Lambert, Sylvie D., Pallant, Julie F., Boyes, Allison W., King, Madeleine T., Britton, Benjamin, Girgis, Afaf, “A Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) among cancer survivors". Originally published in Psychological Assessment Vol. 25, Issue 2, p. 379-390 (2013).

Available from: http://dx.doi.org/10.1037/a0031154

A Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) among cancer survivors

Abstract

Objective: The Hospital Anxiety and Depression Scale (HADS) is commonly used to assess distress among individuals with cancer. However, previous studies cast doubt on the most appropriate dimensional structure for the HADS, suggesting that one- or three-dimensional structures might offer superior fit to the original, two-dimensional one. This paper is the first to use Rasch Analysis to examine the psychometric properties of the subscales corresponding to each of these alternative structures.

Methods: The HADS was completed by 1360 cancer survivors. Rasch analyses were conducted to examine summary and individual model fit statistics, person separation index, response format, item bias, redundancy, and dimensionality.

Results: The HADS-Total scale was found to be multidimensional, and it was necessary to remove almost half of the items to achieve fit. Analyses only partially supported the original structure of the HADS-Anxiety and HADS-Depression as both showed initial model misfit and item deletion was necessary to achieve fit. Within the three-dimensional structures, the Rasch statistics for the anxiety subscales were within acceptable range and no adjustment was needed. Analyses did not support adding item 7 to HADS-Depression.

Conclusions: Results supported modified versions of the HADS-Anxiety and HADS-Depression; however, combining all items to form HADS-Total is not recommended. Numerous studies using classical test theory and Rasch analyses have corroborated the exclusion of some items (e.g., item 7) and appropriateness of the subscales defined by a three-dimensional structure. Further research is required to identify the incremental validity of potential revised subscales.

Keywords: anxiety, depression, cancer survivors, Rasch, psychometrics

In the psycho-oncology literature, the terms anxiety and depression are generally used to refer to clinical diagnoses. Psychological distress is less well-defined and is often used as an umbrella term for any multi-factorial, unpleasant emotional experience (National Comprehensive Cancer Network, 2007). The prevalence of clinically significant psychological distress among individuals diagnosed with cancer ranges from 20% to 40% (Carlson et al., 2004; Mitchell, Meader, & Symonds, 2010; National Breast Cancer Centre, 2003; Zabora, BrintzenhofeSzoc, Curbow, Hooker, & Plantadosi, 2001), with 12% to 40% experiencing clinically significant anxiety and up to half experiencing depression (Meyerowitz & Oh, 2009; National Breast Cancer Centre, 2003; Rabin et al., 2009). Moderate to high distress is typically reported in the immediate aftermath of diagnosis and treatment, followed by a gradual reduction over the following year or two (Kayser & Scott, 2008; Meyerowitz & Oh, 2009; Rabin, et al., 2009). However, a significant portion of patients continue to experience psychological distress well into survivorship. Elevated psychological distress has been associated with lower quality of life (Given et al., 2005; Lynch, Steginga, Hawkes, Pakenham, & Dunn, 2008; Zabora, et al., 2001), lower treatment adherence (Rabin, et al., 2009), and higher severity and incidence of treatment side effects (e.g., increased fatigue) (Rabin, et al., 2009). Psychosocial factors, including stress and depression, have also been equivocally linked to poor cancer prevention practices (Honda, Goodwin, & Neugut, 2005), cancer progression (Armaiz-Pena, Lutgendorf, Cole, & Sood, 2009), and mortality (Lewis, Fletcher, Cochrane, & Fann, 2008).

A range of measures is available to assess anxiety and/or depression in individuals diagnosed with cancer, including the State-Trait Anxiety Inventory (STAI), Profile of Mood States (POMS), Centre for Epidemiologic Studies Depression Scale (CES-D), Beck Depression Inventory (BDI), Brief Symptom Inventory (BSI), and Hospital Anxiety and Depression scale (HADS) (Luckett et al., 2010). A recent review of the performance of these, and other measures, revealed that the HADS performed best, ahead of the POMS (original version and unofficial short-form POMS-37) and the CES-D (Luckett, et al., 2010). Although this review recommends the HADS-Anxiety highly, it cautions against relying solely on HADS-Depression when depression is the primary outcome of interest (Luckett, et al., 2010).

The Hospital Anxiety and Depression Scale (HADS)

The HADS was developed in the 1980s to assess anxiety and depression in medical patients (Zigmond & Snaith, 1983). It purposefully excluded many somatic symptoms (e.g., dizziness, sleep disturbance) to avoid confounding psychological symptoms with disease or treatment. Since its development, the HADS has become a ‘benchmark’ measure of anxiety and depression among diverse clinical and non-clinical hospital populations, including individuals diagnosed with cancer (Bjelland, Dahl, Haug, & Neckelmann, 2002; Luckett et al., 2010; Martin, 2005; Smith et al., 2006; (Martin, 2005; Vodermaier & Roanne, 2011). The HADS is a 14-item self-administered questionnaire, with seven items assigned to each the HADS-Anxiety and HADS-Depression subscales. Each item is rated on a four-point response scale (0 to3 – variable response scale). HADS-Anxiety items 1, 3, 5, 11, and 13 and HADS-Depression items 6, 8, and 10 are negatively scored; the remaining items are positively scored. Subscale scores are typically categorized to indicate the level of anxiety or depression experienced where scores of less than 8 are categorized as *normal*, scores of 8 to 10 as *borderline*, and scores of 11 to 21 as *clinical* (Zigmond & Snaith, 1983). A number of psychometric studies have highlighted the scale’s strengths, including its brevity, reliability, and validity and availability of comparison scores across different populations (Bjelland, Dahl, Haug, & Neckelmann, 2002; Luckett, et al., 2010). These studies have also drawn attention to a number of psychometric issues warranting additional analyses, including a) lack of consensus on the optimal cut-point for cases of anxiety (HADS-Anxiety), depression (HADS-Depression), and emotional distress (HADS-Total) (Bjelland, et al., 2002; Luckett, et al., 2010; Singer et al., 2009; Vodermaier & Roanne, 2011; Walker et al., 2007); b) problematic fit of some items, in particular HADS-Anxiety item 7 (Bjelland, et al., 2002; Hunt-Shanks, Blanchard, Reid, Fortier, & Cappelli, 2009); and c) contradictory results regarding optimal dimensional structure where alternative one-dimensional (Luckett, et al., 2010; Smith et al., 2002; Smith et al., 2006;

Walker, et al., 2007) or three- dimensional (Dunbar, Ford, Hunt, & Der, 2000; Friedman, Samuelian, Lancrenon, Even, & Chiarelli, 2001; Gough & Hudson, 2009; Hunt-Shanks, et al., 2009; Martin, Bonner, Brook, & Luscombe, 2006; Smith, et al., 2002) structures are proposed to provide a better fit than the original, two-dimensional structure.

The alternative dimensional structures of the HADS.

HADS-Total.

A simple structure for the HADS was proposed by Razavi, Delvaux, Farvacques, and Roboye (1990) where all 14 items are considered together as a unidimensional measure of emotional distress (referred to as HADS-Total). Despite its popularity among researchers in psycho-oncology (Jacobsen et al., 2005; Rodgers, Martin, Morse, Kendell, & Verrill, 2005; Sellick & Edwardson, 2007; Singer, et al., 2009; Walker, et al., 2007), support for HADS-Total is divided (Barth & Martin, 2005; Dunbar, et al., 2000; Rodgers, et al., 2005). Several studies, using confirmatory factor analysis, have found that HADS-Total often does not meet the Comparative Fit Index (CFI) $\geq$ 0.95 and the Root Mean Square Error of Approximation (RMSEA) < 0.05 criteria (Barth & Martin, 2005; Desmond & MacLachlan, 2005; Dunbar, et al., 2000; Hunt-Shanks, et al., 2009; Martin, Lewin, & Thompson, 2003; Rodgers, et al., 2005; Schönberger & Ponsford). Furthermore, despite the simplicity inherent in using HADS-Total, combining all HADS item is contradictory to prevailing views that anxiety and depression are distinct phenomena (Dunbar, et al., 2000).

Three-dimensional structures.

