

# Recovery Assessment Scale – Domains and Stages (RAS-DS): Its feasibility and outcome measurement capacity

Nicola Hancock<sup>1</sup>, Justin Newton Scanlan<sup>1</sup>, Anne Honey<sup>1</sup>, Anita C Bundy<sup>1</sup> and Katrina O'Shea<sup>2</sup>

Australian & New Zealand Journal of Psychiatry  
2015, Vol. 49(7) 624–633  
DOI: 10.1177/0004867414564084



© The Royal Australian and  
New Zealand College of Psychiatrists 2014  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
anp.sagepub.com



## Abstract

**Objective:** A self-report instrument of mental health recovery is needed both to facilitate collaborative, recovery-oriented practice and measure recovery-focused outcomes. The Recovery Assessment Scale – Domains and Stages (RAS-DS) has been developed to simultaneously fulfill these goals. The aim of this study was to test the feasibility and measurement properties of the RAS-DS.

**Method:** Feasibility was examined by 58 consumer-staff pairs volunteering from 3 non-government organisations. Consumers completed the RAS-DS, discussed it with staff, and then both completed Usefulness Questionnaires. The psychometric properties were examined using Rasch analysis with the data from these consumer participants and from additional participants recruited from two Partners in Recovery programs (N=324).

**Results:** Over 70% of consumers reported taking 15 minutes or less to complete the RAS-DS and rated the instrument as easy or very easy to use. Qualitative data from both consumers and staff indicated that, for most, the RAS-DS was an easy to use, meaningful resource that facilitated shared understandings and collaborative goal setting. However, for a very small number of consumers, the instrument was too confronting and hard to use. Rasch analysis demonstrated evidence for excellent internal reliability and validity. Raw scores were highly correlated with Rasch-generated overall scores and thus no transformation is required, easing use for clinicians. Preliminary evidence for sensitivity to change was demonstrated.

**Conclusions:** The results provide evidence of the feasibility and psychometric strengths of the RAS-DS. Although further research is required, the RAS-DS shows promise as a potential addition to the national suite of routine outcome measures.

## Keywords

Measure, mental illness, recovery, instrument development, feasibility

## Introduction

Adoption of a recovery-oriented approach to practice has become a prominent feature of mental health policies and system development plans internationally (e.g., Australian Health Ministers' Advisory Council's National Mental Health Working Group, 2009; Department of Health, 2007). Recovery from mental illness, as defined by people with lived experience, is different from cure or symptom amelioration and refers to "a way of living a satisfying, hopeful and contributing life even with limitations caused by illness. Recovery involves the development of new meaning and purpose in one's life as one grows beyond the catastrophic effects of mental illness" (Anthony, 1993: 12).

Recovery involves an individualised journey that each consumer embarks upon. The role of service provision is to be recovery-oriented in order to best facilitate or support consumers on their journey. Recovery-orientation involves a

<sup>1</sup>The University of Sydney, Sydney, Australia

<sup>2</sup>Richmond Fellowship of Queensland, Brisbane, Queensland, Australia

### Corresponding author:

Nicola Hancock, The University of Sydney, Cumberland Campus C42,  
PO Box 170, Lidcombe, NSW 1825, Australia.  
Email: nicola.hancock@sydney.edu.au

service culture that is based upon a belief that recovery is possible and indeed likely for people experiencing mental illness. Based on this belief, practices are strengths-based and foster hope, choice, self-determination, peer-to-peer relationships and positive risk-taking. Primary value is placed on each consumer or service user's perspective (e.g., Slade et al., 2014; Piat and Lal, 2012; Amering and Schmolke, 2009; Piat et al., 2010).

Whilst policies, plans and directives around recovery-orientation abound, there appears to be a lag in practice implementation or systemic change on the 'ground' (Piat and Lal, 2012; Slade et al., 2014). A growing body of literature reports innovative and recovery-oriented initiatives, however many services have been described as "*simply repackaging old wine in the new bottle of recovery language*" (Davidson et al., 2005: 481). The systemic dependence on measures of symptomology, social disability and service-use has been identified as one of the barriers to adoption of a recovery-oriented approach to practice internationally (Thornicroft and Slade, 2014).

In Australia, the outcome measures used routinely across state mental health sectors are predominantly symptom/illness or functioning focused (e.g., Health of Nations Outcome Scale (Wing et al., 1998)). In New South Wales, the only self-report measure included in the suite is the Kessler-10 (Andrews and Slade, 2001). However, consumers dislike its negative orientation and are reluctant to use it, as seen in low completion rates (Pirkis and Callaly, 2010).

Assessments (i.e., the data collected) set the tone for service delivery and influence service planning. Service providers, increasingly called upon to demonstrate effectiveness, are almost compelled to focus treatment on the outcomes they will measure. If the data collected focus on symptoms, then intervention is likely to focus on ameliorating symptoms rather than on recovery defined as living a good and productive life even in the presence of symptoms. Similarly, if data are measured only from the perspectives of clinicians, then service planning is unlikely to include consumers as equal partners. A structure that enables consumers to participate actively in assessment is a fundamental first step toward recovery-oriented practice. Active participation of consumers in their own assessment and treatment planning facilitates self-direction and empowerment, and these are repeatedly described as central recovery processes for the individual (Bird et al., 2014).

