Abstract

Body concealment is an important component of appearance distress for individuals with disfiguring conditions, including scleroderma. The objective was to replicate the validation study of the Body Concealment Scale for Scleroderma (BCSS) among 897 scleroderma patients. The factor structure of the BCSS was evaluated using confirmatory factor analysis and the Multiple-Indicator Multiple-Cause model examined differential item functioning of SWAP items for sex and age. Internal consistency reliability was assessed via Cronbach's alpha. Construct validity was assessed by comparing the BCSS with a measure of body image distress and measures of mental health and pain intensity. Results replicated the original validation study, where a bifactor model provided the best fit. The BCSS demonstrated strong internal consistency reliability and construct validity. Findings further support the BCSS as a valid measure of body concealment in scleroderma and provide new evidence that scores can be compared and combined across sexes and ages.

Keywords: systemic sclerosis, scleroderma, disfigurement, body image avoidance, body concealment, measurement

Validation of the Body Concealment Scale for Scleroderma (BCSS): Replication in the Scleroderma Patient-centered Intervention Network (SPIN) Cohort

Systemic sclerosis (SSc), or scleroderma, is a rare, chronic, autoimmune connective tissue disease with no known cure. Onset of SSc is highest between the ages of 30-50 years, and women are four to five times more likely to be diagnosed with SSc than men (Mayes, 2003). Some research suggests that individuals from non-White racial/ethnic backgrounds have higher prevalence rates of SSc than White individuals; however, this evidence is limited (Mayes, 2005). Disfiguring appearance changes, including hand contractures, digital ulcers, telangiectasia, hyper- and hypo-pigmentation of the skin, and altered facial features, are common among many SSc patients (Boin & Wigley, 2012). These appearance changes often affect highly visible body parts that are important for social interactions, such as the face, mouth, and hands. The psychological and psychosocial impact of visible differences experienced by people with SSc can be significant, and no treatments alleviate disfiguring disease manifestations. Appearance changes in SSc are associated with body image dissatisfaction, poor appearance self-esteem, social discomfort, and depressive symptoms (Benrud-Larson et al., 2003; Jewett et al., 2012; van Lankveld, Vonk, Teunissen, & van den Hoogen, 2007).

Among people with visible differences, appearance concerns can generate fear of negative evaluation and social anxiety, which, in turn, can lead to avoidance behaviors, such as social withdrawal or concealment of body parts (Newell, 1999; Rumsey & Harcourt, 2004). Body concealment is used as a coping strategy among many individuals with visible differences. People with SSc experience unique appearance changes, particularly skin- and hand-related changes; thus, it is important to assess concealment behaviors specific to SSc in order to increase understanding about the etiology and consequences of these concerns and to improve the scope of clinical assessment of such concerns. Jewett et al. (2016) recently developed a brief self-report measure, the Body Concealment Scale for Scleroderma (BCSS), to assess disease-specific body concealment behaviors related to appearance changes in SSc. The initial validation of the BCSS (N = 742) found that a bifactor structure fit best. The BCSS had two identifiable factors (Concealment with Clothing and Concealment of Hands), but a general factor explained the majority of item covariance, supporting the use of a single total score (Jewett et al., 2016). The BCSS total score demonstrated strong internal consistency reliability (α = .88) and good construct validity (Jewett et al., 2016).

All research findings, including those related to measurement, are subject to chance findings and should be replicated (Thompson, 2004). Therefore, the first objective of the present study was to replicate findings from the initial BCSS validation study in a large, international SSc cohort. The second objective was to examine the measurement equivalency of the BCSS across sex and age. Measurement equivalency occurs when individuals from different groups (e.g., females and males) with similar levels of a construct being measured obtain similar scores on the measure and respond similarly to individual items of the measure (Mokkink et al., 2010). Differential item functioning (DIF), on the other hand, occurs when an item has different measurement properties for one group compared to another, apart from any true differences in the construct being measured (Mokkink et al., 2010). DIF is generally evaluated by identifying differences in individual item scores across groups that remain present after controlling for levels of the overall construct being measured (e.g., body concealment behaviours; Mokkink et al., 2010). When DIF is present, it is assumed that scores on the item are affected by group characteristics that are not directly related to the construct being measured (Mokkink et al., 2010). Measurement equivalency is a central component of assessing the validity of any measure, and it becomes especially important in the context of rare diseases like SSc, where researchers increasingly combine results across groups with distinct characteristics in order to attain large enough sample sizes. However, results should only be combined if measurement equivalency is established (Mokkink et al., 2010). To date, the measurement equivalency of the BCSS has not been established for characteristics that might be expected to influence scores, such as age and sex, as it is well documented in the visible difference literature that younger individuals and those of the female sex report more body image distress and appearance-related concerns (Clarke et al., 2013; Rumsey, Clarke, White, Wyn-Williams, & Garlick, 2004; Tiggeman, 2004; Thompson & Kent, 2001).

