Full Title: Refinement and Revalidation of the Demoralization Scale: The DS-II–Internal Validity

Running Title: DS-II: Internal validity

Authors:

Sophie Robinson[1,2] (BPsych(Hons)), David. W. Kissane[1,3,4,6] (MBBS, MPM, MD, FRANZCP, FAChPM, FAPM), Joanne Brooker[1,3] (BSc(Hons), GradDipPsych, GradDipComp, PhD), Natasha Michael[4,7](MBChB, FRACP, FAChPM, MRCPI, MRCGP, MSc), Jane Fischer[5] (MBBS, DipChildHealth, PGradDipMedPallCare, FAChPM), Michael Franco[6,7](MBBS, FRACP, FAChPM), Courtney Hempton[1,3] (BBSc(Hons), MBioethics), Merlina Sulistio[4] (MBBS, BMedSc, FRACP, FAChPM), Julie F. Pallant[8] (PhD), David M. Clarke[1] (MBBS, MPM, PhD, FRACGP, FRANZCP), and Sue Burney[2,3] (BA., Grad.Dip.Couns.Psych., MHlthSc., PhD).

[1] Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia

[2] School of Psychological Sciences, Monash University, Clayton, Australia

[3] Szalmuk Family Psycho-oncology Unit, Cabrini Institute, Malvern, Australia

[4]Cabrini Palliative Care Service, Cabrini Health, Prahran, Australia

[5]Calvary Health Care Bethlehem, Caulfield, Australia

[6]Supportive and Palliative Care Unit, Monash Health, Clayton, Australia

[7]Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia

[8]Department of Rural Health, The University of Melbourne, Melbourne, Australia

Address correspondence to:

David W. Kissane, MD, MPM, FRANZCP, FAChPM, FAPM

Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Level 3, P Block, Monash Medical Centre, 246 Clayton Road, Clayton, Victoria 3186, Australia.

Phone: +61 3 9594 1477;

Email: david.kissane@monash.edu

Total number of:

Text pages=20

Tables=4

Figures=3

Funding sources/grants: Bethlehem Griffiths Research Foundation

Conflicts of interest: None

Author Contributions: Sophie Robinson: Conceptualization, methodology, validation, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, and project administration. **David. W. Kissane:** Conceptualization, methodology, validation, formal analysis, investigation, resources, writing – original draft, writing – review and editing, visualization, supervision, project administration, and funding acquisition. **Joanne Brooker:** Conceptualization, methodology, validation, formal analysis, writing – review and editing, and supervision. **Natasha Michael:** Conceptualization, methodology, investigation, writing – review and editing, visualization, supervision, and project administration. **Jane Fischer:** Conceptualization, methodology, resources, and writing – review and editing. **Michael Franco:** Conceptualization, methodology, investigation, writing – review and editing, and supervision. **Courtney Hempton:** Investigation, data curation, writing – review and editing, and project administration. **Merlina Sulistio:** Conceptualization, methodology, investigation, writing – review and editing, visualization, supervision, and project administration. **Julie F. Pallant:** Methodology, formal analysis, writing – original draft, writing – review and editing, visualization, and supervision. **David M. Clarke:** Conceptualization, methodology, validation, formal analysis, writing – review and editing, and supervision. **Sue Burney**: Conceptualization, writing – review and editing, and supervision.

Precis: The DS-II is a self-report scale comprising 16 items and two components. It is an improved and more practically attractive measure of demoralization.

Keywords: psychometrics, cancer, reliability, validity, adjustment, coping behaviour, demoralization, Rasch modeling.

Abstract:

Background: The Demoralization Scale (DS) was initially validated in 2004 to enable measurement of demoralization in advanced cancer patients. Subsequent shortcomings indicated the need for psychometric strengthening. Here we report on the refinement and revalidation of the DS to form the DS-II, specifically reporting the scale's internal validity.

Methods: Palliative care patients with cancer or other progressive diseases (n = 211) completed a revised version of the 24-item Demoralization Scale and a measure of symptom burden (Memorial Symptom Assessment Scale). Exploratory factor analysis and Rasch modeling were used to evaluate, modify, and revalidate the scale, providing information about dimensionality, suitability of response format, item fit, item bias, and item difficulty. Test-retest reliability was examined for 58 symptomatically stable patients at a five-day follow-up.

Results: Exploratory factor analysis supported a 22-item, two-component model. Separate Rasch modeling of each component resulted in collapsing the response option categories and removing three items from each component. Both final 8-item subscales met Rasch model expectations and were appropriate to sum as a 16-item total score. The DS-II demonstrated internal consistency and test-retest reliability (Meaning and Purpose: $\alpha = 0.84$, ICC = .68; Distress and Coping Ability: $\alpha = 0.82$, ICC = .82; Total: $\alpha = 0.89$, ICC = 0.80).

Conclusion: The DS-II is a three-point-response, self-report scale, comprising 16 items and two subscales. Given our revalidation, psychometric strengthening, and simplification, the DS-II is an improved and more practical measure of demoralization for research and clinical use. External validation of the DS-II will be reported subsequently.

Demoralization has become increasingly recognized in palliative care as a clinical issue requiring assessment and treatment.[1,2] Understood as a state of maladaptive coping, demoralization develops with symptoms of hopelessness and helplessness associated with loss of purpose and meaning in life.[1] In a recent systematic review of 25 studies, clinical prevalence rates for demoralization ranged from 13-18% in patients with progressive diseases like cancer.[3] The morale of any patient fluctuates dimensionally from optimism to mild disheartenment, to greater despondency, to potentially deep despair, which can be associated with a desire for hastened death.[4] Thus, the importance of measuring demoralization has been emphasized with reference to risk of suicide and its potential relevance in end of life decision-making.[1]

Access to a psychometrically sound measure aids the clinical assessment of demoralization.[5] Our preliminary validation of the Demoralization Scale (DS) in 2004 created a 24-item self-report scale, which proved to be a useful measure of a poor coping response.[3] The DS was translated into several languages and further validated with traditional classical test theory (CTT) approaches in four studies.[2,6-8] Psychometric evaluation consistently showed convergent validity with established measures of psychological distress, quality of life, and desire for death, and strong internal reliability for the total scale.[2,5-8] Discriminant validity in relation to depression was more difficult to establish.[3] It's factor structure varied between four and five factors and the test-retest reliability of the scale was not examined.[3] Overall, further validation was required given the inconclusive findings.[5,8]

To further examine the psychometric properties of the DS, the use of item response theory (IRT) models has been recommended.[3] In recent decades, these models have gained popularity in assisting scale development and refinement.[3,9] IRT techniques employ mathematical models to examine the performance of each item and respondent in a scale, with the Rasch model most widely used.[10] In the Rasch

model, dimensionality, category ordering, and item bias (differential item functioning)[11] are tested, while scale length can be reduced as information about items that overlap in difficulty level is provided.[10] From clinical experience, the length of the DS appeared burdensome for some patients. Additionally, Rasch modeling has highlighted the limitations of reversed items as these lead to confusion for respondents, thereby reducing reliability of responses.[12]

Given these issues with the psychometric properties of the DS, scale evaluation, modification and refinement was indicated with a palliative care population. Palliative care is delivered to patients with progressive disease, with advanced cancer the predominate presentation. Progressive disease typically brings increased existential challenge and therefore greater risk of demoralization. We report the revalidation process in two parts. Here, use of CTT and Rasch modeling guide the development of the DS-II, describing its internal validity and reliability. Its external validity (convergent and divergent) is reported in a companion paper (Robinson et al., under joint review CNCR-15-2644).

Materials and Methods

Study Design and Patients

This was a multi-site observational study with ethical approval from each site's Human Research and Ethics Committee. Recruitment occurred from June 2013 to November 2014 in acute, metropolitan hospitals (Monash Health, Cabrini Health, Calvary Health Care Bethlehem), Melbourne, Australia. Patients were eligible if they had advanced cancer (Stage 4, with prognosis < 1 year) or progressive disease of any type (neurological, cardiac, respiratory, etc.), and ineligible if they were cognitively impaired, unable to provide consent or lacked sufficient English. Time since diagnosis, recurrence, and currently being on treatment or being hospitalized were not

inclusion criteria. The study population was typical of the profile of patients seen in palliative care programs. Treating physicians determined patients' eligibility.

Measures

Socio-demographic and medical details included primary diagnosis, illness duration, Karnofsky rating, inpatient status, age, sex, marital status, religion, educational achievement, and employment status.