Recovery-oriented outcome measurement is required within the suite of instruments routinely used in mental health systems (Burgess et al., 2011; Andresen et al., 2010; Thornicroft and Slade, 2014). Since no one knows better than the person experiencing the illness about its effects, a psychometrically-sound self-report measure of recovery is crucial for services to be recovery-oriented.

A second, but related, major challenge with outcome measurement in Australian mental health services is variable completion rates, which severely restrict the capacity to

measure service-level outcomes. Whilst Australia has been at the forefront of international efforts to embed routine assessment into practice (Burgess et al., 2012), completion rates even of mandated outcome measures, remain variable (Kightley et al., 2010; Pirkis and Callaly, 2010). Slade et al. (1999) have suggested that feasibility of an instrument is critical to its uptake and use in routine outcome measurement. Feasibility includes: brevity; simplicity; relevance; acceptability; availability and value. Variable completion rates may well be due to a lack of perceived relevance and poor acceptability of current outcome measures (Patterson et al., 2006; Trauer et al., 2009). Both staff and consumers report that they complete forms for 'bureaucrats' rather than for their contribution to practice (Andrews and Page, 2005; Lakeman, 2004; Happell, 2008). The perceived lack of relevance to practice has been associated repeatedly with poor 'compliance' in outcome measurement (Patterson et al., 2006; Trauer et al., 2009; Trauer et al., 2006). Clearly there is a need for recovery-focused instruments to measure outcomes that are relevant, meaningful and practically useful to both consumers and staff.

### Existing measures of recovery

Over the past 15 years, researchers have worked to establish a sound measure of mental health recovery. In keeping with the principles of the recovery paradigm, recovery-based instruments are invariably self-report measures. To date no instrument has reached the status of a gold standard. Over the last few years a number of authors (Burgess et al., 2011; Shanks et al., 2013; Sklar et al., 2013) have reviewed existing recovery measures. With differing evaluation priorities and criteria, each reviewer reached slightly different conclusions about which measure was most useful or psychometrically strong. All reviews, however, highlighted that: a) limited ongoing assessment and testing of measures had been carried out for most measures after their initial development; b) the Recovery Assessment Scale (RAS) (Corrigan et al., 2004) was the most frequently used and psychometrically tested measure, and c) further work was required to identify a gold standard measure of recovery. The UK-developed Recovery Star (Dickens et al., 2012), not available at the time of the Australian review, has been found to be a useful tool to facilitate collaboration. However, it lacks evidence for the psychometric properties required for outcome measurement (Killaspy et al., 2012). Burgess et al. (2012) concluded that a new or modified recovery instrument might be required for use in Australian mental health services.

### Background development of the RAS-DS

With the aim of developing an instrument that is both practically useful/feasible and psychometrically sound, we began an iterative process of development. Consumers

partnering in our research selected the RAS as their preferred recovery measure (Hancock et al., 2012b). However, after testing, we found several measurement problems with the original instrument (Hancock et al., 2011). We set about reducing these with the RAS-DS. Specifically, we removed redundant and ill-fitting items deemed not to be part of the construct and modified the rating category structure. To address a ceiling effect, we conducted focus groups with consumers who had maximum scores on the RAS and created new items that targeted later stage recovery-based achievement including (a) accepting your illness and gaining control over symptoms, (b) self-love and optimism, (c) doing things for and experiencing pleasure, (d) contributing through meaningful activity, (e) having a diversity of friendships, (f) being needed and valued by others and (g) coming to terms with family relationships (Hancock et al., 2012a; Hancock et al., 2011). We then commenced data collection with a first version of the RAS-DS.

Early on in data collection, qualitative data regarding its feasibility or usefulness revealed the need for an additional point in the rating scale. We suspended data collection, re-defined the scale and piloted our new descriptors with a small group of consumers and staff. Once they indicated increased satisfaction with the modified scale, we re-commenced data collection. The revised RAS-DS has 38 items that consumers rate on a 4-point scale: “untrue”; “a bit true”; “mostly true” and “completely true”.

In this paper we examine its feasibility and psychometric properties. If the RAS-DS was found to be both feasible and psychometrically sound, it would be a potential candidate self-report measure of recovery for use in routine outcome measurement and to enhance consumer voice in the assessment process. Specifically we examined the following attributes of the RAS-DS:

1. Feasibility (brevity; simplicity; relevance and acceptability) as reported by consumers and staff.
2. Measurement properties, including internal validity, internal reliability and responsiveness to change.

## Methods

The results reported here emerged from two studies approved by the Human Research Ethics Committee of the University of Sydney. In Study 1, we explored feasibility or practical usefulness of the RAS-DS together with consumer–staff dyads from three large non-government organisations (NGOs) in two Australian states: New South Wales and Queensland. Each of these organisations provided a wide range of services to individuals living with mental illness in both rural and urban contexts. In Study 2, we explored the measurement properties of data gathered with the RAS-DS, utilising assessment data collected in Study 1 and additional data from two large Partners in Recovery (PIR) programs in New South Wales. PIR programs are a

federally-funded initiative to provide care coordination and support to individuals living with severe and persistent mental illness. The two PIR programs use the RAS-DS as a routine outcome measure and provided de-identified data to the study.

## Procedures and data analysis

### Study 1: Feasibility

**Participants.** Fifty-eight consumer-staff dyads participated in Study 1. As some staff partnered with more than one consumer, there were only 39 staff participants. Demographic data for participants is provided in Table 1.