Consistent with Jewett et al. (2016), we hypothesized that a bifactor model would provide the best fit for the BCSS. Further, we hypothesized that the BCSS total score would correlate strongly (i.e., $r \geq .50$) with the Brief-Satisfaction with Appearance (Brief-SWAP) Social Discomfort subscale score; to a moderate or strong degree (i.e., $r \geq .30$ to $\leq .50$) with the Brief-SWAP Dissatisfaction with Appearance subscale score, Brief Fear of Negative Evaluation (BFNE-II) total score, Social Interaction Anxiety Scale (SIAS-6) total score, and the Patient-Reported Outcomes Measurement Information System (PROMIS-29v2) Health Profile depression domain score; and to a small to moderate degree (i.e., $r < .30$) with the PROMIS-29v2 pain intensity domain score.

Method

Patients and Procedure

The sample consisted of patients enrolled in the Scleroderma Patient-centered Intervention Network (SPIN) Cohort (Kwakkenbos et al., 2013) who completed baseline study questionnaires between April 2014 and February 2016. Patients were enrolled at 27 SPIN centers in Canada, the USA, and the UK. To be eligible for the SPIN Cohort, patients must be classified as having SSc according to the 2013 American College of Rheumatology/European League Against Rheumatism classification criteria (van den Hoogen et al., 2013), be ≥ 18 years of age, have the ability to give informed consent, be fluent in English or French, and be able to respond to questionnaires via the Internet. The SPIN sample is a convenience sample. Eligible patients are invited by physicians or supervised nurse coordinators from SPIN centers to participate in the SPIN Cohort, and written informed consent is obtained. Local SPIN physicians or nurse coordinators complete an online medical data form to initiate registration. An automated welcoming email is sent to participants with instructions for activating their SPIN account and completing SPIN Cohort measures. SPIN Cohort patients complete outcome measures via the Internet upon enrollment and subsequently every three months (Kwakkenbos et al., 2013). SSc patients in the SPIN Cohort who had complete data at their baseline assessment for all BCSS items were included in the present analyses. The SPIN Cohort study was approved by the Research Ethics Committee of the Jewish General Hospital, Montréal, Canada and by the research ethics committees of each participating center.

Measures

Sociodemographic information was collected via patient self-report, and disease characteristics were obtained from physicians' reports. Disease duration was defined as years since onset of the first non-Raynaud's Phenomenon symptoms. Limited SSc was defined as skin involvement distal to the elbows and knees and diffuse SSc as skin involvement proximal to the elbows and knees or the trunk (van den Hoogen et al., 2013).

Body Concealment Scale for Scleroderma (BCSS). The 9-item BCSS assesses body concealment behaviors in SSc (Jewett et al., 2016). BCSS item responses reflect frequency of current behaviors (0 = *never* to 5 = *always*). Higher total scores reflect more concealment behaviors.

Brief-Satisfaction with Appearance Scale (Brief-SWAP). The 6-item Brief-SWAP (Jewett et al., 2010) assesses body image-related distress in SSc and includes two subscales, Dissatisfaction with Appearance and Social Discomfort. Items are scored on a 7-point scale (0 = *strongly disagree* to 6 = *strongly agree*). Higher scores indicate greater dissatisfaction or social discomfort. The Brief-SWAP has strong internal consistency reliability (α = .81 for both subscales) and good construct validity (Jewett et al., 2010). In the present study, Cronbach's alpha for the Brief-SWAP Social Discomfort subscale was .88 and was .80 for the Brief-SWAP Dissatisfaction with Appearance subscale.

Brief Fear of Negative Evaluation Scale (BFNE-II). The 12-item BFNE-II assesses worry about how one is perceived and evaluated by others (Carleton, Collimore, & Asmundson, 2007). Items are rated on a 5-point scale (0 = *not characteristic of me at all* to 4 = *extremely characteristic of me*). Higher summed total scores indicate greater fear of negative evaluation. The BNFE-II has good internal consistency reliability and validity (Carleton et al., 2007). In the present study, Cronbach's alpha for the BFNE-II was .97.