Memorial Symptom Assessment Scale (MSAS) was employed to measure symptom burden by assessing the presence, frequency, severity, and associated distress of 32 symptoms (24 physical and 8 psychological items) over the past week.[13] The scale has demonstrated satisfactory reliability ($\alpha = 0.82$) and is well validated in cancer patients.[13]

Demoralization Scale-II (DS-II). In light of the research indicating reversed items can reduce the reliability of responses,[12] consultation with an expert in the field of outcome evaluation in cancer treatment research (David Cella) was undertaken. After IRT review and confirmation, the five reversed items in the DS were converted to the same valence as the other 19 items. Prior to further scale modification detailed below, this revised DS contained 24 items that were rated on the original 5-point Likert scale, ranging from 0 (*never*) to 4 (*all the time*), with higher scores indicative of higher levels of demoralization. The original validation of the DS found five factors: Loss of Meaning and Purpose; Dysphoria; Disheartenment; Helplessness; and Sense of Failure.[5]

Procedure

Demographic and medical data were obtained from participants' medical records. Consenting patients completed the questionnaires on their own or with assistance from a researcher. Those agreeable to follow-up repeated the measures approximately five days later to examine test-retest reliability. Repeat questionnaires

were completed either face-to-face or via phone follow-up and asked patients to respond in relation to how they had felt in the past day.

Statistical Analyses

A sample size of 150 patients was needed to estimate item difficulty within ± 0.5 logits, with $\alpha = 0.01$ and $\beta = 0.2$.[14]

An exploratory factor analysis (EFA) was undertaken using SPSS Version 22. For the data to be suitable for factor analysis, a Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy exceeding 0.6 was required,[15] as well as a significant alpha value ($p < .05$) for Bartlett's test of sphericity.[16] To extract the factors, a principal components analysis (PCA) was performed, with oblique rotation of the components. Three criteria were used to determine the number of factors to retain: eigenvalues greater than 1, Horn's parallel analysis,[17] and inspection of Cattell's scree plot.[18] Software developed by Watkins (2000) was used to perform parallel analysis.[19]

The RUMM2030 program was used to perform Rasch analyses on the subscales derived from the PCA.[20] Rasch analysis is a form of probabilistic testing that compares a scale against a mathematical measurement model, yielding a detailed assessment of a scale's functioning.[21, 22] Our procedure for Rasch analysis was consistent with guidelines described elsewhere.[10,23,24]

To examine whether response patterns deviated from Rasch model expectations, several fit statistics were calculated.[25] Overall model fit was examined with a non-significant chi-square statistic, using a Bonferroni adjustment ($p = .05/n$ items). Overall person fit and item fit were assessed by examination of the summary fit residual standard deviation, with a value < 1.5 considered acceptable.[10] Chi-square statistics (with a Bonferroni adjustment) and individual fit residual values were used to assess individual item fit and individual person fit. Poor fit was suggested by

significant chi-squared statistics, or residual values exceeding 2.5. Values below 2.5 suggested item redundancy.[25]

An analysis of model misfit was also undertaken by testing for disordered thresholds, differential item functioning, and multidimensionality. We tested the suitability of the 5-point response option format by identifying disordered thresholds for items. Differential item functioning (DIF) was assessed for each item with analysis of variance (with a Bonferroni adjustment) across sex and age. DIF is a form of item bias that occurs when groups (e.g., sex, age) within the sample respond differently to an individual item despite equal levels of the underlying construct.[10,26]

Dimensionality was assessed by examining a PCA of the residual correlation matrix.[27] Subsets of items with high positive or high negative loadings were identified from the PCA's first unrotated factor and these subsets were compared for significant differences using a series of independent *t*-tests.[27] If the lower bound of the binomial confidence interval exceeded 5% (i.e., more than 5% of the tests were significant), the scale was considered multidimensional.[27] Residual correlations exceeding 0.20 in the PCA of residuals were assessed for local dependency. Assessment of how well targeted the items were for participants was undertaken by examining the person-item distribution graphs.

Finally, with the aim of shortening the scale to reduce respondent burden, items appropriate for removal were identified through inspection of item maps. An item map provides information about the relative difficulty of each item.[10] Items listed at the same location on the map are of similar difficulty and potential candidates for removal, when considered in conjunction with other parameters provided by the Rasch analysis and face validity.

The Rasch-derived person separation index (PSI) and the CTT-derived Cronbach's alpha statistic were used to assess internal consistency. These two

statistics are similarly interpreted, with a value above 0.70 considered acceptable.[28] Test-retest reliability was calculated with the CTT intra-class correlation (ICC) coefficient in SPSS employing the two-way random-effects design with relative agreement.[29] An ICC > 0.75 was considered "excellent" and between .40-.75 was deemed "fair-to-good".[30] Patients with symptoms that did not change from baseline to follow up (indicated by a MSAS change score less than +/- ½ SD)[31] were included in the test-retest reliability analysis.

Results

Sample Characteristics

Of the 296 patients approached, 228 patients provided informed consent to participate (77% response rate). As shown in Figure 1, 15 patients were excluded because of incomplete questionnaires, one due to ineligibility (curative disease) and one was an extreme outlier, aged 26 years. Of the 211 patients suitable for analysis, 22 patients had a progressive disease other than cancer. The sample characteristics are summarized in Table 1.

Scale Evaluation

Exploratory Factor Analysis

Bartlett's test of sphericity was significant ($p < .001$) and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.88, indicating suitability for factor analysis. PCA offered five components with eigenvalues greater than 1, but parallel analysis and inspection of Cattell's scree plot suggested a two-factor solution as optimal, explaining a total variance of 46.8%. Oblimin rotation of the two factors was interpretable, but due to low loadings, we deleted one item from each component: Item 6 *'I am not in good spirits'* and item 17 *'I am ashamed of what little I have accomplished'*. The correlation between the components was $r = 0.49$. Component 1 was labelled *'Meaning and Purpose'* to represent the face validity of its items and

Component 2 was labelled *'Distress and Coping Ability',* similarly to represent its items. See Table 2 for the component loadings.

Rasch Analysis

The two components identified using PCA were subjected to Rasch analysis separately using the partial credit model in RUMM2030. Disordered thresholds were detected for all items on both subscales, suggesting an inappropriate response format. This was consistent with our observations during the scale administration that patients showed inconsistent use of the options "seldom" versus "sometimes", and "often" versus "all the time". This was resolved by collapsing response categories[32] to create three response options representing 'never', 'sometimes', and 'often'. This corrected the disordered thresholds on every item for both subscales.

Meaning and Purpose

Rasch analysis of the 11-item *Meaning and Purpose* subscale showed good overall fit (p = .01, fit resid SD = 1.2), with no misfitting items or persons. No DIF was detected for sex or age. Unidimensionality of the subscale was achieved (see Table 3; analysis 1), however, local dependency was found between item 14 *'Life is no longer worth living'* and item 20 *'I would rather not be alive'* (residual correlation = 0.36).

Distress and Coping Ability

The 11-items of the *Distress and Coping Ability* subscale showed good fit to the Rasch model (p = 0.31, fit resid SD = 0.99) with no misfitting items and one misfitting person. Some degree of uniform DIF was found by sex on item 23 *'I feel quite isolated or alone'*, with a significant Bonferonni adjusted alpha value (p = 0.002). Specifically, at equal levels of demoralization, female respondents endorsed a higher level on item 23 than males. This minor level of DIF was consistent with clinical observations and was therefore treated conservatively without further action.

Unidimensionality of the subscale was supported (see Table 3; analysis 3) and there was no evidence of local dependency.

Scale Modification

Following inspection of the item maps for co-location on each subscale, items 3, 14, and 22 were removed from the Meaning and Purpose subscale, while items 10, 16, and 21 were eliminated from the Distress and Coping Ability subscale.

Meaning and Purpose

The revised 8-item Meaning and Purpose subscale demonstrated improved fit to the model ($p = 0.01$, fit resid $SD = 1.02$) with no misfitting items and no misfitting persons. With the removal of item 14, local independence was observed. Non-uniform DIF by sex just reached statistical significance on item 20 *'I would rather not be alive'*. Given this item is of high clinical importance, a conservative approach was taken, with item 20 retained. Unidimensionality of the scale was supported (see Table 3; analysis 2). Overall, the subscale was appropriately targeted, as there were an adequate number of items of varying difficulty to capture the distribution of respondents, as shown in Figure 2.

Distress and Coping Ability

Rasch analysis of the revised 8-item Distress and Coping Ability subscale, showed good overall fit ($p = 0.20$, fit resid $SD = 0.95$), with no misfitting items or persons. Uniform DIF by sex was present for item 23 *'I feel quite isolated or alone'*, ($p = 0.002$). Unidimensionality of the scale was supported and no local dependency was detected (see Table 3, analysis 4). Overall, the subscale was appropriately targeted, as there were sufficient items of varying difficulty to capture the spread of scores from respondents. Refer to Figure 3.

