**­­Psychometric evaluation of the Hope, Agency and Opportunity (HAO);**

**A brief measure of mental health recovery**

**Abstract**

**Background:** The development of safe and effective mental health services is a priority. This requires valid measures of personal recovery, yet these tools are not embedded in routine clinical practice. Brief ‘patient reported measures’ are most likely to be acceptable to service-users and clinicians. The 4-item ‘Hope, Agency and Opportunity’ (HAO) was co-produced to assess recovery outcomes and experience of mental health services.

**Aim:** To evaluate the psychometric properties of the HAO.

**Method:** A clinical sample from secondary healthcare services and a non-clinical sample were assessed at baseline and two weeks, on measures of personal recovery.

**Results:** Factor analysis indicated goodness of fit for the HAO with both clinical and non-clinical samples. The measure demonstrated acceptable internal consistency, moderate to strong construct validity and substantial test-retest reliability over two weeks.

**Conclusions:** The HAO demonstrates satisfactory psychometric properties. Co-production of the measure confers clinical credibility. The brevity of the tool means it can be incorporated into routine clinical practice to drive improvements in service quality.

**Declaration of interest:** None

**Keywords:** Patient reported measures, outcome evaluation, PROMs, PREMs, psychometric evaluation, recovery, questionnaire validation

**Introduction**

***The case for brief measures of recovery***

The quality of mental health services is an international priority (Australian Health Ministers’ Advisory Council, 2013; New Zealand Mental Health Commission, 2012; Piat, Sabetti & Bloom, 2010; President's New Freedom Commission on Mental Health, 2003; UK Department of Health, 2011; 2014). A recovery approach focuses on *‘building a meaningful and satisfying life, whether or not there are ongoing or recurring symptoms or problems’* (Shepherd et al., 2008), and has been adopted as the organising framework for policymakers in Britain, Australasia and North America (*ibid*). This involves a fundamental change in the overarching aims of mental health services, from traditional notions of cure to supporting people to live well with longer term conditions (Slade et al., 2014; cf. Rose, 2014).

The UK Royal College of Psychiatrists (2013) identifies the use of reliable outcome data as key to this change, and to ensure safe and effective healthcare. Routine use of validated ‘patient reported measures’ meets the need to gather information directly from people using services to assess and improve provision (Dawson et al., 2010; cf. Lewis & Killaspy, 2014). With the growing emphasis on personal recovery as well as clinical outcomes (Hasson-Ohayon, Roe, Yanos & Lysaker, 2015; Shepherd et al., 2014), and specifically the experience of hope, agency and opportunity for purposeful activity and social inclusion irrespective of mental health status, reliable self-report measures of personal recovery are essential (Shepherd et al., 2014).

While policy and best practice guidelines prioritise recovery outcomes (Burgess et al., 2011; Thornicroft & Slade, 2014), efforts to embed measures in routine practice have repeatedly failed, largely due to validity and feasibility concerns (Gilbody et al, 2002; Kightley et al., 2010; Pirkis & Callaly, 2010). Reviews reach differing conclusions about whether clinicians believe these measures to be useful, but agree that NHS services have failed to establish an outcome oriented culture despite over 20 years’ policy and linked initiatives (Duncan & Murray, 2012; Jacobs, 2009). Barriers include poor communication of the rationale for measures, limited IT support and lack of feedback (Jacobs, 2009). This is a serious problem. We are not following best practice guidelines because measures are considered irrelevant to therapeutic work or too burdensome. Importantly, the limited evidence from service-users and carers themselves indicates that those receiving services find the tools useful (Black et al., 2009; Law et al., 2012). Duncan and Murray (2012) argue that the current situation is unlikely to change without clinician, team and organisational level commitment.

Using a hierarchical, criterion-based review of available tools, Burgess and colleagues identified four measures of individual recovery suitable for routine use (ranging from 15 to 50 items). The authors concluded that these tools need to be assessed for feasibility in public sector settings, and that modifications and new tools may be required (Burgess et al., 2011). Similarly, Law et al. (2012) reviewed measures of recovery for clinical utility and acceptability, and concluded that the 41-item Recovery Assessment Scale (Giffort et al., 1995) met validity requirements and was most acceptable to service-users. Importantly, no brief measures of less than 15 items were identified in either of the reviews. While 15 items may not seem long to researchers, the evidence to date suggests that measures of this length are unlikely to be implemented in routine service settings.