Social Interaction Anxiety Scale (SIAS-6). The 6-item SIAS-6 assesses anxiety from social interactions (Peters, Sunderland, Andrews, Rapee, & Mattlick, 2012). Items are scored on a 5-point scale (0 = *not at all characteristic of me* to 4 = *extremely characteristic of me*). Higher summed total scores indicate greater interactional anxiety. The SIAS-6 has strong psychometric properties including internal consistency reliability and convergent validity (Peters et al., 2012). In the present study, Cronbach's alpha for the SIAS-6 was .89.

Patient-Reported Outcomes Measurement Information System Health Profile Depression and Pain Domain Scores (PROMIS-29v2). The PROMIS-29v2 assesses patient-reported health status on seven domains, including depression (four items scored on a 5-point scale ranging from 1 = *never* to 5 = *always*), plus a single item for pain intensity (scored on an 11-point rating scale) (Hinchcliff et al., 2011). Higher scores represent more of the measured domain, and raw scores are converted into T-scores standardized based on a general US population sample (M = 50, SD = 10) (Hinchcliff et al., 2011). The PROMIS-29v2 has strong psychometric properties, including construct validity (Hinchcliff et al., 2011). In the present study, Cronbach's alpha for the PROMIS-29v2 Depression Domain Score was .93 and was .97 for the Pain Domain Score.

Data Analysis

The factor structure of the BCSS was evaluated first in the total sample using Confirmatory factor analysis (CFA). Given that in the original BCSS development study, single-factor, two-factor (Concealment with Clothing and Concealment of Hands), and bifactor models were tested, the same three models were replicated in the present study. Bifactor models evaluate the degree to which covariance among a set of item responses can be accounted for by a single general factor that reflects common variance among all items, plus specific factors reflecting additional common variance among clusters of items with similar content (Cook, Kallen, & Amtmann, 2009; Reise, Moore, & Haviland, 2010). Bifactor models can also help assess multidimensionality and evaluate whether a unit-weighted composite score for a single latent trait can be reasonably interpreted, versus creating subscales, in the context of identifiable multidimensionality (Cook et al., 2009; Reise et al., 2010). In the bifactor model, all items were specified to load on the general factor in addition to their designated specific factor

(Concealment with Clothing and Concealment of Hands), and the general and specific factors were specified as orthogonal. To assess the contribution of the general and specific factors to explaining item covariance, we calculated explained common variance (ECV), the ratio of variance explained by the general factor divided by variance explained by the general plus specific factors (Reise, 2012). In addition, coefficient omega, a model-based reliability estimate analogous to coefficient alpha, was calculated for the general and specific factors to evaluate the degree to which the subscales reflected reliable variance beyond variance captured by the general factor (Reise, 2012).

Item responses for the BCSS were modeled as ordinal Likert data, using the weighted least squares estimator with a diagonal weight matrix and robust standard errors and a mean- and variance-adjusted chi-square statistic with delta parameterization (Muthén & Muthén, 2012). A chi-square goodness-of-fit test and three fit indices were used, including the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), Comparative Fit Index (CFI; Bentler, 1990), and Root Mean Square Error of Approximation (RMSEA; Steiger, 1990). Since the chi-square test is highly sensitive to sample size and can lead to the rejection of well-fitting models, practical fit indices were emphasized (Reise, Widaman, & Pugh, 1993). Models with a TLI and CFI close to .95 or higher, and a RMSEA close to .06 or lower, are generally representative of good fitting models (Hu & Bentler, 1999).

In order to determine if the BCSS exhibited DIF for female versus male SSc patients and for younger versus older patients, the Multiple-Indicator Multiple-Cause (MIMIC) model was utilized. MIMIC models for DIF assessment are based on structural equation models, in which the group variable (e.g., sex) is added to the basic CFA model as an observed variable. Therefore, the base MIMIC model consists of the CFA factor model with the additional direct effect of group on the latent factors, which serves to control for group differences on the level of the latent factors (Mokkink et al., 2010). Ideally for DIF assessment, the simplest factor structure with reasonable fit is used.