Total Scale

To test the appropriateness of summing subscale scores to provide a total score representing the underlying construct of demoralization, a subtest analysis was conducted comparing person estimates for the Meaning and Purpose subscale and the Distress and Coping Ability subscale (see Table 3; analysis 5). This test supported the unidimensionality of an underlying construct of demoralization.

Reliability

The internal consistency of the subscales and total 16-item scale are reported in Table 3. With reference to the MSAS scores at baseline (M = .89, SD = .44, range: 0-4), a subsample of patients with stable symptoms (n = 58, M interval days = 4.71, SD = 2.04) demonstrated test-retest reliability of the log of the DS-II total score (ICC = 0.80, 95% CI = 0.66-0.88) and subscales (log of Meaning and Purpose: ICC = 0.68, 95% CI = .45-.81; log of Distress and Coping Ability: ICC = 0.82, 95% CI = .69-.89). Due to the presence of outliers and some indication of non-normality and heteroscedasticity log transformations were performed prior to the calculation of the ICCs.[33]

Descriptive statistics

For ease of clinical utility, ordinal scores were used to report descriptive statistics. The ordinal scores in SPSS were converted to a 3-point scale and items were deleted to form the 16-item scale as determined by the Rasch analyses. Summary scores for Meaning and Purpose and Distress and Coping Ability were calculated by summing the 8 items in each subscale. A total score for demoralization was calculated by summing all 16 items. Both subscales were skewed with a prominent floor effect, indicating low levels of demoralization for many in the sample tested. Refer to Table 4 for a summary of the descriptive statistics.

The Spearman correlation coefficient between Meaning and Purpose and Distress and Coping Ability was $\rho = .61$, $p < 0.001$.

Discussion

We have refined the DS to create a 16-item, 2-component scale with sound psychometric properties. Given the reduced respondent burden, the DS-II should be more user-friendly in advanced disease settings. The *Meaning and Purpose* subscale combines items from the 'Loss of meaning and purpose' and the 'Helplessness' subscales in the original DS into a single factor.[5] This subscale will be a useful response measure in meaning centred therapies. Similarly, the *Distress and Coping Ability* subscale combines items that previously formed the 'Dysphoria', 'Disheartenment', and 'Sense of failure' subscales of the original DS.[5] This subscale will likely be a good indicator of response to cognitive and supportive therapies. The shared variance between the two new subscales was 36%, indicating that although related, the two components measure different aspects of demoralization.

Disordered thresholds indicated respondents were unable to reliably differentiate the five original response options,[10] yet did so satisfactorily when three options (never, sometimes or often) were used. Item reduction was possible as both components satisfied Rasch criteria for model fit, with only minor deviations. This allowed for clinical judgment to be considered in conjunction with psychometric findings. A further three items from each subscale were removed, retaining model fit and resulting in an instrument consisting of 16 items, with two 8-item subscales.

With regard to reliability, the DS-II demonstrated excellent internal consistency when measured with Cronbach's alpha. The IRT derived PSI was lower as it was affected by the skew in the data,[10] given a number of subjects reported zero scores. The magnitude of Cronbach's alpha was unaffected in this manner and is therefore the more relevant measure to cite. The scale demonstrated test-retest reliability among a subset of patients (n = 58) with stable symptoms over time. When measuring test-retest reliability in a palliative care population, symptom stability is

more important than precision of time interval.[34] Although the demonstration of test-retest reliability is an important contribution to the validation of the DS-II, these findings should be replicated in larger samples in the future.

The combination of CTT and Rasch modeling has enabled a comprehensive assessment of the DS-II. We anticipate that the DS-II will be clinically useful as an observational measure of demoralization and a measure of change in trials of meaning-based interventions. Nonetheless, there are limitations to the current findings. Although Rasch analysis is distribution free, it is affected by the spread of respondents across the construct.[10] Our sample was relatively small and skewed, impacting the PSI. However, this likely reflects the distribution of the construct in the targeted population. Given the sample size, we were unable to apply confirmatory factor analysis to the current dataset. For clinical utility, we utilized the ordinal data to report descriptive statistics. In future research, larger studies will enable the development of Rasch conversion tables so that Rasch-derived interval data can be utilized and parametric tests can be employed with confidence. Nonetheless, the difficulty of accessing large numbers of patients in a palliative care cohort is recognized.[35]

Overall, the DS-II is a 16-item, two-component scale that has demonstrated item fit, unidimensionality, and reliability as a measure of demoralization in palliative care patients. Examination of external construct validity with socio-demographic factors and concurrent measures, including quality of life, depression, and attitudes toward end of life is reported in a companion paper (Robinson et al., under review CNCR-15-2644) and completes the revalidation process. The DS-II is likely to be a useful clinical and research tool in meaning-centred therapies and when patient populations are at risk of demoralization (e.g. advanced and serious medical disease,

alcohol and substance dependence, chronic mental illness and low socio-economic groups).

References

1. Kissane DW, Clarke DM, Street AF. Demoralization syndrome: A relevant psychiatric diagnosis for palliative care. J Palliat Care. 2001;17(1):12-21.

2. Mehnert A, Vehling S, Hocker A, Lehmann C, Koch U. Demoralization and depression in patients with advanced cancer: Validation of the German version of the demoralization scale. J Pain Symptom Manage. 2011;42(5):768-776. doi: 10.1016/j.jpainsymman.2011.02.013.

3. Robinson S, Kissane DW, Brooker J, Burney S. A systematic review of the demoralization syndrome in individuals with progressive disease and cancer: A decade of research. J Pain Symptom Manage. 2015;49(3):595-610. doi:10.1016/j.jpainsymman.2014.07.008.

4. Robinson S, Kissane DW, Brooker J, Burney S. A review of the construct of demoralization: History, definitions, and future directions for palliative care. Am J Hosp Palliat Med. In press. doi: 10.1177/1049909114553461.

5. Kissane DW, Wein S, Love A, Lee XQ, Kee PL, Clarke DM. The demoralization scale: A report of its development and preliminary validation. J Palliat Care. 2004;20(4):269-276.

6. Costantini A, Picardi A, Brunetti S, et al. Italian version of demoralization scale: A validation study. Riv Psichiat. 2013;48(3):234-239. doi: 10.1708/1292.14291.

7. Hung H, Chen H, Chang Y, et al. Evaluation of the reliability and validity of the Mandarin version of demoralization scale for cancer patients. J Intern Med Taiwan. 2010;21:427-435.

8. Mullane M, Dooley B, Tiernan E, Bates U. Validation of the demoralization scale in an Irish advanced cancer sample. Palliat Support Care. 2009;7(3):323-330. doi: 10.1017/S1478951509990253.

9. Luo X, Cappelleri JC, Cella D, et al. Using the rasch model to validate and enhance the interpretation of the functional assessment of cancer therapy – kidney symptom index—disease–related symptoms scale. Value in Health. 2009;12(4):580-586. doi: 10.1111/j.1524-4733.2008.00473.x.

10. Pallant JF, Tennant A. An introduction to the rasch measurement model: An example using the hospital anxiety and depression scale (HADS). Br J Clin Psychol. 2007;46(1):1-18. doi: 10.1348/014466506X96931.

11. Holland PW, Wainer H. Differential item functioning. Hilldale, NJ: Lawrence Erlbaum Associates, 2012.

12. Conrad KJ, Wright BD, McKnight P, McFall M, Fontana A, Rosenheck R. Comparing traditional and rasch analyses of the Mississippi PTSD scale: Revealing limitations of reverse-scored items. J Appl Meas. 2004;5(1):15-30. Available from URL: http://www.researchgate.net/profile/Kendon_Conrad [accessed July, 2015].

13. Portenoy RK, Thaler HT, Kornblith AB, et al. The memorial symptom assessment scale: An instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer. 1994;30(9):1326-1336. doi: http://dx.doi.org/10.1016/0959-8049(94)90182-1.

14. Linacre J. Sample size and item calibration [or person measure] stability. Rasch Meas Transact. 1994;7(4). Available from URL: http://www.rasch.org/rmt/rmt74m.htm [accessed February, 2015].

15. Kaiser HF. A second generation little jiffy. Psychometrika. 1970;35(4):401-415. doi: 10.1007/BF02291817.

16. Bartlet M. A note on multiplying factors for various chi-squared approximations. J R Stat Soc Series B Stat Methodol. 1954;16:296-298.

17. Horn JL. A rationale and test for the number of factors in factor analysis. Psychometrika. 1965;30(2):179-185. doi: 10.1007/BF02289447.