Current brief measures of personal recovery include the Questionnaire about the Process of Recovery (QPR; Neil et al., 2009) and versions of the INSPIRE (Williams, Leamy, Bird et al., 2015). The 22-item QPR and shorter 15-item version (Law, Neil, Dunn & Morrison, 2014) are ‘patient reported outcome measures’ (PROMs). The 27-item INSPIRE and 5-item Brief-INSPIRE are ‘patient reported experience measures’ (PREMs). Each demonstrates satisfactory psychometric properties. To incorporate assessment of personal recovery into routine clinical practice, a brief valid measure of both outcome and experience of services is required. The 4-item ‘Hope, Agency and Opportunity’ (HAO) scale is a very brief measure of recovery, combining patient reported outcomes and experience.

***Development of the HAO***

An initial version of the HAO was drafted jointly by the service-user lead for an NHS Trust in the South of England (with lived experience of mental ill-health) and a clinical psychologist, to assess the impact of the local Recovery College. The first three items correspond to the key themes identified in the literature as predictive of recovery: hope, agency and opportunity for purposeful activity and social inclusion (Repper & Perkins, 2003; Shepherd et al., 2014). The final item assumes that the working alliance is integral to effective service provision (Borg & Kristiansen, 2009; Hicks et al., 2012; Slade, 2009), and was included to assess experience of services as well as outcomes. Other frameworks of recovery (eg CHIME; Leamy, Bird, Le Boutillier, Williams & Slade, 2011) incorporate additional components such as connectedness and identity. We decided to use just the three themes identified by Repper and Perkins (2003) in order to keep the measure as brief as possible.

A consultation group of four people with lived experience and four clinicians then worked together to shape the measure prior to wider roll out. This group brought together people from across the Trust who were involved in the Recovery College or other recovery developments across the organisation. The consultation group retained the four questions, and added prompts as a means of encouraging reflection on each item, and facilitating recovery conversations between service-users and clinicians. The group did not seek additional items but decided to improve the design (adding colour and a more professional layout) to increase likelihood of engagement with the tool. This work was done through an iterative process whereby changes were proposed, updated versions were reviewed by the group and local clinicians, and further refinements made. At each step changes were discussed within the consultation group through face-to-face meetings and email, until a final version was agreed by all.

The current study evaluates the psychometric properties of the HAO, a brief co-developed measure of recovery, designed for use in routine clinical practice.

**Method**

***Participants***

The clinical group comprised 139 people utilising secondary care mental health services, recruited via third sector drop-in settings (see Table 1). Participants’ ages ranged from 19 to 69 years (*M* = 45.6, *SD* = 11.3), and included 78 women (56.1%) and 61 men (43.9%), with a broad range of diagnoses (see Table 1). The majority identified themselves as ‘British’ (95.0%). The non-clinical sample comprised 201 individuals recruited from a university pool. Non-clinical participants were excluded if they reported current mental health difficulties. The group’s ages ranged from 17 to 48 years (*M* = 20.35, *SD* = 3.6), and included 173 women (86.1%) and 28 men (13.9%). The majority identified themselves as ‘British’ (73.1%) or ‘Any other White background’ (8%).

Table 1 about here

***Procedure***

Ethical approval was obtained from the local NHS Trust and County Council for the clinical group, and University for the non-clinical group. Clinical participants were recruited through ‘Wellbeing Centres’ which support people accessing local secondary care services. The non-clinical group was recruited through a university pool. Participants were excluded if they were unable to read English (as measures are validated in English). Clinical participants had ongoing mental health difficulties. Non-clinical participants were included if they confirmed that they were not seeking treatment for mental ill-health at the time of the study. All participants were given written information about the study and consented. No one requested assistance to complete the questionnaires.

The clinical group completed the HAO, QPR, Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS) and INSPIRE-CHIME (an earlier version of the Brief-INSPIRE) at baseline, and repeated the HAO at two-week follow-up. The non-clinical group completed the HAO, WEMWBS and INSPIRE-CHIME at baseline, and repeated the HAO at two-week follow-up. Participants were debriefed at the end of the study.