To assess for potential DIF, the direct effect of group on BCSS items was assessed for each item separately, by regressing the items, one at a time, on group (see Figure 1). Each item was tested separately to determine if there was statistically significant DIF, which is represented by a statistically significant association in the model from group (e.g., sex) to the item, after controlling for any differences in the overall level of the latent factor between groups. If statistically significant DIF was present for one or more items, the item with the largest magnitude of DIF was considered to have DIF, and the link between the group variable (e.g., sex) and that item was included in the model. This procedure was then repeated until none of the remaining items showed statistically significant DIF. Hommel's correction for multiple testing (Hommel, 1988) was applied to determine items with statistically significant DIF.

Once all items with significant DIF were identified, the potential magnitude of all DIF items was evaluated collectively, by conducting comparisons of the difference on the latent factor between groups in the baseline model (not correcting for DIF) and after controlling for DIF. The magnitude of this difference was interpreted using Cohen's effect size suggestions, with ≤ .20 *SD* indicating small, .50 *SD* indicating moderate, and .80 *SD* indicating large differences (Cohen, 1988). CFA and DIF analyses were conducted using MPlus, Version 7.

Cronbach's alpha was computed for the BCSS to assess internal consistency reliability. Construct validity was assessed by calculating Pearson's bivariate correlations of the BCSS total score with the Brief-SWAP subscale scores, PROMIS-29v2 depression and pain domain scores, and BFNE-II and SIAS-6 total scores. These analyses were conducted using SPSS, Version 20.

Because some SSc patients from the SPIN Cohort could have been included in the Canadian Scleroderma Research Group (CSRG) Registry sample used in the original BCSS validation study, we also calculated the maximum possible patient overlap between the two samples based on sites that contribute data to both cohorts. As a sensitivity analysis, we replicated analyses after removing patients from CSRG sites who could have possibly been in the SPIN sample, as it was not possible to determine patient-by-patient overlap.

Results

Sample Characteristics

Of 943 patients, 897 had complete BCSS item data and were included in present analyses. The majority of the sample was female (*N* = 782, 87%) and White (*N* = 774, 86%). Mean age was 55.4 years (*SD* = 12.1), and 41% of patients had diffuse SSc, which is a rate similar to that seen among SSc patients more generally (Mayes, 2005). See Table 1 for a breakdown of all sociodemographic and disease characteristics. There were 214 patients (24%) from CSRG sites whose diagnosis was before the date of the initial validation study using CSRG data and who could have possibly been in both cohorts.

Assessment of the Factor Structure (CFA) of the BCSS

Model fit for the single-factor model was suboptimal ($\chi^2(27) = 753.32$, CFI = .94, TLI = .92, RMSEA = .17). Model fit for the two-factor model was substantially better, ($\chi^2(26) = 165.27$, CFI = .99, TLI = .98, RMSEA = .08). The correlation between the Concealment with Clothing and Concealment of Hands latent factors was .75. Model fit was strongest for the bifactor model ($\chi^2(18) = 81.84$, CFI = .99, TLI = .99, RMSEA = .06). The ECV was .79. Coefficient omega for the Concealment with Clothing subscale was .01 and .34 for the Concealment of Hands subscale. See Table 2 for all model fit parameters.

Assessment of the Differential Item Functioning (DIF) of the BCSS

MIMIC base model. Given that for DIF assessment, the simplest factor structure with reasonable fit is used, the base model consisted of the two-factor model of Concealment with Clothing and Concealment of Hands. This two-factor model was extended to include direct effects of the Concealment with Clothing and Concealment of Hands latent factors on sex (female versus male) and age (dichotomized into age bands of less than 60 years, and 61 years and older), separately. These age bands were determined based on an assessment of the age distribution of the present sample.

The MIMIC base model for sex and age both demonstrated strong fit, $\chi^2(33) = 183.60$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .07 and $\chi^2(33) = 183.15$, $p < .001$, CFI = .99, TLI = .98, RMSEA = .07, respectively. Prior to accounting for DIF, male patients had .57 *SD* lower latent scores than female patients for the Concealment with Clothing factor (95% Confidence Interval [CI] -.82 to -.33) and .16 *SD* lower latent scores than female patients for the Concealment of Hands factor (95% CI -.39 to .07). Prior to accounting for DIF, older patients had .22 *SD* lower latent scores than younger patients for the Concealment with Clothing factor (95% CI -.38 to -.06) and .41 *SD* lower latent scores than younger patients for the Concealment of Hands factor (95% CI -.57 to -.25).