18. Cattell RB. The scree test for the number of factors. Multivariate Behav Res. 1966;1(2):245-276. doi: 10.1207/s15327906mbr0102_10.

19. Watkins MW. Monte carlo PCA for parallel analysis [computer software]. State College, PA: Ed & Psych Associates. 2000.

20. Andrich D, Sheridan B, Luo G. Rasch models for measurement: RUMM2030 [computer software]. Perth, Western Australia: RUMM Laboratory Pty Ltd. 2010.

21. Pallant JF, Haines HM, Hildingsson I, Cross M, Rubertsson C. Psychometric evaluation and refinement of the prenatal attachment inventory. J Reprod Infant Psych. 2014;32(2):112-125. doi: 10.1080/02646838.2013.871627.

22. Perline R, Wright BD, Wainer H. The rasch model as additive conjoint measurement. Appl Psychol Meas. 1979;3(2):237-255. doi: 10.1177/014662167900300213.

23. Tennant A, Conaghan PG. The rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a rasch paper? Arthritis Care Res. 2007;57(8):1358-1362. doi: 10.1002/art.23108.

24. Hagquist C, Bruce M, Gustavsson JP. Using the rasch model in nursing research: An introduction and illustrative example. Int J Nurs Stud. 2009;46(3):380-393. Doi: http://dx.doi.org/10.1016/j.ijnurstu.2008.10.007.

25. Shea TL, Tennant A, Pallant JF. Rasch model analysis of the depression, anxiety and stress scales (DASS). BMC Psychiatry. 2009;9(1):21. doi: 10.1186/1471-244X-9-21.

26. Lange R, Thalbourne MA, Houran J, Lester D. Depressive response sets due to gender and culture-based differential item functioning. Personality Individual Differences. 2002;33(6):937-954. doi: 10.1016/S0191-8869(01)00203-3.

27. Smith Jr EV. Understanding rasch measurement: Detecting and evaluating the impact of multidimenstionality using item fit statistics and principal component analysis of residuals. J Appl Meas. 2002;3(2):205-231.

28. Santos JRA. Cronbach's alpha: A tool for assessing the reliability of scales. J Ext. 1999;37(2):1-5. Available from URL: http://www.joe.org/joe/1999april/tt3.php?ref [accessed July, 2015].

29. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005;19(1):231-240. Available from URL: http://www.ncbi.nlm.nih.gov/pubmed/15705040 [accessed February, 2015].

30. Fleiss JL. Design and analysis of clinical experiments. Canada: John Wiley & Sons, 1999.

31. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407-415. doi: 10.1016/0197-2456(89)90005-6.

32. Andrich D, De Jong J, Sheridan BE. Diagnostic opportunities with the rasch model for ordered response categories. In: Rost J, et al, editors. Michigan: University of Michigan Library, 1997:59-70.

33. Tabachnick BG, Fidell LS. Using multivariate statistics. 5th ed. Boston, USA: Pearson Education, 2007.

34. Paiva CE, Barroso EM, Carneseca EC, et al. A critical analysis of test-retest reliability in instrument validation studies of cancer patients under palliative

care: A systematic review. BMC Med Res Methodol. 2014;14(1):8. doi: 10.1186/1471-2288-14-8.

35. McWhinney IR, Bass MJ, Donner A. Evaluation of a palliative care service: Problems and pitfalls. BMJ. 1994;309(6965):1340-1342. doi: http://dx.doi.org/10.1136/bmj.309.6965.1340.

Figure Legends

Figure 1: Patient flow

Figure 2: Person-item threshold distribution graph for the Meaning and Purpose subscale

Figure 3: Person-item threshold distribution graph for the Distress and Coping Ability subscale

Tables

Table 1

Sample characteristics

	Sample	
Variable	***n***	**%**
Total Sample	211	100
Gender		
Male	109	51.7
Female	102	48.3
Age, M(SD)	70.98 (12.00)	
40-59 years	44	21.0
60-79 years	108	51.4
80+ years	58	27.6
Marital Status		
Single	24	11.4
Married/De Facto	113	53.5
Divorced/Separated	36	17.1
Widowed	38	18.0
Religion		
Christianity	116	55.2
Other religion	19	9.1
No religion	75	35.7
Education (highest level achieved)		
Incomplete secondary education	49	23.4
Secondary education	47	22.5
Trade or college training	51	24.4
Tertiary education	62	29.7
Employment Status		
Employed	18	8.6
Retired	144	68.9
Disability pension	47	22.5
Recruitment Site		
Cabrini Health	90	42.6
Calvary Health Care Bethlehem	77	36.5
Monash Health	44	20.9
Type of Patient		
Inpatient	182	86.3
Outpatient	29	13.7
Primary Diagnosis		
Cancer	189	89.6
Cardiorespiratory disease	12	5.7
Neurological disease	9	4.2
Renal failure	1	0.5
Months of illness, M(SD)	34.17 (45.47)	
Karnofsky index, M(SD)	56 (12)	

Note. Missing data present in some categories.

Table 2

Pattern matrix of the two-component PCA solution with oblimin rotation of the DS-II

	Component	
	1	2
Item 02: *My life seems to be pointless*	**.83**	-.05
Item 14: *Life is no longer worth living*	**.80**	-.01
Item 03: *There is no purpose to the activities in my life*	**.80**	-.09
Item 20: *I would rather not be alive*	**.74**	-.14
Item 01: *There is little value in what I can offer others*	**.63**	-.06
Item 04: *My role in life has been lost*	**.61**	.13
Item 08: *I feel that I cannot help myself*	**.55**	.16
Item 07: *No one can help me*	**.55**	.20
Item 22: *I feel discouraged about life*	**.51**	.31
Item 09: *I feel hopeless*	**.50**	.34
Item 19: *I am not a worthwhile person*	**.47**	.15
Item 06: *I am not in good spirits*	.40	.39
Item 11: *I feel irritable*	-.25	**.77**
Item 15: *I tend to feel hurt easily*	-.12	**.75**
Item 16: *I am angry about a lot of things*	-.06	**.75**
Item 18: *I feel distressed about what is happening to me*	.08	**.68**
Item 24: *I feel trapped by what is happening to me*	.20	**.57**
Item 21: *I feel sad and miserable*	.26	**.57**
Item 10: *I feel guilty*	.00	**.53**
Item 12: *I do not cope well with life*	.27	**.51**
Item 13: *I have a lot of regret about my life*	.14	**.48**
Item 05: *I no longer feel emotionally in control*	.24	**.48**
Item 23: *I feel quite isolated or alone*	.24	**.46**
Item 17: *I am ashamed of what little I have accomplished*	.24	.33

Note. Retained values in bold. Refer to Appendix for final version of the DS-II.

Table 3

Summary of results of Rasch analyses of DS-II items

Scale	Analysis	Overall model fit	Item fit residual mean (SD)	Person fit residual mean (SD)	% sig *t*-tests	Internal consistency PSI	Internal consistency α
Meaning and Purpose							
11-items	1	$X^2 = 39.75$, $p = .01$	.13 (1.20)	.27 (1.07)	3.32%	.72	.89
8-items (3, 14, and 22 removed)	2	$X^2 = 31.76$, $p = .01$	.06 (1.02)	.23 (0.86)	0.95%	.64	.84
Distress and Coping Ability							
11-items	3	$X^2 = 24.78$, $p = .31$	.05 (0.99)	.28 (1.24)	2.84%	.73	.87
8-items (10, 16, and 21 removed)	4	$X^2 = 20.48$, $p = .20$	.04 (0.95)	.27 (1.07)	1.90%	.65	.82
Total							
16-items	5	$X^2 = 11.55$, $p = .02$	-0.10 (.13)	-0.39 (.71)	0.97%	.79	.89

Note. Cronbach's alpha statistic for total scale calculated in SPSS

Table 4

Descriptive statistics for the scales

	Meaning and Purpose	Distress and Coping Ability	Total
Mean (SD)	3.75 (3.67)	3.89 (3.45)	7.64 (6.43)
Median (IQ range)	3 (1, 6)	3 (1, 6)	6 (3, 11)
Observed range	0-15	0-16	0-31
Possible range	0-16	0-16	0-32
Skewness	1.02	1.06	1.03
Kurtosis	0.32	0.77	0.70

Full Title: Refinement and Revalidation of the Demoralization Scale: DS-II—

External Validity

Running title: DS-II: External Validity

Authors:
Sophie Robinson[1,2] (BPsych(Hons)), David. W. Kissane[1,3,4,6] (MBBS, MPM, MD, FRANZCP, FAChPM, FAPM), Joanne Brooker[1,3] (BSc(Hons), GradDipPsych, GradDipComp, PhD), Courtney Hempton[1,3] (BBSc(Hons), MBioethics), Natasha Michael[4,7](MBChB, FRACP, FAChPM, MRCPI, MRCGP, MSc), Jane Fischer[5] (MBBS, DipChildHealth, Grad.Dip.Med.Pall.Care, FAChPM), Michael Franco[6,7](MBBS, FRACP, FAChPM), Merlina Sulistio[4] (MBBS, BMedSc, FRACP, FAChPM), David M. Clarke[1] (MBBS, MPM, PhD, FRACGP, FRANZCP), Mehmet Ozmen[1,8] (BComm, BEc(Hons), PhD), and Susan Burney[2,3] (BA., Grad.Dip.Couns.Psych., MHlthSc., PhD).