***Measures***

The HAO is a 4-item self-report measure of recovery developed in consultation with people with mental ill-health (see Figure 1). Consistent with the principle that mental well-being, self-determination and social inclusion are relevant to all of us, the HAO was designed for use in clinical and non-clinical groups. The questionnaire incorporates both outcome (PROM) and experience (PREM) items. Outcome items refer to the three key components of recovery identified in the literature: hope, agency and opportunity for purposeful activity and social inclusion(Repper & Perkins, 2003; Shepherd et al., 2014). The final item refers to experience of services and is therefore only completed by people in contact with services. This final item is based on the assumption that when services are effective, the working alliance is integral to recovery (Borg & Kristiansen, 2009; Hicks et al., 2012; Slade, 2009). The consultation group of people with lived experience of mental ill-health added three prompts to each of the four main questions. The group felt strongly that people should be encouraged to consider different aspects of each of the items before deciding on a rating. Comment boxes were also included to add clinical utility, for notes to guide recovery based conversations and care planning, for example. We recognise the tension between including prompts and usual best practice in questionnaire design (which would include just one question per item to avoid ambiguity). We privileged the advice of our consultation group and acknowledge the methodological limitations resulting from this. Items are rated on a 5-point scale from 0 (none of the time) to 4 (all of the time), and summed to give a total score (0-16). Higher scores indicate greater recovery.

Figure 1 about here

The 15-item QPR (Law et al., 2015) is an outcome measure (PROM) of recovery developed from a qualitative study led by researchers with experience of mental ill-health. Items are rated on a five-point scale (0 = disagree strongly; 4 = agree strongly), and summed to give a total score (0-60). Higher scores indicate greater recovery. In the original sample, test-retest reliability was good (*r* = 0.70) and internal consistency was excellent (*α* = 0.93).

The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS; Tennant et al., 2007) is a 14-item measure of mental wellbeing in the general population. Items are rated on a 5-point scale (1 = none of the time; 5 = all of the time), and summed to give a total score (14-70). Higher scores indicate greater wellbeing. In the original sample, test–retest reliability was excellent (0.83) and internal consistency was good (*α* = 0.89, student sample) to excellent (*α* = 0.91, population sample).

The 5-item INSPIRE-CHIME was used in the present study. This has subsequently been adapted to form the Brief-INSPIRE (Williams, Leamy, Bird et al., 2015). Versions of the INSPIRE measure people’s experience of support in their recovery (PREMs). The INSPIRE-CHIME differs from the Brief-INSPIRE in one way; each item is prefaced with the question ‘An important part of my recovery is…’ to which participants state yes or no. In both the INSPIRE-CHIME and Brief-INSPIRE participants are then asked to rate how much they feel supported by their worker (or ‘others’ for the non-clinical group) in each domain. Items are rated on a 5-point scale (0 = not at all; 4 = very much). The summed total is multiplied by five to give a final score (0-100). Higher scores indicate greater levels of perceived support. Test–retest reliability for the Brief-INSPIRE was 0.72, and internal consistency was good (*α* = 0.86), in the original sample. Given the slight difference between the INSPIRE-CHIME and the Brief-INSPIRE, we cannot assume these levels of reliability, and so assessed internal consistency for the present clinical and non-clinical samples.

***Data Analysis***

Confirmatory factor analysis (CFA) was used to assess how well a single factor model for HAO fitted the full dataset (clinical and non-clinical samples combined) and for the clinical and non-clinical data separately. We computed CFA using AMOS version 24. All other analyses were run using IBM SPSS version 22/24. Cronbach’s alphas and inter-item correlations were calculated to assess internal consistency. Intra-class correlations and weighted Cohen’s Kappa indicated test-retest reliability. Only participants who completed the study 14 (+/- 7) days after Time 1 were included in the test-retest analyses. Construct validity was assessed using Pearson’s correlations between HAO at Time 1 and each of QPR, WEMWBS and INSPIRE. For the non-clinical group, Pearson’s correlations were calculated between HAO and each of WEMWBS and INSPIRE. Exploratory analysis of the differences between the clinical and non-clinical groups used a repeated measure ANOVA with one between-subject variable (*Group*) and one within-subject variable (*Time*), controlling for age and gender.

**Results**

***Confirmatory Factor Analysis.*** A single factor model for the clinical dataset was conducted on the four items of the HAO for the clinical group. The model indicated a good fit, χ2 (2, *n*=139)=0.90, *p*=0.64; CFI=1.00, RMSEA=0.000 (90% CI: 0.00-0.14), for individual paths in the model (see Figure 2).

Figure 2 about here

The three item HAO was also assessed as a single factor model for the clinical and non-clinical groups. When there are no additional constraints specified in a one-factor model with three indicators the model is just-identified and shows individual parameters of the model but does not produce goodness of fit indices. In order to identify the model with the three items, we looked at the individual paths, and constrained the measurement residuals.