DIF assessment. There were no items that demonstrated statistically significant DIF based on sex. Three items were identified with statistically significant DIF based on age: items 2, (*I wear long sleeves to hide skin changes*), 3 (*I avoid wearing revealing clothes (e.g., bathing suits, tank tops, or shorts)*), and 8 (*I hide my hands so that people don't see them*). More specifically, compared to younger patients, older patients had higher scores (more concealment-related behaviours) than would be expected on items 2 ($z = 2.80$, $p = .005$) and 3 ($z = 3.65$, $p <$

.001), based on their latent levels of the Concealment with Clothing factor. Compared to younger patients, older patients had lower scores than would be expected on item 8 (z = -2.87, p = .004), based on their latent levels of the Concealment of Hands factor.

After correcting for DIF, the difference on the Concealment with Clothing latent factor between younger versus older patients increased from -.22 *SD* (95% CI -.38 to -.06) to -.32 *SD* (95% -.49 to -.15). The difference on the Concealment of Hands latent factor between younger versus older patients decreased slightly from -.41 *SD* (95% CI -.57 to -.25) to -.35 *SD* (95% CI -.52 to -.18). Despite the statistically significant DIF found for three BCSS items, the magnitude was small (all ≤ .20 *SD*s) and did not influence scores substantively. Table 3 shows the baseline CFA parameters, before assessing for DIF, as well as the DIF-corrected model parameters for age.

Assessment of the Reliability and Validity of the BCSS Total Score

Cronbach's alpha for the BCSS total score was .89. Consistent with hypotheses, the BCSS total score was strongly correlated with the Brief-SWAP Social Discomfort subscale (r = .63) and moderately correlated with the Brief-SWAP Dissatisfaction with Appearance subscale (r = 0.44), BFNE-II total score (r = .43), SIAS-6 total score (r = .40), and PROMIS-29v2 depression domain score (r = .48). The correlation with the PROMIS-29v2 pain intensity score was slightly stronger than hypothesized (r = .46).

When analyses were replicated removing possibly overlapping patients from CSRG sites, sociodemographic data, factor analysis results and reliability and validity indices were virtually unchanged (see Supplementary Data file linked online to this article). Given this, DIF analyses were only run with the main model, including the full sample, as it is reasonable to assume that no differences in terms of the measurement equivalency of BCSS items would emerge if overlapping patients were removed.

Discussion

Results replicated the initial BCSS validation study (Jewett et al., 2016). The bifactor model fit best, with the general factor explaining 79% of item covariance. Reliability indices for the specific factors were very low, suggesting that they do not reliably measure substantive variance beyond that explained by the general factor (Reise, 2012). Thus, consistent with the initial BCSS validation, evidence from the present study supports the use and interpretation of a single summed score for the BCSS. Beyond this, BCSS items demonstrated measurement equivalency across female and male patients, and while three items had statistically significant DIF for younger versus older patients, the magnitude of the differences was small, and the influence on scores was negligible.

The present study also found evidence of good internal consistency reliability and construct validity for the BCSS, similar to the initial validation study. As expected, scores on the BCSS correlated most strongly with scores on the Brief-SWAP Social Discomfort subscale, and this relationship highlights the social implications of the disfiguring appearance changes experienced by many SSc patients. The moderate correlation found between the BCSS and Brief-SWAP Dissatisfaction with Appearance subscale scores is not surprising, because BCSS items pertain less to body image dissatisfaction, and more to social and behavioral avoidance. Similarly, the moderate correlations found between scores on the BCSS and scores on the BFNE-II, SIAS-6, and PROMIS-29v2 depression domain likely reflect that these latter measures tap into mental health constructs rather than body image-related avoidance mechanisms.

Evidence provided in the current study points to the utility of the BCSS as a patient-reported outcome in programs designed to address body image concerns among people with SSc, with essentially equivalent measurement properties for both female and male patients and those of different ages. Currently, there are programs being developed (Kwakkenbos et al., 2013) that target disfigurement-related distress and concealment behaviors specific to SSc, building on strategies from social interaction skills training and cognitive behavioral therapy (CBT; Bessell et al., 2012; Clarke et al., 2014) that have been used more generally among individuals with visible differences. The BCSS could be used as an outcome measure to assess body concealment behaviors in such programs.