**Data collection.** Consumers who gave informed consent completed the RAS-DS and then discussed their responses with their primary staff member. Both consumers and staff completed an open question about the overall usefulness of the RAS-DS. In addition, each consumer selected the best response to questions about time to complete (“15 minutes or less” or “more than 15 minutes”) and ease of completion (from “very hard” to “very easy”).

**Data analyses.** Responses from categorical questions were calculated as percentages. Comments were coded using constant comparative analysis (Glaser and Strauss, 1968; Charmaz, 2014). This is an inductive method in which codes are developed directly from the data rather than fitting the data into pre-existing codes. Each small chunk of data (word, phrase or sentence) was closely compared to those following to identify underlying concepts or codes. Codes were then compared to each other, and similar codes were grouped into themes. The first and third authors independently coded the full data set, then met to compare and discuss the codes and to find consensus.

### Study 2: Measurement properties

**Participants.** For this study, 298 consumer participants provided 324 data sets. All participants completed the RAS-DS once and 26 participants completed the RAS-DS a second time after 3 months. Participants included 55 of the 58 consumers from Study 1 and 243 consumers from PIR. Participant details are provided in Table 1. Across all collected demographic details there were minimal differences between groups. RAS-DS mean scores did however differ between groups. The PIR group had a significantly lower mean than the NGO group.

**Data analyses.** We used Rasch analysis to examine evidence for validity and reliability of data gathered with the RAS-DS. Rasch analysis is based on two core assumptions (Bond and Fox, 2007; Wright and Stone, 1999). As they apply in this study: all people are more likely to affirm items that are easier to endorse and people who are

**Table 1.** Description of consumer participants.

		NGO consumers	PIR consumers	TOTAL
Gender	Female	28 (48.3%)	128 (52.7%)	156 (51.8%)
	Male	28 (48.3%)	112 (46.1%)	140 (46.5%)
	Unspecified	2 (3.4%)	3 (1.2%)	5 (1.7%)
	Total	58 <sup>a</sup>	243	301
Age		M=42.3; SD=9.8	M=41.8; SD=12.0	
Diagnoses <sup>b</sup>	Anxiety/PTSD /OCD	11 (19.0%)	30 (12.3%)	41 (13.6%)
	Affective/mood disorder	11 (19.0%)	91 (37.4%)	102 (33.9%)
	Schizophrenia/schizo affective disorder	33 (56.9%)	91 (37.4%)	124 (41.2%)
	Personality disorder	4 (6.9%)	30 (12.3%)	34 (11.3%)
	Unspecified	–	41 (16.9%)	41 (13.6%)
RAS-DS Measure Score <sup>c</sup>		M=59.6; SD=11.4	M=51.1; SD=11.2	M=52.7; SD=11.7

<sup>a</sup>missing RAS-DS data from 3 participants means that only 55 data sets used in Study 2; <sup>b</sup>some participants had more than one diagnosis; <sup>c</sup>RAS-DS Mean Measure Score differed significantly between NGO and PIR groups ( $t=5.1, p<.001$ ).

PTSD = Post Traumatic Stress Disorder, OCD = Obsessive Compulsive Disorder, Unspecified = unspecified or 'mental illness' with no specific diagnosis given.

more recovered are more likely than people who are less recovered to affirm items that are hard to endorse. All Rasch analysis procedures were conducted using the program Winsteps (Version 3.75.1, Linacre, 2014) and all other statistical procedures were performed using SPSS (Version 17.0). There was a small amount of missing data in the dataset. Where this was the case, raw scores were averaged from the available items. Rasch analysis procedures are very robust to missing data (Bond and Fox, 2007), so no special management or preparation was required.

**Internal validity.** Rasch analysis provides evidence for a number of different aspects of construct validity. Firstly, we examine point-measure correlations for each item to insure they were positive. Second, we examined goodness of fit statistics for each item (expressed in mean squares [MnSq]) to determine the extent to which data from each item conformed to the expectations of the Rasch model. MnSq values of  $\leq 1.5$  were considered acceptable (Wright and Linacre, 1994). Third, we examined the extent to which the overall spread of items and matched the "recovery levels" of the sample to ensure there were neither gaps which would reduce measurement precision nor item redundancy and that the mean of person measure scores was close to the mean of item measures.

**Reliability.** The Winsteps program provides evidence for several aspects of internal reliability. Two reliability indices, person and item, describe measurement precision;  $\geq .80$  was deemed acceptable (Fisher, 2007). A

person separation statistic provides evidence of the number of levels of recovery represented in the data. To conclude that measure differences are the result of real differences in recovery and not measurement error, the separation should be  $\geq 2.0$  (Fisher, 2007).

**Category structure.** If Rasch assumptions are met, then people scoring "1" on any item would have an average instrument measure lower than those scoring "2" and so on. Ideally, scores are spread across the range and all categories are used regularly ( $n \geq 10$ ). If any of these was violated, then we considered collapsing scoring categories until the assumptions were met.

**Correlations between Rasch-calibrated measure score and raw scores.** For the RAS-DS to have maximal clinical utility, scores would be able to be totaled and interpreted without the need for transformation or Rasch-based calibrations. To examine whether this was possible with the RAS-DS we explored a range of indicators: Pearson's Product Moment and Spearman Rank Order correlations between raw scores and Rasch-calibrated measure scores and the magnitude of changes in rank between raw-score and Rasch-calibrated measure scores.

