysis found that two of the five items – item 16 ‘meeting people who have had similar experiences makes me feel better’ and item 20 ‘I realise that the views of some mental health professionals is not the only way of looking at things’ – had a factor loading below 0.5. This indicates that they are weakly associated with the latent construct and, therefore, do not describe it very well. Anecdotally, we found that item 20 was more difficult for participants to answer, because it

**Table 3** Item loadings in confirmatory factor analysis

QPR item (original 22 version item number)	1-factor model, 15-item total	2-factor model	
		Intrapersonal	Interpersonal
QPR intrapersonal items			
Feel better about myself (1)	0.60**	0.59**	–
Feel able to take chances in life (2)	0.66**	0.65**	–
Able to develop positive relationships (3)	0.71**	0.70**	–
Feel part of society (4)	0.60**	0.57**	–
Able to assert myself (5)	0.61**	0.60**	–
Feel my life has a purpose (6)	0.64**	0.63**	–
Experiences changed me for better (7)	0.67**	0.66**	–
Able to come to terms with past (8)	0.61**	0.61**	–
Strongly motivated to get better (9)	0.73**	0.73**	–
Recognise positive things I have done (10)	0.67**	0.69**	–
Able to understand myself better (11)	0.68**	0.68**	–
Can take charge of my life (12)	0.75**	0.75**	–
Can actively engage with life (19)	0.75**	0.76**	–
Take control of aspects of my life (21)	0.73**	0.74**	–
Find time to do the things I enjoy (22)	0.55**	0.56**	–
Able to access independent support (13)	–	0.53**	–
Make sense of distressing experiences (18)	–	0.62**	–
QPR interpersonal items			
Weigh up pros and cons of treatment (14)	–	–	0.61**
Experiences made me more sensitive (15)	–	–	0.52**
Meeting people with similar experiences (16)	–	–	0.26*
My recovery has challenged others (17)	–	–	0.49**
Views of professionals not only way (20)	–	–	0.38*

QPR, Questionnaire about the Process of Recovery.  
\* $P < 0.05$ , \*\* $P < 0.001$ .

asked about the view of ‘some’ professionals, which participants found confusing. Overall, these results indicate that QPR interpersonal is not well defined, and constructs with five of more items are generally recommended to define a robust construct.<sup>14</sup>

The 15-item QPR total demonstrated excellent internal consistency and test–retest reliability, adequate convergent validity and moderate sensitivity to change. These findings reflect those of Law and colleagues in their paper that recommended the 15-item version of QPR.<sup>8</sup> In their paper the 15-item version had a Cronbach’s alpha of 0.93 and good test–retest reliability ( $r = 0.70$ ). In our paper the CFA found that all items loaded above 0.5. As all of the 15 items are in the QPR intrapersonal subscale, there is a great deal of overlap between QPR total and QPR intrapersonal.

### Strengths and limitations

The study provides a comprehensive and independent psychometric evaluation of the QPR, including being the first study to evaluate sensitivity to change. It is also the first study to compare the two versions of the QPR. A limitation of this study is that one of our samples was relatively small. The use of two data-sets (although similar in demographic characteristics) can also be seen as a limitation, as can the non-collection of important clinical descriptive data such as illness severity and duration. However, each data-set had advantages, with data-set 1 allowing evaluation of test–retest reliability at 2 weeks, and the sample in data-set 2 sufficiently large to allow CFA.

### QPR in clinical practice

We identify three clinical implications. First, the QPR can be used to measure the effectiveness of services in supporting recovery. Evidence on how best to support recovery is still developing,<sup>4,9</sup> and there is as yet little evidence about the impact of recovery support on recovery outcomes. Having a robust tool to measure recovery will contribute to this goal.

Second, the results suggest that the QPR may have a role in benchmarking services and comparing the effectiveness of interventions. Sensitivity to change for QPR total was moderate. Although this provides stronger evidence than for other recovery measures,<sup>5</sup> a robust understanding of processes that have an impact on sensitivity to change is needed. The absence of longitudinal studies of personal recovery mean that the level of likely change in recovery is unknown. At one extreme, recovery may be a highly stable construct, consistent with set point theory that finds that subjective ratings of well-being quickly return to baseline levels, even after life events cause a temporary change.<sup>15</sup> At the other extreme, recovery may be highly unstable and influenced by a range of intrapersonal, interpersonal and social determinants. Given the measurement challenges inherent to subjective rating scales, the minimally important difference estimate of responsiveness should be established before QPR is used for service evaluation. The QPR has been recommended by the Implementing Recovery through Organisational Change (ImROC) initiative as part of a suite of measures to assess recovery support.<sup>16</sup> Robust measures of recovery support such as INSPIRE<sup>17</sup> can be used in conjunction with the QPR to assess the relationship between recovery support and the experience of recovery.