[1] Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia.

[2] School of Psychological Sciences, Monash University, Clayton, Australia.

[3] Szalmuk Family Psycho-oncology Unit, Cabrini Health, Malvern, Australia.

[4]Cabrini Palliative Care Service, Prahran, Australia

[5]Calvary Health Care Bethlehem, Caulfield, Australia

[6]Supportive and Palliative Care Unit, Monash Health, Clayton, Australia

[7]Department of Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia

[8]Department of Econometrics, Monash University, Clayton, Australia

Address correspondence to:
David W. Kissane, MD, MPM, FRANZCP, FAChPM, FAPM
Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Level 3, P Block, Monash Medical Centre, 246 Clayton Road, Clayton, Victoria 3186, Australia.
Phone: +61 3 9594 1477;
Email: david.kissane@monash.edu

Funding sources/grants: Bethlehem Griffiths Research Foundation
Conflicts of interest: None

Author Contributions: Sophie Robinson: Conceptualization, methodology,

validation, formal analysis, investigation, data curation, writing – original draft,

writing – review and editing, and project administration. **David. W. Kissane:** Conceptualization, methodology, validation, formal analysis, investigation, resources, writing – original draft, writing – review and editing, visualization, supervision, project administration, and funding acquisition. **Joanne Brooker:** Conceptualization, methodology, validation, formal analysis, writing – review and editing, and supervision. **Courtney Hempton:** Investigation, data curation, writing – review and editing, and project administration. **Natasha Michael:** Conceptualization, methodology, investigation, writing – review and editing, visualization, supervision, and project administration. **Jane Fischer:** Conceptualization, methodology, resources, and writing – review and editing. **Michael Franco:** Conceptualization, methodology, investigation, writing – review and editing, and supervision. **Merlina Sulistio:** Conceptualization, methodology, investigation, writing – review and editing, visualization, supervision, and project administration. **David M. Clarke:** Conceptualization, methodology, validation, formal analysis, writing – review and editing, and supervision. **Mehmet Ozmen:** methodology, software, formal analysis, writing – review and editing, visualization, and supervision. **Sue Burney**: Conceptualization, writing – review and editing, and supervision.

Total number of:
Print pages = 20
Tables = 3
Figures = 0

Precis: The DS-II demonstrated convergent and discriminant validity. A difference of two points on the DS-II may be clinically meaningful.

Keywords: construct validity, revalidation, external validity; convergent validity, discriminant validity, demoralization, cancer.

Abstract:

Background: The recently refined Demoralization Scale-II (DS-II) is a 16-item, self-report measure of demoralization. Its two factors – Meaning and Purpose, and Distress and Coping Ability, demonstrate sound internal validity, including item fit, unidimensionality, internal consistency, and test-retest reliability. The convergent and discriminant validity of the DS-II with various measures is reported here.

Methods: Palliative care patients with cancer or other progressive diseases (n = 211) completed a battery of questionnaires, including the DS-II and measures of symptom burden, quality of life, depression, and attitudes toward end-of-life. Spearman's rho correlations were performed to assess convergent validity. Mann-Whitney U Tests with calculated effect sizes were used to examine discriminant validity and establish the minimal clinically important difference (MCID). Cross-tabulation frequencies with a chi square analysis were employed to examine discriminant validity with major depression.

Results: The DS-II demonstrated convergent validity with measures of psychological distress, quality of life, and attitudes toward end-of-life. The DS-II demonstrated discriminant validity, as firstly the DS-II differentiated patients with different functional performance levels and high/low symptoms, with a difference of two points between groups on the DS-II considered clinically meaningful. Further, discriminant validity was demonstrated, as co-morbidity with depression was not found at moderate levels of demoralization.

Conclusion: The DS-II has sound psychometric properties and is an appropriate measure of demoralization. Given its structural simplicity and brevity, it is likely to be a useful tool in meaning-centred therapies.

The Demoralization Scale-II (DS-II) is a recently refined and revalidated 16-item, self-report measure of demoralization.[1] Demoralization is a maladaptive coping response conceptualized as a loss of meaning and purpose, with feelings of hopelessness and helplessness.[2] It is understood to arise in response to a stressful event or situation, such as the suffering associated with the diagnosis or experience of an advanced cancer.[2] In our recent systematic[3] and conceptual[4] reviews, we provided a discussion on the differences between demoralization and depression and highlighted that there is a level of overlap between these constructs. We recently reported the internal validity of the DS-II as a two factor model (comprising two 8-item factors: Meaning and Purpose, and Distress and Coping Ability) that demonstrated psychometrically sound item fit, unidimensionality, and reliability in palliative care patients.[1] The reduced number of items and the simplified response format makes the DS-II more user-friendly in the advanced cancer setting than the original 24-item Demoralization Scale (DS).[1]

The original DS demonstrated moderate to strong convergent validity with measures of quality of life, anxiety, depression, hopelessness, hopefulness, adjustment to cancer, and attitudes toward death in a range of cultural contexts.[5-9] The DS also demonstrated discriminant validity, with between 5-23% of patients reporting high demoralization without major depression.[5,6,8,9] However, later research with chi square analyses brought into question the statistical support for discriminant validity of the DS in relation to depression, as demoralized patients were significantly more likely to be depressed than those who were not classified as demoralized.[7] Given the co-morbidity between high demoralization and major depression, further evaluative studies are necessary.[3]

Prior validation studies of the original DS were further limited because they lacked an operational hypothesis for testing discriminant validity. Technically, discriminant validity is suggested when there is no correlation between measures of constructs,[10] yet some level of overlap is to be expected for depression and demoralization.[4,5] Such overlap is seen with depression and anxiety,[11] yet these constructs are recognized diagnostically as separate disorders.[12] Mehnert et al. (2011)[6] found that, compared to patients with moderate demoralization, patients with high levels of demoralization were more likely to experience depression. Thus, co-morbidity is to be expected at the severe end of demoralization, but divergence is more likely with moderate demoralization. If demoralization is conceptualized as an adjustment disorder with limited coping in response to a stressful predicament, greater divergence from depressive features can be anticipated with low or moderate levels of demoralization.[13]

To extend examination of the discriminant validity of the DS-II, its relationship to symptom and performance level is worthwhile. Previous research has found that physical symptoms and demoralization are positively correlated, while activity level and demoralization are negatively correlated.[3] Following methodology employed by Cella and colleagues,[14] we examined differences in the levels of demoralization between high/low functioning and high/low symptomatic patients. Furthermore, this examination can allow for the minimal clinically important difference (MCID) of the scale to be calculated, providing useful information for clinicians or researchers wanting to use the tool as a measure of change.[14]

In this paper the construct (convergent and discriminant) validity of the DS-II is reported. To provide evidence for convergent validity, we expected comparable findings to those reported with the original DS. To provide evidence for discriminant

validity, discrimination between high and low level functioning/symptomatic patients was anticipated, as well as determination of the number of DS-II points needed to demonstrate a MCID. To further examine discriminant validity, we expected co-morbidity with major depression for high levels of demoralization, but divergence at moderate levels.

Materials and Methods

Design and Patients

This observational study was conducted across three sites in Melbourne, Australia: Monash Health, Cabrini Health, and Calvary Health Care Bethlehem, all acute metropolitan hospitals. Approval was received from human research ethics committees at all participating institutions. Patients were recruited from June 2013 to November 2014. Patients' demographic and medical data were obtained from medical records. Patients were eligible if they had advanced progressive disease and excluded if they were too frail/unwell medically to consent; unable to speak English; and/or had cognitive impairment. Eligibility was determined by the patient's treating physician.

Measures

Socio-demographic and medical details. These details included: primary diagnosis, duration of illness, inpatient or outpatient status, treatment type, supportive care status [i.e., receiving counselling], age, sex, marital status, religion, educational achievement, and employment status.