#### Data analyses

**Ability to detect change over time.** We used a paired *t*-test to examine whether RAS-DS scores were typically higher at Time 2 than at Time 1 in a subgroup of participants ( $n=26$ ) who completed the RAS-DS twice. Time 2 measures were completed approximately three months follow-

ing Time 1 measures. The  $\alpha$  value was set at .05. If the paired  $t$ -test indicated that Time 2 measures were typically higher than Time 1 measures, this would provide some initial evidence of the potential of the RAS-DS to measure change over time.

## Results

### Study 1: Feasibility

**Brevity and simplicity.** The majority of consumers ( $n = 41$ , 71%) reported that the RAS-DS took 15 minutes or less to complete. Most rated it as “easy” or “very easy” to use ( $n = 43$ , 74%). Seven consumers (12%) responded in the space between the “easy” and “hard”. Only 8 (14%) reported that the RAS-DS was “hard” or “very hard”; most of those explained the difficulty as stemming from the challenging nature of the questions (e.g., “a bit personal and confronting,” “making deep decisions,” “having to give a great deal of thought to some questions”) rather than problems with the instrument.

**Relevance and acceptability.** The themes in the data from consumers and staff regarding relevance and acceptability are provided in Table 2. These are exemplified with quotes. All themes were positive except for the last three. Numbers of staff and consumers who gave equivalent comments are included to provide a sense of frequency. Of the total number of comments made, 80% of consumer comments and 81% of staff comments were positive.

### Study 2: Measurement properties

**Internal validity.** Point-measure correlations for all items were positive, ranging from .42 to .70. Item fit statistics are shown in Table 3. Fit statistics for 36 of the 38 items from the RAS-DS were within the acceptable range (i.e., MnSq  $\leq 1.5$ ). The two items with poorer fit statistics were “I have friends who have also experienced mental illness” and “I have friends without mental illness.” Outfit MnSq values for these items were 1.97 and 1.64 respectively.

The mean measure score for people was 52.8 and the mean measure for the items was 50 indicating a good match between item difficulty and the level of people’s self-rated recovery. Figure 1 illustrates the relationship between scores of the people and item difficulty. Figure 1 also shows a good representation of items in a 2 standard deviation band from the mean of the items. However the recovery level of 14.5% of the sample is greater than the difficulty of the hardest item. Similarly, the overall scores of 9.3% of participants are lower than the difficulty level of the easiest item. Measure scores for the most and least recovered individuals are less precise than those in the middle because there are no items exactly at their levels.

**Reliability.** Participant and item reliability indices were .93 and .98 respectively. Cronbach’s  $\alpha$  score (calculated using

raw scores) is .96. The person separation index was 3.53, suggesting the RAS-DS can separate people into approximately five statistically distinguishable groups (Wright and Masters, 2002).

**Category structure.** Scoring categories (i.e., 1, 2, 3, 4) progressed in an orderly fashion for all items. Apart from infrequent use of the lowest category (“untrue”) in some of the “easiest” items (e.g., “It is important to have healthy habits”), all categories were typically used with sufficient frequency. The overall percentages of use for each category were well balanced: 1 (16%), 2 (29%), 3 (31%) and 4 (24%).

**Correlations between Rasch-calibrated measure score and raw scores.** The Pearson Product Moment and Spearman Rank Order correlations between raw scores and Rasch-calibrated measure scores were  $r = .95$ ,  $p < .001$  and  $r_s = 1.00$ ,  $p < .01$ , respectively. A very small proportion (11 of 324, 3.4%) of rankings changed more than 3 places when raw scores were converted to Rasch-calibrated measure scores. Considered together, these results suggest that raw scores can be directly interpreted in the clinical setting without the need for transformation. However, it is recommended that raw scores are converted to Rasch-calibrated measure scores prior to further instrumentation analysis of the RAS-DS to ensure optimal accuracy (Linacre, 1998).

**Ability to detect change over time.** The paired  $t$ -test demonstrated that RAS-DS Rasch-adjusted measure scores at Time 2 ( $M = 54.8$ ,  $SD = 10.2$ ) were significantly higher than at Time 1 ( $M = 52.0$ ,  $SD = 9.8$ ),  $t = 2.49$ ,  $p = 0.020$ .

## Discussion

The purpose of this study was to examine the feasibility of use and the measurement properties of data gathered with the RAS-DS, an instrument which has been developed over several iterations from the original RAS (Corrigan et al., 2004). Results presented in this paper provide solid initial evidence for both aspects of the RAS-DS.