Third, government policy in England has become more outcome-focused<sup>2</sup> and the introduction of the payment by results funding system in mental health services may lead to outcome measurement being more widely introduced. As recovery is a policy aim, services may need to routinely measure recovery.

Furthermore, there is a growing interest in the use of Patient Reported Outcome Measures (PROMs) such as the QPR in mental health services.<sup>18</sup>

### QPR in research

The relationship between recovery and other outcomes in mental health is still being explored,<sup>19</sup> and standardised measures of recovery are needed. One issue for the measurement of recovery is ensuring that recovery measures have a robust conceptual underpinning.<sup>20</sup> An empirical understanding of key recovery processes is emerging, such as the CHIME framework, which identifies recovery processes of Connectedness, Hope, Identity, Meaning and purpose, and Empowerment.<sup>3</sup> A systematic review of recovery measures that used the CHIME framework to identify the breadth of coverage of recovery measures, indicated that the QPR had the best spread of items covering the five processes, suggesting that the QPR is measuring these aspects of recovery.<sup>5</sup>

We identify two key knowledge outcomes from this study. First, the 15-item QPR total scale can be recommended for use in research and clinical practice. Our recommendation reflects its adequate psychometrics, lower burden compared with the 22-item version, and easier interpretation (as a total score, rather than two subscales). A shorter version reduces the burden on respondents, which is an advantage<sup>21</sup> as people who use services do not like having to complete long outcome measures.<sup>22</sup> Specifically, a shorter measure makes it more feasible for use, thus increasing the likelihood of completion.<sup>23</sup> Future work could evaluate the implications of measuring recovery as a unidimensional construct, using analytic methods such as item response theory or Rasch measurement theory.

Second, we found preliminary evidence of sensitivity to change. This is a key psychometric property currently under-researched in measures of recovery change.<sup>5</sup> Our results provide some evidence of the ability of the QPR to measure change, suggesting that QPR can be used in longitudinal research and to assess change in a clinical setting. Future research will need to identify a gold standard for evaluating this property, but it is encouraging that there was a correlation with change in two measures assessed at two different time points.

Measuring the recovery of people using mental health services with a standardised measure is one method that can be used to assess the effectiveness of services in supporting recovery. Robust measures of recovery are needed to do this. Overall, the 15-item QPR version is a valuable contribution to the measurement of recovery.

**Julie Williams**, MSc, **Mary Leamy**, PhD, **Francesca Pesola**, PhD, **Victoria Bird**, PhD, **Clair Le Boutillier**, MSc, **Mike Slade**, PhD, King’s College London, Health Service and Population Research Department, Institute of Psychiatry, Psychology & Neuroscience, London, UK

**Correspondence:** Julie Williams, Health Service and Population Research Department (Box P029), Institute of Psychiatry, Psychology & Neuroscience, King’s College London, De Crespigny Park, Denmark Hill, London SE5 8AF. Email: julie.williams@kcl.ac.uk

First received 9 Dec 2014, final revision 6 Feb 2015, accepted 13 Feb 2015

### Funding

This article presents independent research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research (PGfAR) Programme (Grant Reference Number RP-PG-0707-10040), and in relation to the NIHR Biomedical Research Centre for Mental Health at South London and Maudsley NHS Foundation Trust and Institute of Psychiatry, Psychology & Neuroscience, King’s College London. The project will be published in full in the NIHR PGfAR journal. The views expressed in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, Medical Research Council, NIHR Central Commissioning Facility, NIHR Evaluation, Trials and Studies

Co-ordinating Centre, the PGfAR programme or the Department of Health. Further information available at [researchintorecovery.com/refocus](http://researchintorecovery.com/refocus).