Demoralization Scale-II (DS-II). The DS-II is comprised of 16 items rated on a 3-point Likert scale, including: 0 (*never*); 1 (*sometimes*); and 2 (*often*), with higher scores indicative of higher levels of demoralization (score range: 0-32). It contains two 8-item factors: Meaning and Purpose, and Distress and Coping Ability.

The DS-II has demonstrated good internal reliability ($\alpha = 0.89$) for all patients and test-retest reliability in symptomatically stable patients (ICC = 0.80).

McGill Quality of Life Questionnaire (MQOL). In this study, nine items from the MQOL were used, including six from the existential domain, two from the social support domain, and the single global quality of life item.[15-17] The overall scale has demonstrated good internal reliability ($\alpha = 0.83$) and convergent validity with other measures of quality of life in a palliative care setting.[18]

Patient Health Questionnaire (PHQ-9). The PHQ-9 is a self-report measure designed to assess the presence and severity of a major depressive episode (MDE), and is comprised of nine items representing criteria for a MDE.[19-21] Items are rated on a 4-point Likert scale ranging from 0 (*not at* all) to 3 (*nearly every day*) in relation to whether the symptom has been experienced in the past two weeks. Of the nine criteria, a MDE is indicated if five or more criteria are scored a minimum of 2 on the scale (symptoms have been present at least *more than half the days*) during the past two weeks. At least one of the endorsed criteria must be depressed mood or anhedonia. The criterion "Thought that you would be better off dead or hurting yourself in some way" is included as an endorsed item, regardless of duration (score >0).[20] The PHQ-9 has demonstrated good internal reliability ($\alpha = .89$) and construct validity with other health-related measures.[20]

Schedule of Attitudes towards Hastened Death (SAHD). The 20-item SAHD is a self-report measure designed to capture the desire for death in seriously ill patients.[22,23] A dichotomous answer format of *true* or *false* is employed for each item. A strong desire to die corresponds with endorsement of 10 or more items. The SAHD has demonstrated good reliability ($\alpha = .88$-$.89$) and convergent validity with other

measures of desire for death, depression, and hopelessness in a study with people living with HIV and AIDS and a study with terminally ill patients.[22,23]

Will to Live Rating (WTL). Respondents rated the intensity of their current 'will to live' on a scale of 0-10 (0 = no will to live; 10 = strong will to live). The WTL rating was validated in a study by Chochinov et al (1999)[24] in a cohort of palliative care cancer patients. Recordings of the strength of the will to live were examined on a daily basis to show its variability as symptom levels and wellbeing fluctuated.[24]

Memorial Symptom Assessment Scale (MSAS). The MSAS was employed to measure symptom burden of 32 symptoms (physical and psychological).[25] It has demonstrated good reliability ($\alpha = .82$) in cancer patients and has been extensively validated in palliative care.[25] The MSAS provides a total score and three subscale scores, including the Global Distress Index (GDI), physical symptomatology, and psychological distress.

Statistical analyses

Given the inevitable skew present in the DS-II data,[1] non-parametric tests were conducted using SPSS Version 22. Bonferroni adjustments were made for multiple comparisons, otherwise a significance value of $p < .05$ was set.

Firstly, Spearman's rho coefficient correlations between the DS-II and socio-demographic and treatment-related factors were examined. Mann-Whitney U Tests and Kruskal-Wallis Tests were employed to examine group differences in DS-II scores across socio-demographic and treatment-related factors. Next, convergent validity was assessed by examining the patterns of Spearman's rho coefficient correlations between the DS-II and the MSAS, MQOL, PHQ-9, SAHD, and WTL.

To test discriminant validity, Mann-Whitney U Tests were used to determine whether the DS-II differentiated patients with high/low functional performance levels,

global distress, physical symptoms, psychological symptoms, and total symptoms.[14] The MCID of the DS-II was calculated using these symptom measures as clinical anchors and the effect size (ES) of the Mann-Whitney U Test. This was defined as *Z* divided by the square root of sample size (*N*) for each anchor,[26] where an ES of .1 = small, .3 = moderate, and .5 = large.[27] These non-parametric calculation methods were guided by the parametric alternatives used by Cella et al. (2002).[14]

Discriminant validity between demoralization and depression was also examined, by first determining the DS-II cut-off scores with reference to an extreme groups design:[28] low scorers (0-25th percentile), middle scorers (25th-75th percentile), and high scorers (75th percentile +). The closest approximations to these percentile categories allowed by the data were then compared with PHQ-9 categories using cross-tabulation frequencies and a chi square analysis. To aid interpretation standardized residuals were calculated for each category, with a 90% confidence interval set.

Results

Sample characteristics

In the current study, 296 patients were approached and 228 patients provided informed consent to participate (77% response rate). In the consenting group, 15 patients were excluded because of incomplete questionnaires, one due to ineligibility (curative disease), and one because she was an extreme outlier based on age (26-years-old). Of the 211 patients analysed, 51.7% were male and the mean age was 70.98 (12.00 *SD*) years. The sample characteristics are summarized in Table 1.

Associations with socio-demographics and medical characteristics

Age was negatively, albeit weakly, correlated with Total DS-II (ρ = -.14, p = .04, n = 210), unrelated to Meaning and Purpose (ρ = .01, p = .91, n = 210), and

negatively correlated with Distress and Coping Ability (ρ = -.24, p = .001, n = 210). There were no significant differences in DS-II scale scores for sex, marital status, and religion. The MQOL social support subscale, however, was inversely related to both DS-II subscales (Meaning and Purpose; ρ = -.35, p < .001, n = 178; Distress and Coping Ability: ρ = -.29, p < .001, n = 178) and Total DS-II (ρ = -.37, p < .001, n = 178).

Distress and Coping Ability scores were significantly higher in patients with a tertiary education (*Md* = 4.5, n = 62) than those without (*Md* = 3, n = 147; U = 3677, z = -2.22, p = .03) and for patients on a pension (*Md* = 5, n = 47) compared to retired patients (*Md* = 3, n = 144; U = 2426.5, z = -2.93, p = .003 [Bonferroni adjustment .05/4]). There were no significant relationships between DS-II scales and primary diagnosis, cancer tumor type, or supportive care status. No relationship was found between duration of illness and DS-II scales. Total DS-II scores were higher for patients who were receiving or had received radiation therapy (*Md* = 7, n = 104) than patients who had not (*Md* = 6, n = 97; U = 4227, z = -1.99, p = .047). Patients currently taking an anxiolytic also had higher scores on Total DS-II (*Md* = 7; n = 94) than patients who were not (*Md* = 5, n = 105; U = 4054, z = -2.18, p = .029).

Convergent validity

Descriptive statistics for the MSAS, MQOL, PHQ-9, SAHD and WTL, along with their correlation with the DS-II scales are reported in Table 2.

For the Total DS-II scale, there were moderate-to-strong positive relationships with psychological symptom burden, depression, and desire to die. In addition, the results revealed that there were moderate-to-strong negative relationships between Total DS-II scores and quality of life, social support, existential wellbeing, and will to live.

As shown in Table 2, similar patterns were found at a DS-II subscale level. Psychological symptom burden had a higher correlation with Distress and Coping Ability than with Meaning and Purpose, while Meaning and Purpose had a stronger relationship with the MQOL subscales, desire to die, and will to live.

Discriminant validity

Mann-Whitney U Test results indicated that patients reporting higher global distress (>1.0) had significantly higher demoralization scores (*Md* = 8.5; *n* = 110) than patients with lower global distress (*Md* = 3, *n* = 82), with a median difference of 5.5 points (ES = .49; *U* = 1946, *z* = -6.75, *p* < .001). Patients with higher MSAS physical symptoms (>1.0) had significantly higher demoralization scores (*Md* = 7; *n* = 115) than patients with lower physical symptoms (*Md* = 5, *n* = 76; median difference = 2 points; ES = .19; *U* = 3391, *z* = -2.63, *p* = .009). Regarding MSAS psychological distress, patients with higher scores (>1.0) had significantly higher levels of demoralization (*Md* = 11; *n* = 72) than patients with lower scores (*Md* = 4; *n* = 120), with a median difference of seven points (ES = .42; *U* = 1370.5, *z* = -7.93, *p* < .001). Patients with higher levels on MSAS total symptoms (>1.0) had significantly higher levels of demoralization (*Md* = 10; *n* = 73) than patients with lower levels (*Md* = 4; *n* = 118; median difference = 6 points; ES = .39; *U* = 2301.5, *z* = -5.415, *p* < .001). Higher functioning patients (Karnofsky rating > 70, 'able to carry on normal activity')[29] reported lower demoralization scores (*Md* = 3; *n* = 27) than patients with lower performance ratings (*Md* = 7; *n* = 163), with a median difference of four points (ES = .18; *U* = 1538.5, *z* = -2.51, *p* = .012). These observed median differences and subsequent effect sizes suggest that a difference of at least two points on the DS-II between groups may be clinically meaningful (MCID).