Most consumers and staff found the RAS-DS to be an acceptable and relevant instrument. Some commented on its value for facilitating recovery-focused goal setting. Both consumers and staff highlighted the new knowledge and/or awareness that it provided individually and as a partnership. There were, however, some consumers who did not find the RAS-DS relevant or acceptable. These consumers found it either too easy because their recovery was progressing well, or too confronting and personal. This raises two questions or points of discussion. Firstly, is it possible to find any single measure that will meet every individual need and fit every individual’s understanding of his or her experience of mental illness and recovery? The literature describes recovery as a deeply personal and unique process and perhaps therefore there will be no single measure that

**Table 2.** Relevance and acceptability.

Themes	# cons	# staff	Sample quotes
Made consumer think or enhanced their self-understanding	15	13	"It helps me understand my life and my recovery journey better" (consumer) "It gave her a wake up call" (staff)
Useful for goal setting/ recovery planning	6	14	"Sets a template for recovery planning" (consumer) "It helps us to see what we can work on to help improve mental health" (staff)
Positive feelings for consumer (e.g., interest, engagement, enjoyment)	4	4	"I enjoyed answering the questions" (consumer) "It is so nice to see smiles on my clients face when they can relate to a statement and we can talk about it." (staff)
Prompted discussion between consumer and staff	1	20	"It helped to talk about stuff" (consumer) "I found it useful as a conversation starter" (staff)
Helped MH worker to understand consumer	1	11	"My case manager understood me and I understood me" (consumer) "Very useful. Got to know a lot about them and their recovery" (staff)
Helped to track progress in recovery	1	6	"It help me notice a big change in the way I think about recovery – positive change" (consumer) "It encouraged the client to reflect on the progress made over the past six months" (staff)
Other, non-specific positive comments (e.g., it was relevant, useful etc)	15	6	"I was happy with the survey questions – not too personal" (consumer) "I think it's a great tool" (staff)
Caused negative feelings for consumer (e.g., sadness, discomfort)	6	8	"Some questions were uncomfortable to answer. I felt put on the spot" (consumer) "For clients that are isolated and have no friends or family, this survey is very upsetting" (staff)
Not useful	1	8	"Not very useful as I am doing well in my recovery"(consumer) "It didn't really reveal anything new and significant" (staff)
Questions unsatisfactory (e.g., too broad or irrelevant)	4	1	"Inadequate survey" (consumer) "The consumer found that some of the questions were the same"

will be acceptable to all. The addition of harder items might make the instrument more relevant for those further along their recovery journey; however this may well make the instrument even less acceptable for those who found it too hard and confronting. A repeated discussion in the recovery literature on options and choice comes to mind (e.g., Davidson et al., 2006). Second, some consumers said that completing the RAS-DS made them sad and teary. Acknowledging the importance of choice, should we avoid sad and teary? Perhaps this means that we are in fact asking people about things that really matter to them? Perhaps rather than avoiding asking these questions, the challenge is for those supporting consumers to listen, empathise and facilitate them to consider ways to work on what really matters to them?

The instrument was found to be feasible for most in terms of brevity and simplicity. There were however some consumers who took longer than 15 minutes to complete

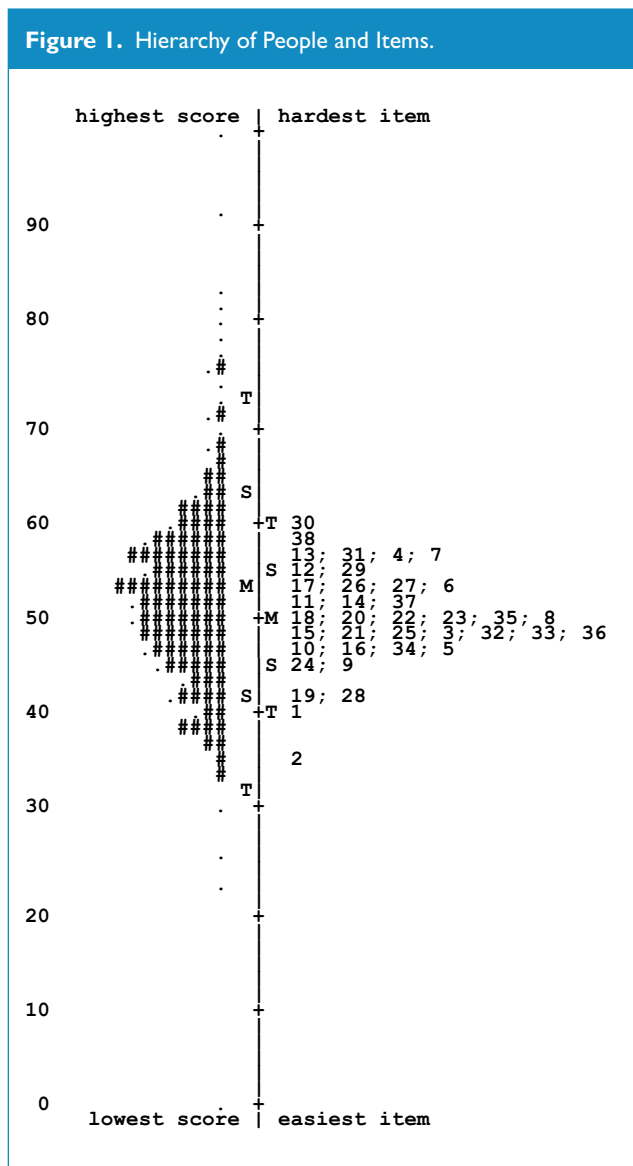
the RAS-DS. For most this was because they had to think deeply about their responses and we consider this to be a positive reason as long as they found it useful overall. For some however, again a small number, it was hard to concentrate and to read all of the items. While we have worked to keep the language as user-friendly as possible, we understand that for some people, having someone to read the statements out to them while they assign the scores will aid completion. This form of administration aids simplicity when needed but, as with other self-report measures, to aid validity and reliability, we caution against "helicoptering" (i.e., hovering over consumers as they rate) or re-wording or re-interpreting the statement. Consumers should be encouraged to rate the statement based on what it means to them (Coombs, 2005).