In 1991, Clark and Watson (1991) proposed a tripartite model of anxiety and depression to explain the overlap in anxiety and depressive symptoms and high comorbidity rates. The tripartite model groups symptoms of depression and anxiety into three subtypes: symptoms of general distress or negative affectivity (nonspecific), physiological hyperarousal (specific to anxiety), and anhedonia and low positive affectivity (specific to depression). This model argues that although anxiety and depression do have distinctive features, both share characteristics with ‘general distress’. Since the publication of this model, there has been considerable interest in examining whether a number of scales conform to the tripartite model of anxiety and depression, including the HADS (Dunbar, et al., 2000). Although the HADS-Anxiety has few items capturing somatic symptom, Dunbar, Ford, Hunt and Der (2000) proposed the following three-dimensional structure for the HADS based on the tripartite model: '*negative affectivity*' (HADS-Anxiety items 1, 5, 7, and 11), '*autonomic anxiety*' (HADS-Anxiety items 3, 9, and 13), and '*anhedonic depression*' (HADS-Depression items and HADS-Anxiety item 7) subscales (see Table 1). Both single order and hierarchical models were put forward, where negative affectivity 'causes' anhedonic depression and autonomic anxiety; however, these models have been found to be largely equivalent in terms of fit by Dunbar et al. (2000) and others (Barth & Martin, 2005; Hunt-Shanks, et al., 2009; Martin, et al., 2003; Rodgers, et al., 2005). Friedman et al. (2001) also proposed a three-dimensional structure, derived from principal component analysis (using orthogonal and oblique rotations), which includes the '*psychomotor agitation*' (HADS-Anxiety items 1, 7, and 11), '*psychic anxiety*' (HADS-Anxiety items 3, 5, 9, and 13) and the original HADS-Depression subscales (see Table 1).

The three-dimensional structures of the HADS stand out not only because of their theoretical foundation, but also because in a number of confirmatory factor analyses they meet the CFI and RMSEA criteria more so than the original, two-dimensional structure and unidimensional structures (Barth & Martin, 2005; Desmond & MacLachlan, 2005; Dunbar, et al., 2000; Martin, et al., 2003; Rodgers, et al., 2005). Some confirmatory factor analyses have suggested that Dunbar et al.'s (2000) model offers a superior fit to the data than Friedman et al.'s (2001) one (Martin, et al., 2003; Rodgers, et al., 2005). Foreseeable advantages of using a three-dimensional structure, includes increased sensitivity and specificity for anxiety and depression (Dunbar, et al., 2000; Martin, 2005), responsiveness to interventions, and predictive validity (Martin, 2005). Although support for the three-dimensional structures is mounting and is suggesting that it might be tapping into the fundamental structure of the HADS (Martin, et al., 2006; Rodgers, et al., 2005), the HADS was not originally designed according to the tripartite model and psychometric analyses are further needed .

Rasch Analysis of the HADS

The widespread use of the HADS and debate surrounding how it should be used underpin the theoretical and practical importance of further examining the psychometric properties of the different item combinations (Desmond & MacLachlan, 2005). To inform clinicians' and researchers' decisions about the most appropriate use of the HADS, recent studies have advocated drawing on contemporary psychometric methods, such as Rasch analysis (Forjaz, Rodriguez-Blázquez, & Martinez-Martin, 2009; Gibbons et al., 2011; Kendel et al., 2010; Lambert, Pallant, & Girgis, 2011; Pallant & Tennant, 2007; Smith, et al., 2006; Tang, Wong, Chiu, Lum, & Ungvari, 2008; Tang, Wong, Chiua, & Ungvari, 2007). Rasch analysis is a rigorous psychometric approach increasingly used to identify measurement issues not easily detected by classical test theory analyses (e.g., item bias, response format) (Hagquist, Bruce, & Gustavsson, 2009; Tennant & Conaghan, 2007). Rasch analyses of the HADS have been performed on data from individuals with stroke (Tang, et al., 2007), Parkinson's disease (Forjaz, et al., 2009), chronic obstructive pulmonary disease (Tang, et al., 2008), or motor neuron disease (Gibbons, et al., 2011), individuals undergoing coronary artery bypass graft surgery (Kendel, et al., 2010), or attending an outpatient musculoskeletal rehabilitation program (Pallant & Tennant, 2007), and caregivers of cancer survivors (Lambert, et al., 2011). Only one study has used Rasch to analyze the HADS among individuals with cancer (Smith, et al., 2006). Findings of these studies are summarized in Table 2. Unidimensionality of HADS-Anxiety is typically supported; however, misfit for items 3 '*Frightened feeling as if something awful is about to happen*' (Tang, et al., 2008), 7 '*Sit at ease and feel relaxed*' (Pallant & Tennant, 2007), and 11 '*Have to be on the move*'(Gibbons, et al., 2011; Lambert, et al., 2011; Smith, et al., 2006; Tang, et al., 2008) have been noted. Studies have also supported the unidimensionality of HADS-Depression, with items 8 '*Slowed down*'(Gibbons, et al., 2011; Kendel, et al., 2010; Lambert, et al., 2011), 10 '*Lost interest in my appearance*'(Kendel, et al., 2010; Smith, et al., 2006), and 14 '*Enjoy a good book or radio or TV program*'(Gibbons, et al., 2011; Kendel, et al., 2010; Smith, et al., 2006) showing misfit. Three out of four studies examining the unidimensionality of HADS-Total supported its use as a measure of psychological distress (Forjaz, et al., 2009; Pallant & Tennant, 2007; Smith, et al., 2006). None of these Rasch studies have examined the fit of the subscales derived from aforementioned three-dimensional structures.

Aims

The present paper examines the psychometric properties of the HADS using Rasch analysis, in a large-scale, population-based sample of cancer survivors. Our three aims are to examine the psychometric properties of the subscales corresponding to each of the alternative structures of the HADS (see Table 1 for overview of subscales tested):

a) Original two HADS-Depression and HADS-Anxiety subscales proposed by Zigmond et al. (1983);
b) Anxiety and depression subscales derived from Dunbar et al.'s (2000) and Friedman et al.'s (2001) three-dimensional structure; and
c) HADS-Total proposed by Razavi et al. (1990).

Despite previous evidence suggesting that the HADS-Total is less fitting than the original two- and three-dimensional models, it is still increasingly popular in psycho-oncology and to provide a comprehensive overview of the psychometric advantages and disadvantages of competing HADS subscales, it was included in this analysis.

Methods

This paper is based on the HADS data collected at six months post-diagnosis as part of a longitudinal *Cancer Survival Study.* The study protocol has been reported elsewhere (Boyes, Girgis, D'Este, & Zucca, 2011).

Participants

As notification of cancer to the Cancer Registry is a statutory requirement in Australia, and 60% of all new cancer cases are diagnosed amongst people living in the states of New South Wales (NSW) or Victoria (VIC), the state-based cancer registries of NSW and VIC were used as the population sampling frame (Australian Institute of Health and Welfare & Australasian Association of Cancer Registries, 2006). Eligibility was restricted to those diagnosed with their first histologically confirmed primary cancer of any of the top eight incident cancer types in Australia (colorectal, female breast, prostate, melanoma of the skin, lung, Non-Hodgkin's lymphoma, head and neck, or leukaemia); aged 18-80 years at diagnosis; resident of NSW or VIC; considered physically and mentally capable of participating by their clinician; having adequate English skills to complete a survey; and aware of their cancer diagnosis. Of the 3315 eligible individuals identified by the registries, 1691 consented to be contacted about the study by the research team and were mailed a self-administered, scannable survey. A total of 1360 survivors completed a 6-month (referred to as Time 1) survey. Participants' median age was 63 years (range 19 to 81), 59% were male and 79% were married or living with a partner. Table 3 shows that the most common diagnosis was prostate cancer (26%) followed by breast cancer (16%) and melanoma (15%). Slightly more than half of participants (52%) were diagnosed with early stage disease and 72% had received surgical treatment. Almost three-quarters (73%) had not received any active treatment in the last month. The study was approved by the relevant Human Research Ethics Committees.

Data Collection

The *Cancer Survival Study* survey assessed a comprehensive range of factors, including socio-demographic characteristics and anxiety and depression as measured by HADS (Zigmond & Snaith, 1983). For each of the 14 HADS items, participants were asked to rate how they felt in the past *month*. Each item is on a 4-point scale from 0 to 3, with the response options varying across items (Zigmond & Snaith, 1983). Non-responders received one mailed reminder package three weeks later and one reminder phone call another three weeks later.