To test for discriminant validity between major depression and demoralization, the DS-II data were divided into three categories: low scorers 0-3 (65 patients, 30.8%), middle scorers 4-10 (85 patients, 40.3%), and high scorers 11+ (61 patients, 28.9%). The PHQ-9 was completed by 183 patients, with 21 (11.5%) meeting diagnostic criteria for a MDE. Table 3 shows the cross-tabulation frequencies and associated standardized residuals (>+/-1.64 significant, 90% confidence interval) between demoralization and the presence of major depression. The chi-square analysis found a significant relationship between PHQ-9 categories and DS-II categories, X^2 (2, n = 183) = 32.41, p < .001, indicating demoralized patients were significantly more likely to be depressed than those who were not demoralized. The standardized residuals indicated that individuals with low and moderate demoralization were more likely than chance to not be experiencing a MDE, whereas individuals with high demoralization were more likely than chance to be experiencing a MDE.

Discussion

The DS-II is a refined measure of demoralization, consisting of 16 items and two subscales: Meaning and Purpose, and Distress and Coping Ability. In this paper, the DS-II demonstrated convergent validity with measures of psychological symptom burden, quality of life, existential wellbeing, depression, and attitudes toward end-of-life (including desire to die and will to live).

Compared to the Distress and Coping Ability scale, Meaning and Purpose yielded a stronger relationship with desire to die and will to live. This observed difference suggests that loss of meaning and purpose has a more profound impact on attitudes toward end-of-life, and perhaps the development of suicidal ideation,[2] than a breakdown in coping and general distress. This is consistent with research that found

that meaninglessness and hopelessness were mediators of the relationship between depression and suicidality,[3,24,30] and with recent research by Fang and colleagues,[31] who reported demoralization had more influence on suicidal ideation than depression.

Furthermore, Meaning and Purpose had a stronger observed relationship with quality of life indicators than Distress and Coping Ability. Thus it appears that a loss of meaning and purpose has a more profound effect on quality of life than general dysphoria, disheartenment, and sense of incompetence, as measured by the Distress and Coping Ability subscale. In contrast, Distress and Coping Ability yielded a stronger correlation with psychological symptom burden than Meaning and Purpose. This outcome suggests that the Distress and Coping Ability subscale is a better measure of global psychological distress, than the Meaning and Purpose subscale.

Previously, the establishment of the discriminant validity of the DS was hampered by the lack of a clear hypothesis. Mehnert and colleagues (2011)[6] highlighted that divergence between depression and demoralization is evident at moderate levels of the construct. Here we have shown that co-morbidity between depression and demoralization exists at high levels of demoralization but not at moderate levels, thus supporting Mehnert et al.'s findings. From anecdotal evidence, co-morbidity can be well understood when patients with major depression develop suicidality, as prominent demoralization is a key component of this presentation. However, in the moderate demoralization range, many patients do not meet criteria for major depression, yet will have a constellation of symptoms that constitute a form of adjustment disorder. These features may be more usefully conceptualized as Adjustment Disorder with demoralization than Adjustment Disorder with depressive symptoms. We have previously suggested that both the patient and their clinician might better understand such a classification, as it more accurately describes the

patient's experience and could allow for the patient to feel more clearly heard.[3] Thus, one approach would be to consider a separate diagnosis of Demoralization Syndrome,[2] however, the present data does not support this argument. Alternatively, demoralization could be added as a specifier to Adjustment Disorder.[3] Overall, future research is required to continue to clarify demoralization's diagnostic role in mental and physical health.[32]

We recognize the limitation of utilizing interquartile ranges to determine cut-offs for the DS-II. Unfortunately, we had no alternative analytic means of establishing cut-off points at this time. To determine reliable clinical significance levels for demoralization, scores on the DS-II need to be compared to a 'gold standard' measure, typically a validated diagnostic interview. Such an interview has yet to be developed. Nonetheless we chose to examine the divergence of demoralization from depression in an exploratory manner, as we believe the question of co-morbidity is of strong clinical interest. Discriminant validity was strengthened by showing different functional performance levels and high/low symptoms with high versus low DS-II scores. Finally, the estimation of a minimally important difference on the DS-II has provided clinically relevant information.

Examination of the associations between DS-II scores and socio-demographic and treatment-related factors revealed both consistencies and disparities with previous findings for the DS.[3] For example, no differences were found in DS-II scale scores across marital status, although previous research has generally found that being in a relationship or living with other people is associated with lower levels of demoralization.[3] However, one study reported no differences in demoralization by marital status.[33] Nonetheless, we found a small inverse relationship between social support and demoralization, in line with previous findings.[3] Furthermore, the weak

inverse relationship between age and demoralization found in previous studies was replicated here. Being on a disability support pension was associated with increased distress and reduced coping compared with those who were retired. The finding that patients with a tertiary education reported slightly higher distress and poorer coping than patients without a tertiary education adds to the mixed evidence for the association between education and demoralization.[3] Patients who were receiving radiation therapy or had done so previously, reported slightly higher demoralization than patients receiving other treatments, possibly reflecting more serious disease status. The finding that patients currently taking an anxiolytic medication were slightly more demoralized is another possible marker of illness complexity. Finally, we found no relationship between demoralization and primary diagnosis, cancer type, duration of illness, or supportive care status.[7] Caution is warranted in the interpretation of these associations, however, as they are weak.

There are additional study limitations that need to be considered. The homogeneity of the sample was problematic, in that the majority of patients were of older age, retired, Caucasian, and Christian. Therefore, future research is required to replicate the present findings in a heterogenous sample, longitudinally, and cross-culturally. Longitudinal research will assist in understanding the causal nature of the various associations identified between demoralization, socio-demographic and treatment-related factors. A question also remains as to whether the DS-II is suitable for detecting intervention-related changes. Change in DS-II scores following treatment of physical symptoms will help clarify the measure's responsiveness as a measure of state rather than trait.

Overall, in conjunction with the recently reported internal validity of the DS-II,[1] the present findings indicate that the scale is a suitable measure of demoralization

that has demonstrated sound psychometric properties. It is hoped this tool will be utilized in research, timely in an era when trials of meaning-based interventions are emerging. Clinically, it is anticipated that the brevity of the DS-II means it can be utilized readily to assist clinicians when making a clinical judgment about a patient's mental state.

References

1. Robinson S, Kissane DW, Brooker J, et al. Refinement and revalidation of the demoralization scale: The DS-II–internal validity. Cancer. Under joint review.
2. Kissane DW, Clarke DM, Street AF. Demoralization syndrome: A relevant psychiatric diagnosis for palliative care. J Palliat Care. 2001;17(1):12-21.
3. Robinson S, Kissane DW, Brooker J, Burney S. A systematic review of the demoralization syndrome in individuals with progressive disease and cancer: A decade of research. J Pain Symptom Manage. 2015;49(3):595-610. doi:10.1016/j.jpainsymman.2014.07.008.
4. Robinson S, Kissane DW, Brooker J, Burney S. A review of the construct of demoralization: History, definitions, and future directions for palliative care. Am J Hosp Palliat Med. In press. doi: 10.1177/1049909114553461.
5. Kissane DW, Wein S, Love A, Lee XQ, Kee PL, Clarke DM. The demoralization scale: A report of its development and preliminary validation. J Palliat Care. 2004;20(4):269-276.
6. Mehnert A, Vehling S, Hocker A, Lehmann C, Koch U. Demoralization and depression in patients with advanced cancer: Validation of the German version of the demoralization scale. J Pain Symptom Manage. 2011;42(5):768-776. doi: 10.1016/j.jpainsymman.2011.02.013.
7. Mullane M, Dooley B, Tiernan E, Bates U. Validation of the demoralization scale in an Irish advanced cancer sample. Palliat Support Care. 2009;7(3):323-330. doi: 10.1017/S1478951509990253.
8. Costantini A, Picardi A, Brunetti S, et al. Italian version of demoralization scale: A validation study. Riv Psichiat. 2013;48(3):234-239. doi: 10.1708/1292.14291.

9. Hung H, Chen H, Chang Y, et al. Evaluation of the reliability and validity of the Mandarin version of demoralization scale for cancer patients. J Intern Med Taiwan. 2010;21:427-435.