If, as (Slade et al., 1999) suggested, feasibility is critical to the uptake of measures, then the RAS-DS should be easy for mental health services to adopt. It is freely accessible,

**Table 3.** RAS-DS Item Measure and Fit Statistics.

Abbreviated item	M	SE	Infit MnSQ	Outfit MnSQ
1. It is important to have fun	41.0	.8	1.18	1.17
2. It is important to have healthy habits	35.9	.9	1.16	1.20
3. I do things that are meaningful to me	48.9	.7	.97	.96
4. I continue to have new interests	56.1	.7	1.14	1.12
5. I do things ... valuable and helpful to others	47.2	.7	1.02	1.07
6. I do things that give me ... great pleasure	53.4	.7	.96	.97
7. I can handle it if I get sick again	56.8	.7	1.02	1.02
8. I can help myself become better	49.9	.7	.69	.69
9. I have the desire to succeed	45.5	.7	.86	.86
10. I have goals in life that I want to reach	46.0	.7	1.07	1.07
11. I believe ..I can reach ... personal goals	51.9	.7	.77	.75
12. I can handle what happens in my life	55.1	.7	.72	.76
13. I like myself	55.8	.7	.88	.92
14. I have a purpose in life	51.4	.7	.94	.91
15. If people really knew me they would like me	48.8	.7	.97	1.01
16. If I keep trying, I will continue to get better	46.5	.7	.70	.68
17. I have an idea of who I want to become	52.6	.7	1.18	1.22
18. Something good will eventually happen	49.4	.7	.80	.78
19. I am ...most responsible for my own improvement	42.3	.8	.80	.77
20. I am hopeful about my own future	49.8	.7	.66	.66
21. I know when to ask for help	48.8	.7	1.03	1.05
22. I ask for help, when I need it	49.5	.7	1.04	1.01
23. I know what helps me get better	50.3	.7	.77	.74
24. I can learn from my mistakes	45.1	.7	.80	.82
25. I... identify the early warning signs ...	48.7	.7	1.18	1.20
26. I have ... plan for how to stay or become well	53.4	.7	.99	.99
27. There are things ... that help me deal with ...symptoms	52.7	.7	.76	.75
28. I know ... mental health services that help me	42.3	.8	.95	.94
29. Although ... symptoms may get worse, can handle it	55.6	.7	.87	.89
30. My symptoms interfere ... less with my life	60.0	.8	1.06	1.03
31. My symptoms ...problem for shorter periods ...	57.2	.7	.92	.91
32. I have people that I can count on	48.2	.7	1.14	1.09
33. ...when I don't believe in myself, other people do	49.0	.7	1.10	1.21
34. It is important to have a variety of friends	46.7	.7	1.30	1.34
35. ...friends who ...also experienced mental illness	50.1	.7	1.58	1.97
36. I have friends without mental illness	47.7	.7	1.51	1.64
37. I have friends that can depend on me	51.9	.7	1.25	1.21
38. I feel OK about my family situation	58.4	.7	1.43	1.41

Items presented in original order. M = measure, SE = standard error, MNSQ = mean square.



M = mean, S = 1 standard deviation, T = 2 standard deviations, # refers to 3 participants, . refers to 1 or 2 participants.

(<http://ses.library.usyd.edu.au/handle/2123/9317>), quick to administer with no training required. A user-manual is freely available from the corresponding author. It is readily understood by most consumers and contributes to positive relationships, understanding and goal setting. Further, the high correlation between raw scores and Rasch-generated measure scores suggests that raw scores are adequate for describing recovery. This also contributes to its acceptability and simplicity.

Examination of the measurement properties of the RAS-DS undertaken as part of this project has provided promising initial evidence for its suitability as a measure of mental health recovery. Examination of internal validity suggests that items are all related to the underlying construct of recovery. Recovery has been repeatedly described

in the literature as a multi-dimensional journey that is unique for each individual. Indeed the title of this instrument includes the word ‘domains’ to reflect the breadth of the construct of recovery that has been described by consumer participants during the iterative cycles of instrument development (Hancock, 2012a). Consumers described functional, personal, social and clinical aspects. We began our testing of the instrument by assuming that while each individual is unique, the RAS-DS comprises items that describe 4 aspects or domains of a broader unidimensional hierarchy of mental health recovery. We then tested that hypothesis and found good evidence that it is true. Of course, no construct is perfectly unidimensional but the evidence for multiple dimensions found was actually relatively weak given the good fit of the items. Rather, our analyses demonstrate that these four domains are contained within a larger single recovery construct.

Most items were used consistently by consumers. The two items which demonstrated poorest “fit” to the expectations of the Rasch model (“I have friends who have also experienced mental illness” and “I have friends without mental illness”) were added recently as a result of a study that explored the perspectives of consumers considered to be at “later stages” of their recovery (Hancock et al., 2012a). Given that consumers identified these as important aspects of their conceptualisation of recovery, and given the ‘border-line’ nature of the fit statistics meaning that they do not denigrate the data, we believe it would be premature to suggest that these be removed.