Data Analysis

Rasch analysis involves the testing of an outcome scale against a mathematical measurement model developed by the Danish mathematician Georg Rasch (Rasch, 1960). The Rasch measurement model assumes that the probability of a participant endorsing an item is a logistic function of the relative difference between the item's (difficulty of the item) and the person's (ability of the person) location. The mathematical Rasch model is considered the formal representation of 'proper' measurement against which data are examined. Hence, the overall objective of the analysis is to test the extent to which the observed pattern of item responses conforms to Rasch model expectations (Pallant & Tennant, 2007; Tennant & Conaghan, 2007). When the observed response pattern coincides with, or does not deviate too much from, the expected response pattern, the items are said to 'fit' the measurement model. In Rasch analysis, deviation from model expectations or 'misfit' points to potential improvements and revisions of the measure (Hagquist, et al., 2009). Initially, because the HADS has polytomous response options, a likelihood ratio test was conducted for each subscale and examined to determine whether it was more appropriate to use the Rating Scale (Andrich, 1978) or the Partial Credit model (Masters, 1982). The probability value of the likelihood ratio test was significant for all subscales ($p<.001$), which means that the distances between threshold varies across items and it is more appropriate to use the Partial Credit model than the Rating Scale model.

HADS data were initially entered into SPSS and then exported into the Rasch Unidimensional Models for Measurement (RUMM) software version 2030 (Andrich, Lyne, Sheridan, & Luo, 2010). Separate Rasch analyses were conducted for each of the aims, according to the established Rasch analysis protocols published by Pallant and Tennant (2007) and Tennant and Conaghan (2007), including examining summary and individual model fit statistics, person separation index (PSI), appropriateness of the response format (item thresholds), item bias (differential item functioning or DIF), redundancy (local dependency), and dimensionality. These analyses are further detailed below and the criteria used to assess model fit are summarised in Table 4.

Overall fit to the Rasch model was initially assessed using three statistics: a) item-trait interaction chi-square probability value, b) item fit residual standard deviation (SD), and c) person fit residual SD. As an indication of good fit, it was expected that the chi-square probability value would be non-significant (using Bonferroni alpha value adjusted to the number of items) and that the fit residual SD for items and persons would be less than 1.5 (Shea, Tennant, & Pallant, 2009). Given the sensitivity of the chi-square statistics to large sample sizes, the residual statistics were used primarily to guide decision-making concerning fit.

The Person Separation Index (PSI) provides an indication of the internal consistency of the scale and the power of the measure to discriminate amongst respondents with different levels of the trait being measured. The PSI is interpreted in the same way as a Cronbach's alpha coefficient where 0.70 is considered a minimal value for group or research use and 0.85 for individual or clinical use (Tennant & Conaghan, 2007). Measures of internal consistency (PSI, Cronbach Alpha) are influenced by the number of items in the scale; therefore some caution should be exercised when interpreting the results for subscales with few items. For these short subscales it is often recommended that the mean inter-item correlation be reported, with a minimum recommended value of 0.20, with an optimal range of 0.20 to 0.40 (Briggs & Cheek, 1986).

Individual item and person fit residual values were also inspected to identify items and/or persons that might be contributing to misfit (i.e., values outside the range ± 2.5). High positive fit residual values indicate misfit, while high negative fit residuals suggest item redundancy.

Threshold maps were examined to identify disordered thresholds, which would indicate problems with the response format. For good model fit, it is expected that respondents with high levels of anxiety or depression would endorse high scoring response options on each of the items, while individuals with low levels of anxiety or depression would consistently endorse low scoring responses. In Rasch analysis terms, this would be indicated by a set of ordered response *thresholds* for each item. The term *threshold* refers to the point between two response categories where either response is equally probable. That is the point where, for example, the probability of scoring a 0 on the item or scoring a 1 is 50/50. Respondents not using the response categories in a manner that is consistent with the level of the trait being measured is a common source of item misfit and results in *disordered thresholds*. If a disordered threshold was detected, item rescoring was considered, informed by the item's category probability curve.

Item bias, or differential item functioning (DIF), can occur when different groups within the sample, despite equal levels of anxiety or depression, respond in a different manner to an individual item (Pallant & Tennant, 2007). Essentially, the scale should work in the same way, irrespective of the group assessed (Tennant, McKenna, & Hagell, 2004). This does not preclude a different score between younger and older participants or male and females, but rather indicates that, given the same level of anxiety or depression, the expected score on any item should be the same irrespective of age or gender (Tennant & Conaghan, 2007). When one group shows a consistent difference in their responses to an item, across the whole range of the attribute being measured, this is referred to as *uniform DIF*. When there is non-uniformity in the differences between the group, or DIF varies across levels of the attribute, this is referred to as *non-uniform DIF*. To detect DIF in RUMM2030, analysis of variance (ANOVA) (Bonferroni adjusted alpha level) of the standardized response residuals was conducted for each item across each level of the person factors and class interval (i.e., at different levels of trait). Respondents were classified into four age groups for DIF analysis: 18-49, 50-59, 60-69, and 70 and older. It is recognized that the younger age group includes a broad age range; however, this is consistent with other studies (Smith, et al., 2002) and ensures enough participants to carry out the analysis in this age group. When an item was found to exhibit DIF (statistically and graphically), deletion was considered, particularly if removal improved overall model fit (Tennant & Pallant, 2007).

Assumptions of local independency and unidimensionality were also assessed (Pallant & Tennant, 2007). Local independence means that the response to any item is unrelated to any other item when the level of the construct is controlled for. To identify local dependency, the residual correlation matrix was examined and pairs of items with correlations exceeding 0.3 were taken to indicate dependency. Unidimensionality implies that only one construct is measured by a set of items in a scale (or subscale). To examine the unidimensionality of the subscales and HADS-Total, principal component analysis of the residuals was performed to identify the two subsets of items that show the most difference from one another. Differences between person estimates (location values) derived from these two subsets of items were compared using a series of t-tests. If more than 5% of these tests are significant (or specifically the lower bound of the binomial confidence interval is above 5%), the scale is multidimensional (Tennant & Conaghan, 2007). This approach has been shown to be robust to simulated levels of multidimensionality in polytomous scales (Tennant & Pallant, 2006).

In this analysis, 1354 survivors were included in the analysis (six survivors excluded due to more than 50% of questions on HADS-Anxiety or HADS-Depression missing), which is adequate for the Rasch analyses conducted (Linacre, 1994).

Results

Aim 1 - Analysis of Original HADS Subscales

Analysis of HADS-Anxiety.

Although no significant misfit for persons was detected, a high summary fit residual SD for items (SD=3.84) suggested the presence of misfitting items (i.e., value exceeds 1.5) (Analysis 1 Table 5). No disordered thresholds were detected. The fit residual for item 11 '*Have to be on the move*' of 8.04 suggested that removing it might improve model fit. However, even after this item was excluded, the overall fit residual SD for items still exceeded 1.5 (SD=1.95) (Analysis 2 Table 5). The high positive item fit residual value for item 7 '*Sit at ease and feel relaxed*' (fit residual=3.37) suggested that it should also be excluded. Deleting item 7 resolved model misfit (Analysis 3 Table 5). After removing these items, the PSI value was low 0.73; however, the mean inter-item correlation was acceptable (r=0.56).

No DIF was observed for age or sex, and no local dependency was present. There was no evidence of multidimensionality with a series of independent t-tests, comparing person estimates from subtests identified using PCA of the residual, indicating only 2.44% statistically significant tests.

Analysis of HADS-Depression.

Analysis of the HADS-Depression items revealed initial misfit to the Rasch model expectations, as indicated by an item fit residual SD of 2.81 (Analysis 4 Table 5). Although items 2 '*Enjoy the things I used to enjoy*' and 14 '*Enjoy a good book or radio or TV program*' displayed disordered thresholds, attempts at rescoring these items (by collapsing responses 2 and 3) did not improve model fit and the original scoring was retained (Analyses 5 and 6 Table 5).