A multi-group analysis of the model was conducted where only the measurement residuals were restricted to be equal between groups. Goodness of fit indices for this baseline model indicated a good model fit: χ2 (3, *n*=340)=1.99, *p=*0.57; CFI=1.00, RMSEA=0.00 (90% CI: 0.00-0.08). Additionally constraining the variance of the latent variable and the regression weights on the latent factor ‘recovery’ between clinical and non-clinical samples showed a similar model fit: χ2 (7, *n*=340)=7.56, *p=*0.27; CFI=0.99, RMSEA=0.03 (90% CI: 0.00-0.08). The specific paths in the model for the clinical and non-clinical sample are shown in figure 3.

Figure 3 about here

The model testing scalar invariance (where the intercepts are also constrained) revealed a weak model fit, χ2 (9, *n*=340)=83.240, *p<*.001; CFI=0.64, RMSEA=0.16 (90% CI: 0.13-0.19). We can conclude that the HAO has partial measurement invariance and the only difference lies in the intercepts of the groups. In general, the non-clinical sample scored higher on each item of the HAO then the clinical sample (see figure 3).

***Clinical Sample***

***Internal consistency.*** Cronbach’s alpha at Time 1 (*n*=138) showed acceptable internal consistency (*α*=0.69)[[1]](#footnote-1). Alphas may be low without indicating poor internal consistency for brief measures, whereas mean inter-item correlations estimate internal consistency independent of scale length (Spiliotopoulou, 2009). Cronbach’s formula for inter-item correlations (Cronbach, 1951) was therefore used, and yielded a value of 0.36 for the HAO, which is within the acceptable range of 0.15 to 0.50 (Clark & Watson, 1995). Inter-item correlations are reported in Table 2.

Table 2 about here

***Test-retest reliability.*** Two-way mixed effects intra-class correlation between Time 1 and Time 2 (*n*=88) indicated substantial reliability (ICC=0.63, 95% CI: 0.48–0.74). Weighted Cohen's Kappa coefficients ranged from 0.38 to 0.59 for individual items (fair-moderate agreement), and was 0.66 for both the three and four-item HAO (good agreement). See Table 3.

***Construct validity.*** Internal consistency was excellent for the QPR (α=0.94), WEMWBS (α=0.94) and INSPIRE (α=0.92) in the clinical sample. HAO correlated strongly with QPR (*r*=0.79, 95% CI: 0.71–0.85, *n*=121, *p* <0.001) and WEMWBS (*r*=0.77, 95% CI: 0.69–0.83, *n*=126, *p*<0.001), and moderately with INSPIRE (*r*=0.56, 95% CI: 0.41–0.69, *n*=100, *p*<0.001).

***Non-Clinical Sample***

***Internal consistency.*** Cronbach’s alpha at Time 1 (*n*=201) indicated acceptable internal consistency (*α*=0.74). The mean inter-item correlation was also within the acceptable range (*p* =0.49). Inter-item correlations are reported in Table 2.

***Test-retest reliability.*** Two-way mixed effects intra-class correlation between Time 1 and Time 2 (*n*=75) indicated substantial reliability (ICC=0.69, 95% CI: 0.55–0.79). Weighted Cohen's Kappa coefficients ranged from 0.42 to 0.58 for individual items (fair-moderate agreement), and was 0.69 for the three-item HAO (good agreement). See table 3.

Table 3 about here

***Construct validity.*** Internal consistency was excellent for the WEMWBS (α=0.91) and good for the INSPIRE (α=0.81) in the non-clinical sample. HAO correlated strongly with WEMWBS (*r*=0.71, 95% CI: 0.63–0.77, *n*=200, *p*<0.001) and moderately with INSPIRE (*r*=0.62, 95% CI: 0.53–0.70, *n*=192, *p*<0 .001).

***Post hoc analysis of group differences.*** A repeated-measures ANOVA with one between-subject variable, *Group* (Clinical vs Non-clinical), and one within-subject variable, *Time* (Time 1 vs Time 2), was conducted for the HAO (excluding the final ‘working relationships’ item which is only completed by clinical participants). There was a main effect of time, *F*(1, 164)=4.24, *p*<0 .05, np2=0.03, group, *F*(1, 164)=32.4, *p*<0.001, np2=0.17, and time by group interaction, *F*(1, 164)=9.03, *p*=0.005, np2=0.05.