Results from the investigation of rating scale structure and reliability indicators also suggest that the RAS-DS is useful in distinguishing differences in recovery between more and less recovered individuals. The term ‘stages’ is included in the title of this instrument as a result of previous work aiming to rectify the lack items addressing later stage recovery achievements (Hancock, 2012a) in the original RAS (Corrigan, 2004). The hierarchy of items and the reliability of the rating scale structure demonstrate that there are earlier and later recovery achievements. However, as can be seen by the essentially normal curve of people distributions in Figure 1, there is a fluidity of stages rather than distinct or discrete steps. This reflects the broader and fluidic stages of recovery that are described in previous literature (e.g., Andresen et al., 2006). Even with the fluidic nature of these stages, our data demonstrates that there were 5 stages or groups of relative recovery within our sample.

Although only preliminary, the examination of change in RAS-DS scores over time is also promising. In a previous iteration of the instrument, results of an examination of test-re-test reliability with a small number of people demonstrated that when changes occurred, the direction of changes were consistent with reported life-changes (Hancock, 2012b). Additionally, in this study, the good item fit statistics suggests that changes are a result of



changes in participants' self-reported measure of recovery rather than instrument instability. However, further examination is required to ensure that this change shown over time is the result of actual change in level of recovery rather than any measurement error caused by scale instability.

Considered together, these results suggest that the RAS-DS holds great promise as an instrument to measure mental health recovery. If further testing continues to demonstrate the feasibility of the RAS-DS and the soundness of the data it generates, it would be a strong candidate for routine adoption within the Australian mental health system.

One final point deserves discussion. As can be seen in Figure 1, a number of items are at the same level of difficulty (i.e., redundancy; Baghaei, 2008). If the aim was to develop the briefest possible measure of recovery that generated psychometrically strong data, these items should be examined for potential removal. The development of a 'short-form' version of the RAS-DS that can be used as a screening tool with the long-form reserved for therapeutic purposes is one potential future endeavour. Perhaps more valuable would be the development of a computer assisted version of the RAS-DS that would enable the richness of the domains of recovery to be maintained, but reduce burden by sparing consumers from completing items that were not relevant to their current stage of recovery. However, caution needs to be taken in attempts to reduce burden. As demonstrated in the qualitative, first component of this study, the majority of consumer participants did not find the RAS-DS burdensome. Further, both consumers and staff valued the opportunity to reflect and discuss their positive achievements to date as well as the possible successes ahead. For now we have chosen to retain all the items and will continue to explore the possibility of additional items at the most difficult and least difficult ends of the hierarchy.

There are limitations with this study and further work is required. In terms of the psychometric testing, this study has focused upon the internal workings of the measure. Whilst it may be challenging given there is no gold-standard measure of recovery, future work is required to examine aspects of convergent as well as divergent validity. Study examining further aspects of validity such as criterion and known-groups would also contribute to our understanding of the RAS-DS and its capacity to measure the recovery construct. In relation to the sample, all of the data have been collected from consumers and staff working within the non-government or community-managed mental health sector. Additional study is needed to ensure findings are replicated when the RAS-DS is used within other service sectors such as in-patient contexts. Further, data were all collected within the New South Wales and Queensland states of Australia and we did not record the cultural background of participants. Testing in other states and within other countries is needed to ensure that data gathered in other places are equally robust and to examine how culture

affects use of the RAS-DS. Finally, the small data set used to examine sensitivity to change shows positive trends over a relatively short period of time. This however requires further testing.

## Conclusions

These promising results suggest that the RAS-DS is an efficient and effective instrument that (a) generates valid and reliable scores reflecting mental health recovery and (b) promotes understanding and partnerships between consumers and staff as a basis for collaborative intervention planning. While further testing is warranted, it appears that the RAS-DS can make an important contribution to service delivery in mental health.

## Acknowledgements

We thank the mental health consumers and staff who not only completed the RAS-DS but provided the rich expert knowledge required to understand the relevance and acceptability of the RAS-DS. We thank the Richmond Fellowship Queensland, the Schizophrenia Fellowship of New South Wales, RichmondPRA and both the Hunter and South Western Sydney PIR programs for their participation.

## Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## References

- Amering M and Schmolke M (2009) *Recovery in mental health: reshaping scientific and clinical responsibilities*. West Sussex: John Wiley & Sons.
- Andresen R, Caputi P and Oades L (2006) The Stages of Recovery Instrument: Development of a measure of recovery from serious mental illness. *Australian and New Zealand Journal of Psychiatry* 40: 972–980.
- Andresen R, Caputi P and Oades L (2010) Do clinical outcome measures assess consumer-defined recovery? *Psychiatry Research* 177: 309–317.
- Andrews G and Page AC (2005) Outcome measurement, outcome management and monitoring. *Australian and New Zealand Journal of Psychiatry* 39: 649–651.
- Andrews G and Slade T (2001) Interpreting scores on the Kessler psychological distress scale (K10). *Australian and New Zealand Journal of Public Health* 25: 494–497.
- Anthony W (1993) Recovery from mental illness: The guiding vision of the mental health service system in the 1990s. *Psychosocial Rehabilitation Journal* 16: 11–23.
- 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*.
- Baghaei P (2008) The Rasch model as a construct validation tool. *Rasch Measurement Transactions* 22: 1145–1146.
- Bird V, Leamy M, Tew J, Le Boutillier C, Williams J and Slade M (2014). Fit for purpose? Validation of a conceptual framework for personal