To achieve fit to the Rasch model, it was necessary to remove item 12 '*Look forward with enjoyment to things*', which recorded a high, negative fit residual value (fit residual =-7.25) (Analysis 7 Table 5). DIF by sex was evident for item 10 '*Lost interest in my appearance*'; that is, at equivalent levels of depression, women endorsed the item at higher levels than men. DIF by age was also identified for item 6 '*Feel cheerful*' indicating that at equal levels of depression, participants of different age vary on the likelihood of endorsing the item. As the level of DIF was relatively minor and deleting these items did not improve model fit (lower PSI), no additional action was taken (Analyses 8 and 9 Table 5). No local dependency was detected. A series of t-tests performed on the person estimates from two subsets of items identified from principal component analysis of the residuals revealed that only 1.85% of cases had statistically significant t-values. The PSI value was low at 0.62; however, the mean inter-item correlation was acceptable (r=0.39).

Aim 2 - Analysis of the Subscales Derived from the Three-Dimensional Structure Models

Dunbar et al.'s (2000) three-dimensional structure.

The *autonomic anxiety* subscale (items 3, 9, and 13) showed fit to model expectation (Analysis 10 Table 5). However, the *negative affectivity* subscale (items 1, 5, 7, and 11) showed misfit as indicated by a high summary item fit residual SD (SD=2.66) (Analysis 11 Table 5). A high individual item fit residual value for item 11 '*Have to be on the move*' (fit residual = 4.20) indicated misfit and suggested possible item deletion. Once item 11 was removed, no additional items showed misfit (Analysis 12 Table 5). There was no evidence of disordered thresholds, local dependency, or DIF across these subscales. However, the PSI was relatively low for these anxiety subscales, ranging from 0.48 to 0.65; however, the mean inter-item correlations (r= 0.56-0.57) suggested adequate internal consistency given the small number of items involved.

Analysis of the *Anhedonic depression* subscale (items 2, 4, 6, 7, 8, 10, 12, and 14) revealed initial misfit to the Rasch model expectations as indicated by a high summary item fit residual SD (SD=2.90) (Analysis 13 Table 5). Item 14 ‘*Enjoy a good book or radio or TV program*’ displayed a disordered threshold; however, rescoring by collapsing responses 2 ‘*not often*’ and 3 ‘*very seldom*’ did not improve model fit (SD=2.90) and original scoring was retained (Analysis 14 Table 5). To improve model fit, it was necessary to remove item 12 (Analysis 15 Table 5), followed by item 6 (Analysis 16 Table 5). No local dependency was detected. DIF by sex was evident for items 2 and 10. DIF by age was also identified for items 7 and 8. As the level of DIF was minor and deleting these items did not improve fit, no additional action was taken. The PSI was relatively low, but the mean inter-item correlation was acceptable.

A series of t-tests performed on the person estimates from two subsets of items identified from principal component analysis of the residuals revealed that only 2.66% of cases had statistically significant t-values.

Friedman et al.’s (2001) three-dimensional structure.

Both the *psychic anxiety* (items 3, 5, 9, and 13) and *psychomotor agitation* (items 1, 7, and 11) subscales showed appropriate model fit as indicated by summary item fit residual SDs of less than 1.5: *psychic anxiety* SD= 0.55 and *psychomotor agitation* SD = 0.92 (Analyses 17 and 18 Table 5). All individual item fit residuals were within the $\pm$ 2.5 range. Although the PSI values were relatively low, ranging from 0.54 to 0.62, the mean inter-item correlations were adequate. No disordered thresholds, no local dependency, and no DIF for age or sex were detected across these subscales.

As the depression subscale proposed in this model mirrors the original HADS-Depression, analyses are not repeated and the reader is referred to Analyses 4-9, Table 5.

Aim 3- Analysis of HADS-Total

Analysis of all HADS items combined revealed initial misfit to the Rasch model expectations, as indicated by a high summary item fit residual SD (SD=3.43) (Analysis 19 Table 6). Dimensionality analysis revealed two subsets of items, which closely reflected the distribution of items across the HADS-Anxiety and HADS-Depression subscales, with the exception of item 7. A series of t-tests performed on the person estimates from these two subsets of items revealed that 11.09% (95% CI= 9.90 – 12.2) of cases had statistically significant t-values, indicating multidimensionality. Even when the HADS-Anxiety and HADS-Depression subscales were modified as suggested in the previous sections (i.e., delete HADS-Anxiety items 7 and 11 and HADS-Depression item 12), unidimensionality was not supported (Analysis 20 Table 6). In addition, a high summary fit residual SD value for the items suggested overall model misfit (SD=2.76) (Analysis 20 Table 6). Disordered thresholds were detected for items 2 and 14; however, rescoring (collapsing responses 2 '*not often*' and 3 '*very seldom*') only marginally improved overall model fit (SD = 2.72; Analysis 21 Table 6) and the original response scale was retained.

Different sequences of item deletion were examined to improve overall model fit (Analyses 22 and 23 Table 6), with the goal of achieving fit with minimal deletion. An optimal solution was achieved by deleting items 5 '*Worrying thoughts go through my mind*', 6 '*Feel cheerful*', and 9 '*like 'butterflies' in the stomach'* (Analysis 24 Table 6). A series of t-tests performed on the person estimates from the two subsets of items identified from principal component analysis of the residuals revealed that 4.51% of cases had statistically significant t-values, supporting its unidimensionality. The final solution (Analysis 24 Table 6), consisting of three HADS-Anxiety (i.e., 1, 3, 13) and five HADS-Depression (i.e., 2, 4, 8, 10, 14) items, showed adequate fit to the model. Although the PSI was low, the mean inter-item correlation was acceptable.

Several items showed DIF for age. HADS-Anxiety items 1 and 3 were more likely to be endorsed by the younger participants and HADS-Depression item 8 was more likely to be endorsed by the older participants. A number of items showed DIF for sex where HADS-Anxiety items 1, 3, and 13 were more likely to be endorsed by women and HADS-Depression items 2, 4, and 14 were more likely to be endorsed by men. As almost half of the items had already been deleted and the other statistics did not support the HADS-T, no further action was taken.

Discussion

Since its development, the psychometric properties of the HADS have been extensively studied, mainly using classical test theory approaches (Barth & Martin, 2005; Bjelland, et al., 2002; Gough & Hudson, 2009; Hunt-Shanks, et al., 2009; Smith, et al., 2002) and more recently using Rasch analysis (Forjaz, et al., 2009; Kendel, et al., 2010; Lambert, et al., 2011; Pallant & Tennant, 2007; Smith, et al., 2006; Tang, et al., 2008; Tang, et al., 2007). To our knowledge, this is the first Rasch analysis to examine model fit of the unidimensional (Razavi, Delvaux, Farvacques, & Robaye, 1990), the original, two-dimensional (Zigmond & Snaith, 1983) and three-dimensional (Dunbar et al., 2000; Friedman et al., 2001) structures among cancer survivors. Although the autonomic anxiety, psychic anxiety, and psychomotor agitation subscales showed initial fit to the model expectations, all other subscales showed model misfit and item deletion was necessary, with the HADS-Total requiring the most modifications.

Unidimensionality analyses supported the original HADS-Anxiety and HADS-Depression subscales; however, some modifications were necessary to achieve adequate model fit. For HADS-Anxiety, misfitting items 7 '*Sit at ease and feel relaxed*' and 11 '*Have to be on the move'* were removed. Both of these items assess restlessness, agitation, or tension and were problematic in several other studies using classical test theory (Barth & Martin, 2005; Bjelland, et al., 2002; Gough & Hudson, 2009; Hunt-Shanks, et al., 2009) and Rasch analysis (Lambert, et al., 2011; Pallant & Tennant, 2007). Numerous factor analyses have found item 7 to jointly load on HADS-Anxiety and HADS-Depression (Barth & Martin, 2005; Bjelland, et al., 2002; Gough & Hudson, 2009; Smith, et al., 2002), with some analyses showing higher loading on HADS-Depression (Barth & Martin, 2005; Gough & Hudson, 2009; Smith, et al., 2002). The positive wording of item 7 might explain these results, as the HADS-Anxiety has few positively worded items, whereas most HADS-Depression items are positively worded (Schönberger & Ponsford). In addition to item 7, item 11 was also removed in this analysis to improve the fit of the HADS-Anxiety. A recent Rasch analysis of the HADS among caregivers of cancer survivors undertaken by members of the research team (Lambert, et al., 2011) also found that removing this item resolved model misfit. Schönberg and Ponsford (2010) used confirmatory factor analysis to examine fit of the original two-dimensional structure among individuals with traumatic brain injury and found that whereas anxiety items 9 and 13 had the highest loadings, item 11 had the lowest. Moreover, although restlessness and not feeling relaxed are indicators of anxiety, these symptoms might also be due to other factors, including pain, physical discomfort or commonly used agitation-inducing medications such as corticosteroids. In the present analysis, after removing items 7 and 11, the remaining five HADS-Anxiety items showed good internal consistency and there was no evidence of misfitting items, disordered thresholds, or DIF.