Given the differences in age and gender between the two samples, we repeated the analyses controlling for these variables. While the main effect of time was no longer significant, *F*(1, 151)=2.66, *p*<0.105, np2=0.17, the main effect of group, *F*(1, 151)=6.27, *p*<0.05, np2=0.040, and time by group interaction, *F*(1, 151)=5.338, *p*=0.05, np2=0.034) remained significant. None of the interaction terms with age or gender were significant (all *p*s>0.05).

To understand the time effect in the repeated measures analyses, paired samples *t*-tests with a Bonferroni corrected significance level (0.025) indicated that recovery improved in the clinical group from Time 1 (*M*=5.85, *SD*=2.13) to Time 2 (*M*=6.49, *SD*=2.21), *t*(90)=3.56, *p*=0.001, and that no change was observed in the non-clinical group between Time 1 (*M*=7.91, *SD*=2.05) and Time 2 (*M*=7.79, *SD*=1.75), *t*(74)=0.69 (see Figure 4). These simple effects indicate that the main effect of Time was driven solely by the difference in the clinical group between Time 1 and Time 2.

Figure 4 about here

**Discussion**

***Overview of findings***

This study recruited a clinical and non-clinical sample to assess the psychometric properties of the HAO, a very brief patient reported measure of mental health recovery. The HAO demonstrates goodness of fit, acceptable internal consistency, moderate to strong construct validity and substantial test-retest reliability.

The single factor found with both clinical and non-clinical samples suggests that the HAO assesses one underlying construct that we can term ‘recovery.’ Internal consistency is typically estimated using Cronbach’s alpha. However, alpha values can be low without indicating poor internal consistency as a result of short scale length. Cronbach’s alpha, and the conventionally accepted value of 0.70 for acceptability, may therefore fail to provide an adequate assessment of internal consistency for brief measures and where small samples are used (Spiliotopoulou, 2009). Inter-item correlations evaluate internal consistency independent of scale length (Cronbach, 1951; Spiliotopoulou, 2009), and are in the acceptable range for HAO.

The HAO shows strong convergence with the comparator outcome (QPR) and wellbeing (WEMWBS) measures, and moderate convergence with the experience measure (Brief-INSPIRE). It is likely that this reflects the make-up of the HAO, which includes three outcome-based items and a single experience-based item. Together with test-retest reliability, these results show that the HAO can be used to measure recovery outcomes and experience of services, and over time.

***Limitations***

By privileging the advice of our consultation group to include prompts for each of the four main items, we have not kept to usual standards of questionnaire design which elicit a response to one unambiguously defined question per item (Furr, 2011; Shaugnessy, Zechmeister & Zechmeister, 2012, chapter 5). We acknowledge that the inclusion of these prompts could lead to ambiguity and reduced confidence in the resulting data. However, we would argue that the co-production of the measure, along with the specific changes made, enhance clinical credibility and utility of the HAO, and that the psychometric properties suggest that people did not find the prompts confusing or ambiguous.

 Assessment of validity is limited by the use of an earlier version of the INSPIRE to the one now available, though internal consistency for the INSIRE-CHIME was excellent in the clinical sample and good in the non-clinical sample. Also, we do not yet know if the HAO can reliably measure sensitivity to change.

The clinical and non-clinical samples were not demographically comparable. The clinical group was, on average, aged in the mid-40s, with a broadly equal gender split. The non-clinical group was younger, predominantly female and recruited from a university pool. There are also likely to have been educational differences between the two samples. Importantly, our aim was not to compare the groups but to determine reliability of the HAO in these two different populations. If the concept of recovery is relevant to all of us, whether using services or not, we need tools that can be used by non-clinical as well as clinical groups. This might include people with physical healthcare needs, vulnerable groups such as people who are homeless, and general population groups who may be struggling with sub-clinical problems (accepting the need to validate with specific groups separately).

***Implications***

The routine use of patient reported measures has been identified as key to improving mental healthcare provision, and addressing organisational failures (Burgess et al., 2011; 2012; Thornicroft & Slade, 2014); we need measures that are of *clinical value* and *feasible in practice* in order to improve the quality of care provided (Royal College of Psychiatrists, 2011; cf. Smith et al, 2016).