- recovery with current mental health consumers. *Australian and New Zealand Journal of Psychiatry* 48: 644–653.
- Bond TG and Fox CM (2007) *Applying the Rasch model: Fundamental measurement in the human sciences*, Mahwah, NJ: Lawrence Erlbaum.
- Burgess P, Coombs T, Clarke A, et al. (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: 4.
- Burgess P, Pirkis J, Coombs T, et al. (2011) Assessing the value of existing recovery measures for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry* 45: 267–280.
- Charmaz K (2014) *Constructing grounded theory* (2nd ed). London: SAGE.
- Coombs T (2005) *Australian Mental Health Outcomes and Classification Network: Behaviour and Symptom Identification Scale (BASIS) - 32 Training Manual NSW Institute of Psychiatry*.
- Corrigan P, Salzer M and Ralph RO (2004) Examining the factor structure of the recovery assessment scale. *Schizophrenia Bulletin* 30: 1035–1041.
- Davidson L, O'Connell M, Tondora J, et al. (2006) The top ten concerns about recovery encountered in mental health system transformation. *Psychiatric Services* 57: 640–645.
- Davidson L, O'Connell MJ, Tondora J, et al. (2005) Recovery in Serious Mental Illness: A New Wine or Just a New Bottle? *Professional Psychology - Research & Practice* 36: 480–487.
- Department of Health (2007) *Commissioning framework for health and well-being*. Available at: [www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_072604](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_072604).
- Dickens G, Weleminsky J, Onifade Y, et al. (2012) Recovery Star: validating user recovery. *The Psychiatrist* 36: 45–50.
- Fisher WP (2007) Rating scale instrument quality criteria. *Rasch Measurement Transactions* 21:1095.
- Hancock N, Bundy A, Honey A, et al. (2012a) Measuring the Later Stages of the Recovery Journey: Insights Gained from Clubhouse Members. *Community Mental Health Journal*: 1–8.
- Hancock N, Bundy A, Honey A, et al. (2011) Improving measurement properties of the Recovery Assessment Scale with Rasch analysis. *American Journal of Occupational Therapy* 65: e77–e85.
- Hancock N, Bundy A, Tamsett S, et al. (2012b) Participation of mental health consumers in research: Training addressed and reliability assessed. *Australian Occupational Therapy Journal* 59: 218–224.
- Happell B (2008) Determining the effectiveness of mental health services from a consumer perspective: Part 1: Enhancing recovery. *International journal of mental health nursing* 17: 116–122.
- Kightley M, Einfeld S and Hancock N (2010) Routine outcome measurement in mental health: feasibility for examining effectiveness of an NGO. *Australasian Psychiatry* 18: 167–169.
- Killaspy H, White S, Taylor TL, et al. (2012) Psychometric properties of the Mental Health Recovery Star. *The British Journal of Psychiatry* 201: 65–70.
- Lakeman R (2004) Standardized routine outcome measurement: Pot holes in the road to recovery. *International Journal of Mental Health Nursing* 13: 210–215.
- Linacre J (2014) Winsteps® Rasch measurement computer program. Beaverton, Oregon: Winsteps.com.
- Linacre JM (1998) Do correlations prove scores linear? *Rasch Measurement Transactions* 12: 605–606.
- Patterson P, Matthey S and Baker M (2006) Using mental health outcome measures in everyday clinical practice. *Australasian Psychiatry* 14: 133–136.
- Piat M and Lal S (2012) Service providers' experiences and perspectives on recovery-oriented mental health system reform. *Psychiatric Rehabilitation Journal* 35: 289.
- Piat M, Sabetti J and Bloom D (2010) The transformation of mental health services to a recovery-orientated system of care: Canadian decision maker perspectives. *International Journal of Social Psychiatry* 56: 168–177.
- Pirkis J and 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.
- Shanks V, Williams J, Leamy M, et al. (2013) Measures of personal recovery: a systematic review. *Psychiatric Services* 64: 974–980.
- Sklar M, Groessler EJ, O'Connell M, et al. (2013) Instruments for measuring mental health recovery: a systematic review. *Clinical Psychology Review* 33: 1082–1095.
- Slade M, Amering M, Farkas M, et al. (2014) Uses and abuses of recovery: implementing recovery-oriented practices in mental health systems. *World Psychiatry* 13: 12–20.
- Slade M, Thornicroft G and Glover G (1999) The feasibility of routine outcome measures in mental health. *Social psychiatry and psychiatric epidemiology* 34: 243–249.
- Thornicroft G and Slade M (2014) New trends in assessing the outcomes of mental health interventions. *World Psychiatry* 13: 118–124.
- Trauer T, Callaly T and Herrman H (2009) Attitudes of mental health staff to routine outcome measurement. *Journal of Mental Health* 18: 288–297.
- Trauer T, Gill L, Pedwell G, et al. (2006) Routine outcome measurement in public mental health—what do clinicians think? *Australian Health Review* 30: 144–147.
- Wing J, Beevor A, Curtis R, et al. (1998) Health of the Nation Outcome Scales (HoNOS). Research and development. *The British Journal of Psychiatry* 172: 11–18.
- Wright BD and Masters GN (2002) Number of person or item strata (4G11)/3. *Rasch Measurement Transactions* 16: 888.
- Wright BD and Stone M (1999) *Measurement essentials*. Wilmington, Delaware: Wide Range, Inc.
- Wright BD and Linacre JM (1994) Reasonable mean square fit values. *Rasch Measurement Transactions* 8: 370.