For HADS-Depression to achieve model fit, it was necessary to delete item 12 '*Look forward with enjoyment to things*'. Other Rasch analyses have typically identified misfit for items 8 '*Slowed down*' (Gibbons, et al., 2011; Kendel, et al., 2010; Lambert, et al., 2011), 10 '*Lost interest in my appearance*'(Kendel, et al., 2010; Smith, et al., 2006), and 14 '*Enjoy a good book or radio or TV program*' (Kendel, et al., 2010; Smith, et al., 2006). Traditional factor analyses have found item 12 to load highly on the factor corresponding to HADS-Depression and, along with item 2, has been identified as a marker item of this factor (Gough & Hudson, 2009; Smith, et al., 2002). The misfit of item 12 in this study might be due to an altered perception of the future unique to cancer survivors who have faced potential mortality. Responding from this context, perhaps the question regarding '*looking forward*' was imbued with a grander, more existential meaning by survivors rather than a question solely about mood, therefore generating a misfit.

Contemporary studies of the HADS are increasingly suggesting that the three-dimensional structure as proposed by Dunbar et al. (2000) or Friedman et al. (2001) offers superior fit to data compared to the original two- dimensional structure. Friedman et al.'s '*psychic anxiety*' and '*psychomotor agitation*' subscales showed model fit without any adjustment. Although Dunbar's '*autonomic anxiety*' subscale showed initial model fit, item deletion was necessary to fit the '*negative affectivity*' subscale. Friedman et al.'s three- dimensional structure has been supported among individuals with heart disease (Barth & Martin, 2005), the homeless and individuals socially marginalised (Martin, et al., 2006), and post-myocardial infarction (Martin, et al., 2003) and across cultures (Martin, Thompson, & Barth, 2008). A study by Rodgers et al. (2005) tested seven structures, using classical test theory analyses, among women diagnosed with breast cancer and found that Dunbar et al's (2000) three- dimensional model (single-order) offered best fit followed by Friedman et al's (2001) model. Zigmond and Snaith's (1983) original two- dimensional model was fourth and the unidimensional structure of Razavi et al. (1990) was last. One notable limitation to the three-dimensional structures in the present study is the subscales' low PSI, which may be due in part to the small number of items. Others have also found low internal consistency for the '*psychomotor agitation*' (Friedman, et al., 2001; Hunt-Shanks, et al., 2009).

Although HADS-Total is increasingly used as an outcome measure in psycho-oncology research (Jacobsen, et al., 2005; Sellick & Edwardson, 2007; Singer, et al., 2009; Walker, et al., 2007), the present findings caution against combining all HADS items to form a measure of overall distress. These findings also contribute to the general debate about the conceptualization of anxiety and depression and suggest these exist as distinct phenomena. Clinically, a concern of using the HADS-Total is that high scores on one subscale might be masked by low score on the other (resulting in an overall moderate HADS-Total score). This conclusion is consistent with a commentary by Snaith (1991) stressing that the HADS was originally developed to distinguish between the constructs of anxiety and depression and not as a screening tool for overall distress.

Research and Clinical Implications

The present analysis provides further insights into the strengths and weaknesses of the different HADS subscales across the original one, two, and three dimensional structures. Our findings corroborate those of others regarding misfit of some HADS-Anxiety (e.g., item 11) and HADS-Depression (e.g., item 7) items. These two subscales might benefit from modifications, including rewording misfitting items, removing them altogether, and/or proposing alternate items. In particular, negatively wording item 7 in line with most of the other HADS-Anxiety items might be fruitful in resolving misfit. However, further research on the psychometric properties of these revised HADS subscales is required before they can be applied with confidence in either clinical or research settings.

Findings also add to the mounting support for the subscales part of the three- dimensional structure. No studies have examined the psychometric properties of these subsclaes among individuals with cancer. Future studies are needed to examine the incremental validity of using the three-dimensional structure, particularly in terms of sensitivity and specificity, responsiveness, and predictive validity, including the extent to which it might lead to more effective treatment (Dunbar, et al., 2000). Although continued reliance on the two- dimensional structure has raised concerns regarding the potential of under-estimating caseness (Martin, 2005), it is possible that the incremental validity of using the three-dimensional structure over the two-dimensional one is so small that it might not be worth the additional time that might be needed to score a more complex scale (Hunsley & Meyer, 2003). Also, future studies need to determine whether the three-dimensional structure decreases the ease of use of the HADS in clinical settings (Rodgers, et al., 2005).

Strengths and Limitations

Using the two largest state-based cancer registries in Australia as the sampling frame is a major strength of this study. Although the study sample is generally representative of its source population, the use of rapid case ascertainment procedures and registry policies prohibiting individuals being approached for more than one study meant that the sampling frame from which the sample was recruited was incomplete. Ideally, the sample recruited from both states would have been stratified by cancer type proportionate to its incidence in that state. The 44% recruitment rate, although superior to other similar studies conducted elsewhere (Smith et al., 2007), might raise concerns about sample bias. This is an inherent consequence of the multistep recruitment process routinely used by cancer registries to identify potential study participants on behalf of the research team due to privacy, confidentiality, and adverse event concerns. The *Cancer Survival Study* is unique in the diversity of the cancer survivors included in terms of primary cancer type, extent of disease, and geographic location. However, despite the multicultural nature of the Australian population, survivors who were not proficient in English were excluded due to prohibitive costs involved in translation of the study survey into other languages. In addition, the authors acknowledge that the baseline, 6-month data collection time point might not be representative of the four others, and future analyses are needed to examine temporal stability. Furthermore, Rasch analysis is one of three potential Item Response Theory approaches and although these approaches have fundamental paradigm differences, there is considerable debate about the potential advantages of one approach over another with Rasch at times criticized for being more restrictive than the other approaches (Khalid, Hussain, Hussain, & Riaz, 2011).

Conclusion

The HADS is one of the most commonly used tools for measuring anxiety and depression among individuals diagnosed with cancer. However, there is considerable debate concerning the scales' most appropriate structure. The present paper assessed the strengths and weaknesses of the different subscales derived for the one, two, and three dimensional strucutures of the HADS. Findings can guide clinicians' and researchers' decision-making on the most appropriate use of the HADS among cancer survivors. To our knowledge, this is the first Rasch analysis to examine the HADS within a large sample of cancer survivors in the early phases of survivorship. The present analysis provides further insights into some of the contemporary developments regarding the scales' best use. Overall, results support modified versions of the HADS-Anxiety and HADS-Depression subscales; however, before these can be integrated in any future use of the HADS, further research is needed to identify revised cut-points. While findings supported the use of the subscales proposed by three-dimensional models of the HADS, the combination of all items to form a HADS-Total is not recommended. Additional studies are also needed to explore the clinical and diagnostic utility of the three dimensional structure.