The majority of clinicians do not use outcome measures in day-to-day practice, mainly due to validity and resource concerns (Gilbody et al., 2002; Kightley et al., 2010; Pirkis & Callaly, 2010). The HAO directly addresses the question of *clinical value* to patients by targeting the elements identified as essential to living well with mental ill-health: hope, agency and opportunity for valued activity and social inclusion (Repper & Perkins, 2003; Shepherd, 2014). The brevity of the HAO addresses the question of *feasibility.*

In addition to point in time assessment of individuals’ recovery, regular use of ‘patient reported’ outcome and experience measures is likely to be an effective means of shaping service delivery. Used prior to clinic appointments, brief PROMs and PREMs provide key information about current presentation and prompt collaborative discussion about the factors known to be associated with living well with mental ill-health. In this way, these tools can structure conversations, monitor change and benchmark service level performance when pooled (cf. Williams, Leamy, Bird et al., 2015). If used at care plan reviews, the measures can be included at established review points, and compared with previous scores, minimising administrative burden. The HAO is designed to be completed independently, and the validation process suggests that this was not problematic for those who took part. While training in the use of measures may not be necessary, sessions targeting individual, team and organisational level commitment to the use of routine outcome measures is likely to be needed to effect wider cultural change (Duncan & Murrey, 2012).

In summary, the HAO demonstrates satisfactory psychometric properties, co-production of the measure confers clinical credibility, and the brevity of the tool enables use in routine clinical practice. This addresses the validity and feasibility concerns currently hindering the use of recovery measures in services. The modest range of validated tools now available, including the HAO, QPR, WEMWBS and INSPIRE, allows us to select patient reported measures of recovery to fit service requirements, and thereby meet the need to incorporate these measures in routine practice, and drive service improvement.

**References**

Australian Health Ministers’ Advisory Council’s National Mental Health Working Group (2009) *4th National Mental Health Plan: An agenda for collaborative government action in mental health 2009–2014.*

Bellack, A.S. (2006) Scientific and consumer models of recovery in schizophrenia: Concordance, contrasts, and implications. *Schizophrenia Bulletin, 32,* 432–442.

Black, J., Lewis, T., McIntosh, P., Callaly, T., Coombs, T., Hunter, A. & Moore, L. (2009) It's not that bad: The views of consumers and carers about routine outcome measurement in mental health. *Australian Health Review, 33,* 93-99.

Borg, M. & Kristiansen, K. (2004) Recovery-oriented professionals: Helping relationships in mental health services. *Journal of Mental Health, 13*(5), 493-505.

Burgess, P., Coombs, T., Clarke, A., Dickson, R., & Pirkis, J. (2012) Achievements in mental health outcome measurement in Australia: Reflections on progress made by the Australian Mental Health Outcomes and Classification Network (AMHOCN). *International Journal of Mental Health Systems, 6*(1), 1.

Burgess, P., Pirkis, J., Coombs, T. & Rosen, A. (2011) Assessing the value of existing recovery measures for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry, 45*(4), 267-280.

Clark, L.A. & Watson, D. (1995) Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*(3), 309-319. http://dx.doi.org/10.1037/1040-3590.7.3.309

Cronbach, L. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334. http://dx.doi.org/10.1007/bf02310555

Dawson, J., Doll, H., Fitzpatrick, R., Jenkinson, C. & Carr, A.J. (2010) The routine use of patient reported outcome measures in healthcare settings. *British Medical Journal*, *340*, c186. http://dx.doi.org/10.1136/bmj.c186

Department of Health (2014) *Closing the gap: priorities for essential change in mental health*. Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment\_data/file/281250/Closing\_the\_gap\_V2\_-\_17\_Feb\_2014.pdf

Duncan, E. & Murray, J. (2012) The barriers and facilitators to routine outcome measurement by allied health professionals in practice: A systematic review. *BMC Health Services Research, 12,* 1-9.

Furr, M. (2011). *Scale construction and psychometrics for social and personality psychology.* London: SAGE.

Giffort, D., Schmook, A., Woody, C., Vollendorf, C. & Gervain, M. (1995) *Construction of a scale to measure consumer recovery.* Springfield, IL: Illinois Office of Mental Health.

Gilbody, S.M., House, A.O. & Sheldon, T.A. (2002) Psychiatrists in the UK do not use outcomes measures: National survey. *The British Journal Of Psychiatry*, *180*(2), 101-103. http://dx.doi.org/10.1192/bjp.180.2.101

Hasson-Ohayon, I., Roe, D., Yanos, P. T. & Lysaker, P. H. (2015) The trees and the forest: Mixed methods in the assessment of recovery based interventions' processes and outcomes in mental health. *Journal of Mental Health, 25*(6), 543-549.