10. DeVellis RF. Scale development: Theory and applications. 3rd ed. USA, North Carolina: Sage Publications, 2012.

11. Zbozinek TD, Rose RD, Wolitzky□Taylor KB, et al. Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. Depress Anxiety. 2012;29(12):1065-1071.

12. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). 5th ed. Washington, DC: American Psychiatric Publishing, 2013.

13. Kissane DW. Demoralization: A life-preserving diagnosis to make for the severely medically ill. J Palliat care. 2014;30(4):255.

14. Cella D, Eton DT, Lai J-S, Peterman AH, Merkel DE. Combining anchor and distribution-based methods to derive minimal clinically important differences on the functional assessment of cancer therapy (FACT) anemia and fatigue scales. J Pain Symptom Manage. 2002;24(6):547-561. doi: 10.1016/S0885-3924(02)00529-8.

15. Cohen SR, Hassan SA, Lapointe BJ, Mount BM. Quality of life in HIV disease as measured by the McGill quality of life questionnaire. Aids. 1996;10(12):1421-1427.

16. Cohen SR, Mount BM, Strobel MG, Bui F. The McGill quality of life questionnaire: A measure of quality of life appropriate for people with advanced disease. A preliminary study of validity and acceptability. Palliat Med. 1995;9(3):207-219. doi: 10.1177/026921639500900306.

17. Cohen SR, Mount BM, Tomas JJ, Mount LF. Existential well□being is an important determinant of quality of life: Evidence from the McGill quality of life questionnaire. Cancer. 1996;77(3):576-586. doi: 10.1002/(SICI)1097-0142(19960201)77:3<576::AID-CNCR22>3.0.CO;2-0.

18. Cohen SR, Mount BM, Bruera E, Provost M, Rowe J, Tong K. Validity of the McGill quality of life questionnaire in the palliative care setting: A multi-centre canadian study demonstrating the importance of the existential domain. Palliat Med. 1997;11(1):3-20. doi: 10.1177/026921639701100102.

19. Kroenke K, Spitzer RL. The PHQ-9: A new depression diagnostic and severity measure. Psychiatr Ann. 2002;32(9):1-7.

20. Kroenke K, Spitzer RL, Williams JB. The PHQ□9. J Gen Intern Med. 2001;16(9):606-613.

21. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD. J Am Med Assoc. 1999;282(18):1737-1744. doi: 10.1001/jama.282.18.1737.

22. Rosenfeld B, Breitbart W, Galietta M, et al. The schedule of attitudes toward hastened death. Cancer. 2000;88(12):2868-2875. doi: 10.1002/1097-0142(20000615)88:12<2868::AID-CNCR30>3.0.CO;2-K.

23. Rosenfeld B, Breitbart W, Stein K, et al. Measuring desire for death among patients with HIV/AIDS: The schedule of attitudes toward hastened death. Am J Psychiatry. 1999;156(1):94-100. doi: http://dx.doi.org/10.1176/ajp.156.1.94.

24. Chochinov HM, Tataryn D, Clinch JJ, Dudgeon D. Will to live in the terminally ill. The Lancet. 1999;354(9181):816-819. doi: 10.1016/S0140-6736(99)80011-7.

25. Portenoy RK, Thaler HT, Kornblith AB, et al. The memorial symptom assessment scale: An instrument for the evaluation of symptom prevalence, characteristics and distress. Eur J Cancer. 1994;30(9):1326-1336. doi: http://dx.doi.org/10.1016/0959-8049(94)90182-1.

26. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. J Exp Psychol Gen. 2012;141(1):2. doi: http://dx.doi.org/10.1037/a0024338.

27. Coolican H. Research methods and statistics in psychology. 6th ed. London and New York: Psychology Press, 2014.

28. Preacher KJ. Extreme groups designs. Encyclopedia of Clinical Psychology. 2015:1-4. doi: 10.1002/9781118625392.wbecp190.

29. Mor V, Laliberte L, Morris JN, Wiemann M. The karnofsky performance status scale: An examination of its reliability and validity in a research setting. Cancer. 1984;53(9):2002-2007. doi: 10.1002/1097-0142(19840501)53:9<2002::AID-CNCR2820530933>3.0.CO;2-W.

30. Beck AT, Kovacs M, Weissman A. Hopelessness and suicidal behavior. J Am Med Assoc. 1975;234(11):1146-1149. doi: 10.1001/jama.1975.03260240050026.

31. Fang CK, Chang MC, Chen PJ, et al. A correlational study of suicidal ideation with psychological distress, depression, and demoralization in patients with cancer. Support Care Cancer. 2014;22(12):3165-3174. doi: 10.1007/s00520-014-2290-4.

32. De Figueiredo JM. Distress, demoralization and psychopathology: Diagnostic boundaries. Eur J Psychiat. 2013;27(1):61-73. doi: http://dx.doi.org/10.4321/S0213-61632013000100008.

33. Lee CY, Fang CK, Yang YC, et al. Demoralization syndrome among cancer outpatients in Taiwan. Support Care Cancer. 2012;20(10):2259-2267. doi: 10.1007/s00520-011-1332-4.

Tables

Table 1

Sample characteristics and medical information

Variable	Sample n	%
Total Sample	211	100
Gender		
Male	109	51.7
Female	102	48.3
Age, M(SD)	70.98 (12.00)	
40-59 years	44	21.0
60-79 years	108	51.4
80+ years	58	27.6
Marital Status		
Single	24	11.4
Married/De Facto	113	53.5
Divorced/Separated	36	17.1
Widowed	38	18.0
Religion		
Christianity	116	55.2
Other religion	19	9.1
No religion	75	35.7
Education		
Incomplete secondary education	49	23.4
Completed secondary education	47	22.5
Trade or college training	51	24.4
Tertiary education	62	29.7
Employment Status		
Employed	18	8.6
Retired	144	68.9
Disability Pension	47	22.5
Type of Patient		
Inpatient	182	86.3
Outpatient	29	13.7
Primary Diagnosis		
Cancer	189	89.6
Breast	25	(13.2)
Prostate	21	(11.1)
Gynecologic	11	(5.8)
Digestive system	48	(25.4)
Lung	32	(17.0)
Other	52	(27.5)
Cardiorespiratory disease	12	5.7
Neurological disease	9	4.2
Renal failure	1	0.5
Duration of illness (months), M(SD)	34.17 (45.47)	
Karnofsky index, M(SD)	56 (12)	
Received/ing treatment		
Palliative chemotherapy	131	62.1
Radiation therapy	104	49.3

Table 1 cont.

Sample characteristics and medical information

Surgery	121	57.3
Medications (currently taking)		
Anxiolytic	94	47.2
Anti-depressant	45	23.1
Opioid	167	81.9
NSAID	102	54.0
Receiving supportive psychosocial care	79	37.4

Note. Missing data present in some categories.

Table 2

Descriptive statistics of Memorial Symptom Assessment Scale (MSAS), McGill Quality of Life Questionnaire (MQOL), Patient Health Questionnaire (PHQ-9), Schedule of Attitudes toward Hastened Death (SAHD) and Will to Live Rating (WTL) and Spearman's correlations with DS-II

Scale	Content	N	Min	Max	Mean	SD	DS-II Meaning & Purpose	DS-II Coping & Personal Sensitivity	DS-II Total
MSAS	Psychological	192	0	3.67	0.95	0.8	.49**	.65**	.64**
MQOL	QoL	180	0	10	7.59	2.47	-.40**	-.34**	-.41**
	Existential Wellbeing	181	0.33	10	7.45	2	-.57**	-.45**	-.57**
PHQ-9	MDE	183	0	1			.37**	.41**	.41**
SAHD	Desire to die	162	0	15	4.02	3.7	.43**	.23*	.39**
WTL	Will to live	120	0	10	8.28	2.29	-.49**	-.25*	-.44**

Note. n values vary due to missing data. QoL = quality of life; MDE = major depressive episode. * = $p < .01$, ** = $p < .001$.

Table 3

Cross-tabulation frequencies between demoralization and major depression

Demoralization		Major Depressive Episode		Total
		No	Yes	
Low (0-3)	% of total	32.8%	.5%	33.3%
	Count	60	1	
	Expected Count	54	7	
	Std. Residual	.8	**-2.3**	
Moderate (4-10)	% of total	36.6%	1.6%	38.3%
	Count	67	3	
	Expected Count	62	8	
	Std. Residual	.6	**-1.8**	
High (11+)	% of total	19.1%	9.3%	11.5%
	Count	35	17	
	Expected Count	46	6	
	Std. Residual	-1.6	**4.5**	

Note. Standardized residuals >+/-1.65 deemed significant (90% confidence interval). Expected cell frequency met (>5).