Table 1

Subscales corresponding to Each HADS Structure Tested

Study Aims	**Labels of subscales and allocation of items to subscales**
1. Two-dimensional structure of Zigmond et al. (1983)	Anxiety – 1, 3, 5, 7, 9, 11, 13 Depression– 2, 4, 6, 8, 10, 12, 14
2a. Three-dimensional structure of Dunbar et al. (2000)	Autonomic anxiety - 3, 9, 13 Negative affectivity - 1, 5, 7, 11 Anhedonic depression - 2, 4, 6, 7, 8, 10, 12, 14
2b. Three-dimensional structure of Friedman et al. (2001)	Psychic anxiety - 3, 5, 9, 13 Psychomotor agitation – 1 ,7 , 11 Depression– 2, 4, 6, 8, 10, 12, 14
3. Unidimensional structure of Razavi et al. (1990)	All items combined

Table 2

Summary of the Findings from Rasch Studies of the HADS

Author	**Sample**	**Findings**		
		HADS-Anxiety	**HADS-Depression**	**HADS-Total**
Forjaz et al. (2009)	387 Parkinson's disease patients	Good fit to model expectations Item 3 rescored by adjoining the 2nd and 3rd categories No DIF; PSI = .80 Unidimensionality supported	Failed to fit model expectations No DIF All items displayed ordered thresholds Unidimensionality not explored	Failed to fit model expectations Misfit for items 2,4,7, and 8 Rescoring of item 3 and 8 by adjoining the 2nd and 3rd categories No DIF; PSI = .87 Unidimensionality supported
Gibbons et al. (2011)	298 individuals with motor neurone disease	Failed to fit model expectations Item 11 removed Local dependency between items 3 and 5	Failed to fit model expectations Item 8 removed No DIF Items 2 and 14 rescored	Failed to fit model expectations Items 2 and 14 rescored Item 10 removed Items with local dependency combined Unidimensionality supported
Kendel et al. (2010)	1271 patients undergoing coronary artery bypass	-	Failed to fit model expectations Unidimensionality supported	-

Author	Sample	Findings		
		HADS-Anxiety	HADS-Depression	HADS-Total
	graft surgery		Items 8,10, and 14 removed DIF gender for items 2 and 14	
Lambert et al. (2011)	541 partners or caregivers of individuals with cancer	Failed to fit model expectations Item 11 removed No disordered threshold maps No DIF Unidimensionality supported	Failed to fit model expectations Item 14 rescored by adjoining the 2nd and 3rd response categories and removing item 8; Unidimensionality supported	Unidimensionality not supported
Pallant et al. (2007)	296 patients attending an outpatient musculoskeletal rehabilitation program	Misfit to model expectation, which seems attributable to item 7	Good fit to the model expectations	Initial misfit to model expectations Misfit for item 11; Rescoring of item 2 by adjoining the 2nd and 3rd categories; Minor DIF item 13 Unidimensionality supported
Tang et al. (2007)	100 chinese with acute stroke	-	Unidimensionality supported Dichotomous response set might be more appropriate Revised version similar screening	-

Author	Sample	Findings		
		HADS-Anxiety	HADS-Depression	HADS-Total
			performance as original	
Tang et al. (2008)	166 chinese patients with chronic obstructive pulmonary disease	Unidimensionality supported Borderline misfit items 3 and 11 Item 13 found to be redundant No DIF for age, sex, education, six-minute walking distance, and forced expiratory volume. Item 13 DIF for arterial oxygen saturation	-	-
Smith et al. (2006)	381 individuals with cancer	Unidimensionality supported Misfit item 11 Removal of misfitting item led to a small reduction in screening performance	Unidimensionality supported Misfit items 10 and 14 Removal of misfitting item led to a small reduction in screening performance	Unidimensionality supported Misfit items 10, 11, and 14 Screening efficacy marginally better than that of the HADS-A and HADS-D Removal of misfitting items had little impact on screening efficacy

Table 3

Selected Socio-Demographic and Disease Characteristics of Study Sample Compared to National Cancer Incidence Data

	Study sample*		National[†]	
	n	%	n	%
Gender	**1360**		**58 665**	
Male	806	59	34 223	58
Female	554	41	24 442	42
Age (years)	**1360**		**58 665**	
18-39	49	4	2826	5
40-49	137	10	6032	10
50-59	327	24	13 049	22
60-69	487	36	18 451	31
70 or more	360	26	18 307	31
Primary cancer	**1360**		**58 665**	
Prostate	357	26	13 886	24
Breast (female)	212	16	10 896	19
Melanoma	208	15	9197	16
Bowel	162	12	10 108	17
Non Hodgkin's Lymphoma	158	12	3127	5
Lung	135	10	7312	12
Head & neck	97	7	2331	4
Leukaemia	31	2	1808	3
Stage of disease at diagnosis	**1360**			
Early/less progressed	709	52		
Late/more progressed	263	19		
Not applicable	189	14		

	Study sample*		National†	
	n	%	n	%
Unknown	199	15		
Treatment received in first 6 months‡				
Surgery (n=1346)	976	72		
Radiotherapy (n=1347)	395	29		
Chemotherapy (n=1346)	447	33		
Hormone treatment (n=1344	230	17		
Marital status	**1356**			
Married/defacto/living with partner	1074	79		
Separated/divorced/widowed/never married	282	21		
Highest level of education	**1353**			
Primary school	79	6		
Secondary school	656	48		
Certificate or Diploma	344	25		
University	274	20		
Health insurance status	**1346**			
Public	454	34		
Private	892	66		
State	**1360**			
Victoria	828	61		
New South Wales	532	39		

* number of observations varies across characteristics due to missing data

† restricted to 8 most incident cancers and those aged 20-79 years; data not available for all characteristics

‡ multiple responses permitted

Table 4

Rasch Statistics Used to Examine HADS Items

Rasch statistic	Fit Criteria
Overall Model Fit	
Item-trait interaction chi-square probability value	Non-significant Bonferroni adjusted probability*
Summary items mean and SD	Perfect fit =Mean of zero and SD of one
	Acceptable fit = SD < 1.5 (Shea, et al., 2009)
Summary persons mean and SD	Perfect fit =Mean of zero and SD of one
	Acceptable fit = SD < 1.5 (Shea, et al., 2009)
Person separation index (PSI)	Values greater or equal to .85
Individual Person Fit Residuals	Fit residual = ± 2.5
Individual Item Fit Residuals	Fit residual = ± 2.5
Thresholds	Ordered thresholds
Local dependency	No positive correlations greater than 0.3
Differential item functioning (uniform and non-uniform examined)	Non-significant Bonferroni adjusted probability
Dimensionality	Less than 5% of t-tests are significant

Note. SD= standard deviation. * Given the sensitivity of the chi-square statistics to large sample sizes (in this case n=1354), the residual statistics were used primarily to guide decision-making concerning fit.

Table 5

Model Fit Statistics for the Original and Revised Two- and Three-Dimensional Structures (Aims 1 and 2)

Action	**Analysis**	**Overall Model Fit**	**Items Fit Residual Mean (SD)**	**Persons Fit Residual Mean (SD)**	**PSI**	**Mean inter-item correlation**	**% Significant T-tests**
HADS-Anxiety							
Original	1	X^2 = 238.62 p < .001	- 0.13 (3.84)	-0.32 (1.04)	0.77	0.50	3.18
Item 11 removed	2	X^2 = 104.17 p < .001	-0.21 (1.95)	-0.37 (1.01)	0.77	0.54	3.47
Items 11 and 7 removed	3	X^2 =87.60 p < .001	-0.30 (1.23)	-0.39 (0.96)	0.73	0.56	2.44
HADS-Depression							
Original	4	X^2 =219.76 p < .001	-1.77 (2.81)	-0.45 (0.88)	0.67	0.42	4.51
Item 2 rescored (0122)	5	X^2 =231.17 p < .001	-1.72 (2.77)	-0.44 (0.88)	0.67	-	1.63
Item 14 rescored (0122)	6	X^2 =208.96 p < .001	-1.80 (2.82)	-0.44 (0.89)	0.67	-	1.77
Item 12 removed	7	X^2 = 171.34 p < .001	-1.76 (1.24)	-0.44 (0.82)	0.62	0.39	1.85
Items 12 and 6 removed±	8	X^2 = 144.02 p < .001	-1.94 (1.25)	-0.42 (0.75)	0.60	0.36	-
Items 12 and 10 removed±	9	X^2 = 186.68 p < .001	-1.79 (1.29)	-0.44 (0.77)	0.61	0.40	-
Dunbar et al.'s (2000) three-dimensional structure							
Autonomic anxiety (3, 9, 13)	10	X^2 = 48.87 p < .001	0.87 (0.44)	-0.27 (0.77)	0.48	0.57	--
Negative affectivity (1, 5, 7,	11	X^2 = 89.40 p < .001	0.43 (2.66)	-0.45 (1.20)	0.65	0.49	--