Hicks, A.L., Deane, F.P. & Crowe, T.P. (2012) Change in working alliance and recovery in severe mental illness: An exploratory study. *Journal of Mental Health, 21*(2), 127-134.

Jacobs, R. (2009) *Investigating patient outcome measures in mental health.* York: Centre for Health Economics.

Kirby, M.J.L. & Keon, W.J. (2006) *Out of the shadows at last: Transforming mental health, mental illness and addition services in Canada.* Retrieved from http://www.parl.gc.ca/Content/SEN/Committee/391/soci/rep/pdf/rep02may06part1-e.pdf

Kightley, M., Einfeld, S., & Hancock, N. (2010) Routine outcome measurement in mental health: Feasibility for examining effectiveness of an NGO. *Australasian Psychiatry, 18,* 167–169.

Law, H., Morrison, A., Byrne, R. & Hodson, E. (2012) Recovery from psychosis: A user informed review of self-report instruments for measuring recovery. *Journal of Mental Health, 21*(2), 192-207.

Law, H., Neil, S., Dunn, G. & Morrison, A. (2014) Psychometric properties of the Questionnaire about the Process of Recovery (QPR). *Schizophrenia Research, 156*(2), 184-189.

Leamy, M., Bird, V., Le Boutillier, C., Williams, J. & Slade, M. (2011) Conceptual framework for personal recovery in mental health: Systematic review and narrative synthesis. *British Journal of Psychiatry, 199*(6), 445-452.

Lewis, G. & Killaspy, H. (2014) Getting the measure of outcomes in clinical practice. *Advances in Psychiatric Treatment, 20*(3)*,* 165-171.

Pesola, F., Williams, J., Bird, V., Freidl, M., Le Boutillier, C., Leamy, M., . . . Slade, M. (2015) Development and evaluation of an Individualized Outcome Measure (IOM) for randomized controlled trials in mental health. *International Journal of Methods in Psychiatric Research*, *24*(4), 257 -265. http://dx.doi.org/10.1002/mpr.1480

Pirkis, J. & Callaly, T. (2010) Mental health outcome measurement in Australia. In: Trauer, T. (ed). *Outcome Measurement in Mental Health: Theory and Practice.* Cambridge: Cambridge University Press, pp.15–25.

President's New Freedom Commission on Mental Health (2003) *Achieving the promise: Transforming mental health care in America*. Retrieved from http://govinfo.library.unt.edu/mentalhealthcommission/reports/FinalReport/downloads/FinalReport.pdf

Repper, J. & Perkins, R. (2003) *Social inclusion and recovery*. Edinburgh: Baillière Tindall.

Rose, D. (2014) The mainstreaming of recovery. *Journal of Mental Health, 23*(5), 217-218.

Royal College of Psychiatrists (2011) *Outcome measures recommended for use in adult psychiatry* (Occasional Paper OP78). Retrieved from http://www.rcpsych.ac.uk/files/pdfversion/op78x.pdf

Royal College of Psychiatrists (2013) *Driving quality implementation in the context of the Francis report* (Occasional Paper OP92). Retrieved from http://www.rcpsych.ac.uk/files/pdfversion/OP92.pdf

Shepherd, G., Boardman, J., Rinaldi, M. & Roberts, G. (2014) *Supporting Recovery in*

 *Mental Health Services: Quality and Outcomes*. Retrieved from http://www.imroc.org/wp-content/uploads/8Supporting-recovery-quality-and-outcomes-briefing-final-for-website-3-March.pdf

Schrank, B. & Slade, M. (2007) Recovery in psychiatry. *The Psychiatrist, 31*, 321–5.

Shaughnessy, J.J., Zechmeister, E.B. & Zechmeister J.S. (2012). *Research methods in psychology* (9th ed.). New York: McGraw Hill.

Slade, M. (2009). The contribution of mental health services to recovery. *Journal of Mental Health, 18*(5), 367-371.

Slade, M., Amering, M., Farkas, M., Hamilton, B., O'Hagan, M., Panther, G. ... & Whitley, R. (2014) Uses and abuses of recovery: Implementing recovery‐oriented practices in mental health systems. *World Psychiatry, 13*(1), 12-20.