## Acknowledgements

We acknowledge the support of the PICuP Clinic at the Maudsley Hospital and the other mental health teams in recruiting participants.

## References

- 1 Anthony W. Recovery from mental illness: the guiding vision of the mental health service system in the 1990s. *Psychosoc Rehab J* 1993; **16**: 11–23.
- 2 Department of Health. *No Health Without Mental Health: A Cross-Government Mental Health Outcomes Strategy for People of all Ages*. Department of Health, 2011.
- 3 Leamy M, Slade M, Le Boutillier C, Williams J, Bird V. A conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. *Br J Psychiatry* 2011; **199**: 445–52.
- 4 Le Boutillier C, Leamy M, Bird V, Davidson L, Williams J, Slade M. What does recovery mean in practice? A qualitative analysis of international recovery-oriented practice guidance. *Psychiatr Serv* 2011; **62**: 1470–6.
- 5 Shanks V, Williams J, Leamy M, Bird V, Le Boutillier C, Slade M. Measures of personal recovery: a systematic review. *Psychiatr Serv* 2013; **64**: 974–80.
- 6 Sklar M, Groessl E, O'Connell M, Davidson L, Aarons GA. Instruments for measuring mental health recovery: a systematic review. *Clin Psychol Rev* 2013; **33**: 1082–95.
- 7 Neil S, Kilbride M, Pitt L, Nothard S, Welford M, Sellwood W, et al. The questionnaire about the process of recovery (QPR): a measurement tool developed in collaboration with service users. *Psychosis* 2009; **1**: 1–11.
- 8 Law H, Neil S, Dunn G, Morrison A. Psychometric properties of the Questionnaire about the Process of Recovery (QPR). *Schizophr Res* 2014; **156**: 184–9.
- 9 Slade M, Bird V, Le Boutillier C, Williams J, McCrone P, Leamy M. REFOCUS trial: protocol for a cluster randomised controlled trial of a pro-recovery intervention within community based mental health teams. *BMC Psychiatry* 2011; **11**: 185.
- 10 Corrigan P, Gifford D, Rashid F, Leary M, Okeke I. Recovery as a psychological construct. *Community Ment Health J* 1999; **35**: 231–9.
- 11 Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes* 2007; **5**: 63.
- 12 Maheswaran H, Weich S, Powell J, Stewart-Brown S. Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): group and individual level analysis. *Health Qual Life Outcomes* 2012; **10**: 156.
- 13 Muthén L, Muthén B. *Mplus User's Guide*. Muthén and Muthén, 1998–2011.
- 14 Costello A, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assess Res Eval* 2005; **10**: 1–9.
- 15 Lykken D, Tellegen A. Happiness is a stochastic phenomenon. *Psychol Sci* 1996; **7**: 186–9.
- 16 Shepherd G, Boardman J, Rinaldi M, Roberts G. *Supporting Recovery in Mental Health Services: Quality and Outcomes*. Centre for Mental Health, NHS Confederation, 2014.
- 17 Williams J, Leamy M, Bird V, Le Boutillier C, Norton S, Pesola F, et al. Development and evaluation of the INSPIRE measure of staff support for personal recovery. *Soc Psychiatry Psychiatr Epidemiol* 2015; **50**: 777–86.
- 18 McCabe R, Saidi M, Priebe S. Patient-reported outcomes in schizophrenia. *Br J Psychiatry* 2007; **191** (suppl 50): s21–8.
- 19 Andresen R, Caputi P, Oades L. Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Res* 2010; **177**: 309–17.
- 20 Burgess P, Pirkis J, Coombs T, Rosen A. Assessing the value of existing recovery measures for routine use in Australian mental health services. *Aust NZ J Psychiatry* 2011; **45**: 1–14.
- 21 Reininghaus U, Priebe S. Measuring patient-reported outcomes in psychosis: conceptual and methodological review. *Br J Psychiatry* 2012; **201**: 262–7.
- 22 Crawford M, Robotham D, Thana L, Patterson S, Weaver T, Barber R, et al. Selecting outcome measures in mental health: the views of service users. *J Ment Health* 2011; **20**: 336–46.
- 23 Slade M, Thornicroft G, Glover G. The feasibility of routine outcome measures in mental health. *Soc Psychiatry Psychiatr Epidemiol* 1999; **34**: 243–9.