Action	Analysis	Overall Model Fit	Items Fit Residual Mean (SD)	Persons Fit Residual Mean (SD)	PSI	Mean inter-item correlation	% Significant T-tests
11)							
Negative affectivity - item 11 removed	12	$X^2 = 37.58$ $p < .001$	0.24 (1.70)	-0.56 (1.22)	0.65	0.56	--
Anhedonic depression (2, 4, 6, 7, 8, 10, 12, 14)	13	$X^2 = 205.88$ $p < .001$	-0.92 (2.90)	-0.39 (0.89)	0.75	0.43	4.21
Anhedonic depression-Item 14 rescored (0122)	14	$X^2 = 195.19$ $p < .001$	-0.95 (2.90)	-0.39 (0.89)	0.75	-	3.47
Anhedonic depression - Item 12 removed	15	$X^2 = 150.48$ $p < .001$	-0.79 (1.77)	-0.38 (0.84)	0.72	0.40	2.07
Anhedonic depression- Item 12 and 6 removed[¥]	16	$X^2 = 122.44$ $p < .001$	-0.77 (1.15)	-0.37 (0.78)	0.69	0.38	2.66
Friedman et al.'s (2001) three-dimensional structure							
Psychic anxiety (3, 5, 9, 13)	17	$X^2 = 52.20$ $p < .001$	0.56 (0.55)	-0.30 (0.89)	0.62	0.58	--
Psychomotor agitation (1, 7, 11)	18	$X^2 = 49.43$ $p < .001$	0.59 (0.92)	-0.55 (1.17)	0.54	0.47	--

Note. Bonferonni adjustment to the alpha level for the number of items; X^2= Chi-Square; p = probability. SD= Standard Deviation; PSI= Person Separation Index. [¥]Items listed in the order they were deleted.

Table 6

Model Fit Statistics for Original and Revised HADS-Total (Aim 3)

Action	Analysis	Overall Model Fit	Items Fit Residual Mean (SD)	Persons Fit Residual Mean (SD)	PSI	Mean inter-item correlation	% Significant T-tests
All 14 original items	19	X^2 =464.77 p < .001	-0.12 (3.43)	-0.31 (1.06)	0.84	0.40	11.09
Remove items 11, 7, and 12[¥]	20	X^2 =278.86 p < .001	-0.53 (2.76)	-0.34 (1.00)	0.79	0.40	9.31
Remove items 11 , 7, and 12 and rescore item 14 (0122)	21	X^2 =276.76 p < .001	-0.57 (2.72)	-0.34 (1.00)	0.79	-	9.02
Remove items 11, 7, 12, 8, 2, 14, 10, and 5[¥]	22	X^2 =125.04 p< .001	-0.70 (1.25)	-0.40 (.93)	0.67	0.46	2.59
Remove items 11, 7, 12, 5, 6, 14, 9[¥]	23	X^2 =154.92 p< .001	-0.69 (1.04)	-0.33 (.85)	0.72	0.39	4.73
Remove items 11, 7, 12, 5, 6, 9[¥]	24	X^2 = 134.06 p< .001	-0.59 (1.37)	-0.34 (.88)	0.72	0.37	4.51

Note. Bonferonni adjustment to the alpha level for the number of items; X^2= Chi-Square; p = probability. SD= Standard Deviation; PSI= Person Separation Index. [¥]Items listed in the order they were deleted.

References

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*(4), 561-573.

Andrich, D., Lyne, A., Sheridan, B., & Luo, G. (2010). *RUMM2030*. Perth, Western Australia.

Armaiz-Pena, Guillermo N., Lutgendorf, Susan K., Cole, Steve W., & Sood, Anil K. (2009). Neuroendocrine modulation of cancer progression. *Brain, Behavior, and Immunity, 23*(1), 10-15.

Australian Institute of Health and Welfare & Australasian Association of Cancer Registries. (2006). Cancer in Australia: An overview: Cancer series no 37. Cat no CAN 32. Canberra: AIHW, 2007.

Barth, Jurgen, & Martin, Colin. (2005). Factor structure of the Hospital Anxiety and Depression Scale (HADS) in German coronary heart disease patients. *Health and Quality of Life Outcomes, 3*(1), 15.

Bjelland, I, Dahl, AA., Haug, TT., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research, 52*(2), 69-77.

Boyes, A.W., Girgis, A., D'Este, C., & Zucca, A.C. (2011). Flourishing or floundering? Prevalence and correlates of anxiety and depression among a population-based sample of adult cancer survivors 6 months after diagnosis. *Journal of Affective Disorders, 135* 184-192.

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality, 54*(1), 106-148. doi: 10.1111/j.1467-6494.1986.tb00391.x

Carlson, L.E., Angen, M., Cullum, J., Goodey, E., Koopmans, J., Lamont, L., . . . Bultz, B.D. (2004). High levels of untreated distress and fatigue in cancer patients. *British Journal of Cancer, 90*, 2297-2304.

Clark, Lee Anna, & Watson, David. (1991). Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*(3), 316-336.

Desmond, Diedre M., & MacLachlan, Malcolm. (2005). The factor structure of the Hospital Anxiety and Depression Scale in older individuals with acquired amputations:a comparison of four models using confirmatory factor analysis. *International Journal of Geriatric Psychiatry, 20*, 344-349. doi: DOI: 10.1002/gps.1289

Dunbar, M., Ford, G., Hunt, K., & Der, G. (2000). A confirmatory factor analysis of the Hospital Anxiety and Depression Scale: Comparing empirically and theoretically derived structures. *British Journal of Clinical Psychology, 39*, 73-94.

Forjaz, Maria João, Rodriguez-Blázquez, Carmen, & Martinez-Martin, Pablo. (2009). Rasch Analysis of the Hospital Anxiety and Depression Scale in Parkinson's disease. *Movement Disorders, 24*(4), 526-532.

Friedman, S., Samuelian, J-C., Lancrenon, S., Even, C., & Chiarelli, P. (2001). Three-dimensional structure of the Hospital Anxiety and Depression Scale in a large French primary care population suffering from major depression. *Psychiatry Research, 104*(3), 247-257.

Gibbons, C. J., Mills, R. J., Thornton, E. W., Ealing, J., Mitchell, J. D., Shaw, P. J., . . . Young, C. A. (2011). Rasch analysis of the Hospital Anxiety and Depression Scale (HADS) for use in motor neurone disease. [Research Support, Non-U.S. Gov't]. *Health and Quality of Life Outcomes, 9*, 82. doi: 10.1186/1477-7525-9-82

Given, B., Wyatt, G., Given, CW., Gift, A., Sherwood, A., DeVoss, D., & Rahbar, M. (2005). Burden and depression among caregivers of patients with cancer at the end-of-life. *Oncology Nursing Forum, 31*(6), 13.

Gough, K., & Hudson, P. (2009). Psychometric properties of the Hospital Anxiety and Depression Scale in family caregivers of palliative care patients. *Journal of Pain and Symptom Management, 37*(5), 797-806. doi: DOI:10.1016/j.jpainsymman.2008.04.012

Hagquist, C., Bruce, M., & Gustavsson, J. (2009). Using the Rasch model in nursing research: An introduction and illustrative example. *International Journal of Nursing Studies, 46*, 380-393.

Honda, K., Goodwin, R. D., & Neugut, A. I. (2005). The associations between psychological distress and cancer prevention practices. *Cancer Detection and Prevention, 29*(1), 25-36.

Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: conceptual, methodological, and statistical issues. *Psychological assessment, 15*(4), 446-455.

Hunt-Shanks, T., Blanchard, C., Reid, R., Fortier, M., & Cappelli, M. (2009). A psychometric evaluation of the Hospital Anxiety and Depression Scale in cardiac patients: addressing factor structure and gender invariance. [Corrected proof]. *British Journal of Health Psychology*. doi: DOI:10.1348/135910709X432745

Jacobsen, P.B., Donovan, K.A, Trask, P.C, Fleishman, S.B, Zabora, J, Baker, F, & Holland, J.C. (2005). Screening for psychologic distress in ambulatory cancer patients: A multicenter evaluation of the distress thermometer. *American Cancer Society, 103*, 1494-1502.

Kayser, K., & Scott, J. L. (2008). *Helping couples cope with women's cancers: an evidence-based approach for practitioners*. New York: Springer Science + Business Media.

Kendel, F., Wirtz, M., Dunkel, A., Lehmkuhl, E., Hetzer, R., & Regitz-Zagrosek, V. (2010). Screening for depression: Rasch analysis of the dimensional structure of the PHQ-9 and the HADS-D. *Journal of Affective Disorders, 122*(3), 241-246.