Slade, M., Amering, M. & Oades, L. (2008) Recovery: An international perspective. *Epidemiology and Psychiatric Sciences, 17*(2), 128–37.

Smith, W., Patel, A., McCrone, P., Jin, H., Osumili, B. & Barrett, B. (2016) Reducing outcome measures in mental health: A systematic review of the methods. *Journal of Mental Health, 25*(5), 461-472.

Spiliotopoulou, G. (2009) Reliability reconsidered: Cronbach's alpha and paediatric assessment in occupational therapy. *Australian Occupational Therapy Journal*, *56*(3), 150-155. http://dx.doi.org/10.1111/j.1440-1630.2009.00785.x

Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., . . . Stewart-Brown, S. (2007) The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health Quality Life Outcomes, 5,* 63. http://dx.doi.org/ 10.1186/1477-7525-5-63.

Williams, J., Leamy, M., Bird, V., Le Boutillier, C., Norton, S., Pesola, F., & Slade, M. (2015). Development and evaluation of the INSPIRE measure of staff support for personal recovery. *Social Psychiatry And Psychiatric Epidemiology*, *50*(5), 777-786. http://dx.doi.org/10.1007/s00127-014-0983-0

Williams, J, Leamy, M., Pesola, F., Bird, V., Le Boutillier, C. & Slade, M. (2015) Psychometric evaluation of the Questionnaire about the Process of Recovery (QPR). *The British Journal Of Psychiatry*, *207*(6), 551-555. http://dx.doi.org/10.1192/bjp.bp.114.161695

**Figure 1**



Hope, Agency and Opportunity (HAO) measure of recovery

**Figure 2**



CFA model for the four-item HAO in the clinical sample

**Figure 3**





CFA model for the three-item HAO in the non-clinical (top frame) AND clinical sample (bottom frame)

All variances and regression weights have been restricted to be equal between groups, with only intercept weights varying between groups

**Figure 4**

HAO scores for clinical and non-clinical groups at Time 1 and Time 2

Error bars represent +/-*SEM*

**Table 1**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Clinical Group** **(*N* = 139)** | **Non-Clinical Group** **(*N* = 201)** |
| Gender: *n* (%) | male | 61 (43.9) | 28 (13.9) |
|  | female | 78 (56.1) | 173 (86.1) |
| Age, years: *M* (*SD*) |  | 45.6 (11.3) | 20.35 (3.6) |
| Ethnicity: *n* (%) (where > 5%) |  |
| English/ Welsh/ Scottish/ Northern Irish/ British | 132 (95.0) | 147 (73.1) |
| any other White background | - | 16 (8.0) |
| Diagnosis: *n* (%) | anxiety disorder | 85 (61.2) | - |
|  | depression | 92 (66.2) | - |
|  | eating disorder | 7 (5.0) | - |
|  | schizophrenia / psychosis | 32 (23.0) | - |
|  | bipolar disorder  | 24 (17.3) | - |
|  | personality disorder | 26 (18.7) | - |
|  | other | 1 (0.7) | - |
|  | decline to report | 5(3.6) |  |

Participant characteristics

**Table 2**

|  |
| --- |
| **Clinical Group** |
|  | Hope  | Agency | Opportunity | Working relationships |
| Hope  | 1.000 |  |  |  |
| Agency | 0.420 | 1.000 |  |  |
| Opportunity | 0.498 | 0.367 | 1.000 |  |
| Working relationships | 0.299 | 0.298 | 0.312 | 1.000 |
| **Non-Clinical Group** |
|  | Hope  | Agency | Opportunity |   |
| Hope  | 1.000 |  |  |  |
| Agency | 0.558 | 1.000 |  |  |
| Opportunity | 0.469 | 0.449 | 1.000 |   |

Inter-item correlations

**Table 3**

|  |  |
| --- | --- |
|  | **Weighted Cohen's Kappa** |
|  | **clinical group****(*n* = 91)** | **non-clinical group****(*n* = 75)** |
| Hope | 0.546 | 0.555 |
| Agency | 0.382 | 0.579 |
| Opportunity | 0.589 | 0.423 |
| Working Relationships | 0.425 |  |
| HAO (3 item) | 0.656 | 0.686 |
| HAO (4 item) | 0.659 |  |

Weighted Cohen’s Kappa

1. Thanks to a blind reviewer for noting that this represents an acceptable level of consistency given length of scale. [↑](#footnote-ref-1)