Lambert, S. D., Pallant, J. F., & Girgis, A. (2011). Rasch analysis of the Hospital Anxiety and Depression Scale among caregivers of cancer survivors: Implications for its use in psycho-oncology. *Psycho-Oncology, 20*(9), 919–925.

Lewis, F. M., Fletcher, KA, Cochrane, BB, & Fann, JR. (2008). Predictors of depressed mood in spouses of women with breast cancer. *Journal of Clinical Oncology, 26*(8), 7.

Linacre, JM. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions 7*(28), 328.

Luckett, T., Butow, P. N., King, M. T., Oguchi, M., Heading, G., Hackl, N. A., . . . Price, M. A. (2010). A review and recommendations for optimal outcome measures of anxiety, depression and general distress in studies evaluating psychosocial interventions for English-speaking adults with heterogeneous cancer diagnoses. *Supportive Care Cancer, 18*, 1241-1262.

Lynch, B. M., Steginga, S. K., Hawkes, A. L., Pakenham, K. I., & Dunn, J. (2008). Describing and predicting psychological distress after colorectal cancer. *Cancer, 112*(6), 1363-1370.

Martin, C. R. (2005). What does the Hospital Anxiety and Depression Scale (HADS) really measure in liaison psychiatry settings? *Current Psychiatry Reviews, 1*, 69-73.

Martin, C. R., Bonner, A., Brook, A., & Luscombe, C. (2006). Factor structure and use of the Hospital Anxiety and Depression Scale in the homeless and socially marginalized. *Psychology, Health & Medicine, 11*(2), 190-197. doi: DOI:10.1080/13548500500155883

Martin, C. R., Lewin, R. J. P., & Thompson, D. R. (2003). A confirmatory factor analysis of the Hospital Anxiety and Depression Scale in coronary care patients following acute myocardial infarction. *Psychiatry Research, 120*(1), 85-94.

Martin, C. R., Thompson, D.R., & Barth, J. (2008). Factor structure of the Hospital Anxiety and Depression Scale in coronary heart disease patients in three countries. *Journal of Evaluation in Clinical Practice, 14*(2), 281-287. doi: 10.1111/j.1365-2753.2007.00850.x

Masters, GN. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47,* 149-174.

Meyerowitz, Beth E., & Oh, Sindy. (2009). Psychosocial response to cancer diagnosis and treatment. In S. M. Miller, D. J. Bowen, R. T. Croyle & J. H. Rowland (Eds.), *Handbook of Cancer Control and Behavioral Science: A Resource for Researchers, Practitioners, and Policymakers* (pp. 361-389). Washington, DC: American Psychological Association.

Mitchell, A. J., Meader, N., & Symonds, P. (2010). Diagnostic validity of the Hospital Anxiety and Depression Scale (HADS) in cancer and palliative settings: A meta-analysis. [Meta-Analysis]. *Journal of Affective Disorders, 126*(3), 335-348. doi: 10.1016/j.jad.2010.01.067

National Breast Cancer Centre, National Cancer Control Initiative. (2003). *Clinical Practice Guidelines for the Psychosocial Care of Adults with Cancer.* Canberra: Commonwealth of Australia.

National Comprehensive Cancer Network. (2007). Clinical practice guidelines in oncology: Distress management.

Pallant, J. F., & Tennant, A. (2007). An introduction to the Rasch measurement model: An example using the Hospital Anxiety and Depression Scale (HADS). *British Journal of Clinical Psychology, 46*, 1-18.

Rabin, E. G., Heldt, E., Hirakata, V. N., Bittelbrunn, A. C., Chachamovich, E., & Fleck, M. P. A. (2009). Depression and perceptions of quality of life of breast cancer survivors and their male partners. *Oncology Nursing Forum, 36*(3), 6.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press.

Razavi, D., Delvaux, N., Farvacques, C., & Robaye, E. (1990). Screening for adjustment disorders and major depressive disorders in cancer in-patients. *The British Journal of Psychiatry, 156*(1), 79-83. doi: 10.1192/bjp.156.1.79

Rodgers, J. , Martin, C., Morse, R., Kendell, K., & Verrill, M. (2005). An investigation into the psychometric properties of the Hospital Anxiety and Depression Scale in patients with breast cancer. *Health and Quality of Life Outcomes, 3*(1), 41.

Schönberger, M., & Ponsford, J. (2010). The factor structure of the Hospital Anxiety and Depression Scale in individuals with traumatic brain injury. *Psychiatry Research, 179*(3), 342-349.

Sellick, S. M., & Edwardson, A. D. (2007). Screening new cancer patients for psychological distress using the Hospital Anxiety and Depression Scale. *Psycho-Oncology*, *16*(6), 534-542.

Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the Depression, Anxiety and Stress Scales (DASS). *BMC Psychiatry, 9*(21).

Singer, S., Kuhnt, S., Götze, H., Hauss, J., Hinz, A., Liebmann, A., . . . Schwarz, R. (2009). Hospital anxiety and depression scale cutoff scores for cancer patients in acute care. *British Journal of Cancer, 100*(6), 908-912.

Smith, A. B., Selby, P. J., Velikova, G., Stark, D., Wright, E. P., Gould, A., & Cull, A. (2002). Factor analysis of the Hospital Anxiety and Depression Scale from a large cancer population. *Psychology and Psychotherapy: Theory, Research and Practice, 75*(2), 165-176.

Smith, A. B., Wright, E. P., Rush, R., Stark, D. P., Velikova, G., & Selby, P. J. (2006). Rasch analysis of the dimensional structure of the Hospital Anxiety and Depression Scale. *Psycho-Oncology, 15*(9), 817-827.

Tang, W. K., Wong, E., Chiu, H. F. K., Lum, C. M., & Ungvari, G. S. (2008). Examining item bias in the anxiety subscale of the Hospital Anxiety and Depression Scale in patients with chronic obstructive pulmonary disease. *International Journal of Methods in Psychiatric Research, 17*(2), 104-110. doi: DOI: 10.1002/mpr.234

Tang, W. K., Wong, E., Chiua, H. F. K., & Ungvari, G. S. (2007). Rasch analysis of the scoring scheme of the HADS Depression subscale in Chinese stroke patients *Psychiatry Research, 150*(1), 97-103.

Tennant, A., & Conaghan, P. G. (2007). 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 and Rheumatism, 57*(8), 1358-1362.

Tennant, A., McKenna, S. P., & Hagell, P. (2004). Application of Rasch analysis in the development and application of quality of life instruments. *Value in health: The journal of the International Society for Pharmacoeconomics and Outcomes Research, 7*(Suppl 1), S22-26. doi: 10.1111/j.1524-4733.2004.7s106.x

Tennant, A., & Pallant, J. F. (2006). Unidimensionality Matters! *Rasch Measurement Transactions, 20*, 1048-1051.

Vodermaier, A., & Roanne, D. M. (2011). Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: A systematic review and meta-analysis. *Supportive Care in Cancer, 19*(12), 1899-1908.

Vodermaier, A., & Roanne, D. M. (2011). Accuracy of the Hospital Anxiety and Depression Scale as a screening tool in cancer patients: A systematic review and meta-analysis. *Supportive Care in Cancer*, *19*(12), 1899-1908.

Walker, J., Postma, K., McHugh, G. S., Rush, R., Coyle, B., Strong, V., & Sharpe, M. (2007). Performance of the Hospital Anxiety and Depression Scale as a screening tool for major depressive disorder in cancer patients. *Journal of Psychosomatic Research, 63*(1), 83-91.

Zabora, J., BrintzenhofeSzoc, K., Curbow, B., Hooker, C., & Plantadosi, S. (2001). The prevalence of psychological distress by cancer site. *Psycho-Oncology, 10,* 19-28.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361-370.

Acknowledgments. The *Cancer Survival Study* is led by [name removed] and [name removed] and funded by the National Health & Medical Research Council (NHMRC - #252418), Cancer Council NSW, Hunter Medical Research Institute, Honda Foundation and the University of Newcastle. The Victorian Cancer Registry (Cancer Council Victoria) and NSW Central Cancer Registry (NSW Department of Health and Cancer Institute NSW) assisted with recruitment. Our sincere thanks to the survivors who provided the survey data and [name removed] for research support. [name removed] is supported by the Australian Government through Cancer Australia and [name removed] by an NHMRC Research Training Fellowship.