There are limitations to consider. First, the SPIN Cohort constitutes a convenience sample of SSc patients receiving treatment at SPIN centers, and patients at these centers may differ from those in other settings. SPIN Cohort patients complete questionnaires online, which may further limit the generalizability of findings. Next, the present sample combined both female and male patients, and body concealment behaviors may differ by sex. However, there were not enough men to examine measurement properties separately by sex. Additionally, the majority of our sample (86%) was White; therefore, results may not be generalizable to individuals from non-White racial/ethnic backgrounds. Finally, there was overlap between sites that recruited SSc patients for the SPIN Cohort and the CSRG Registry used in the original BCSS validation study, and it is possible that some patients were included in both. However, even if all patients from those sites were enrolled in both studies, the overlap would have been < 25%. The actual amount of overlap is likely much smaller because the studies were conducted several years apart. Furthermore, results were virtually identical when possibly overlapping patients were excluded.

In summary, the present study replicated previous findings and provided further evidence for the validity of the BCSS by showing that scores can be compared and combined across females and males with SSc, and those of different ages without concern that measurement differences may substantially influence results. The BCSS assesses disease-specific concealment behaviors related to the unique skin and hand appearance changes experienced by SSc patients and highlights body concealment as a pertinent issue for body image and social interactions for individuals with the disease. The BCSS can be used as a clinical assessment tool and as a patient-reported outcome measure to assess body concealment behaviors among people with SSc.

References

Bessell, A., Brough, V., Clarke, A., Harcourt, D., Moss, T., & Rumsey, N. (2012). Evaluation of FACE IT, a computerized psychosocial intervention for disfigurement related distress. *Psychology, Health & Medicine,* 17, 565-577. doi:10.1080/13548506.2011.647701

Benrud-Larson, L. M., Heinberg, L. J., Boling, C., Reed, J., White, B., Wigley, F. M., & Haythornthwaite, J. A. (2003). Body image dissatisfaction among women with scleroderma: Extent and relationship to psychosocial function. *Health Psychology, 22,* 130-139.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Boin, F., & Wigley, F. M. (2012). Clinical features and treatment of scleroderma. In G. S. Firestein, R. C. Budd, S. E. Gabriel, I. B. McInnes, & J. R. O'Dell (Eds.), *Kelley's textbook of rheumatology* (9th ed., pp. 1366-1403). Philadelphia (PA): Elsevier.

Carleton, R. N., Collimore, K. C., & Asmundson, G. J. (2007). Social anxiety and fear of negative evaluation: Construct validity of the BFNE-II. *Journal of Anxiety Disorders, 21,* 131-141. doi:S0887-6185(06)00047-8 [pii]

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York, New York: Lawrence Erlbaum Associates.

Cook, K. F., Kallen, M. A., & Amtmann, D. (2009). Having a fit: Impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Quality of Life Research, 18,* 447-460. doi:10.1007/s11136-009-9464-4 [doi]

Clarke, A., Jenkinson, E., Rumsey, N., Newman, S., Newell, R., Thompson, A., … Members of the ARC Research Collaboration. (2014). *CBT for appearance anxiety: Psychosocial interventions for visible difference in appearance.* London, England: Wiley.

Hinchcliff, M., Beaumont, J. L., Thavarajah, K., Varga, J., Chung, A., Podlusky, S., . . . Cella, D. (2011). Validity of two new patient-reported outcome measures in systemic sclerosis: Patient-reported Outcomes Measurement Information System 29-item Health Profile and Functional Assessment of Chronic Illness Therapy-Dyspnea short form. *Arthritis Care & Research, 63*, 1620-1628. doi:10.1002/acr.20591 [doi]

Hommel, G. (1988). A stagewise rejective multiple test procedure on a modified Bonferroni test. *Biometrika, 75,* 383-386. doi: 10.1093/biomet/75.2.383 [doi]

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 131-152.

Jewett, L. R., Hudson, M., Haythornthwaite, J. A., Heinberg, L., Wigley, F. M., Baron, M., . . . Canadian Scleroderma Research Group. (2010). Development and validation of the Brief-Satisfaction with Appearance scale for systemic sclerosis. *Arthritis Care & Research, 62*, 1779-1786. doi:10.1002/acr.20307 [doi]

Jewett, L. R., Hudson, M., Malcarne, V. L., Baron, M., Thombs, B. D., & Canadian Scleroderma Research Group. (2012). Sociodemographic and disease correlates of body image distress among patients with systemic sclerosis. *PloS One, 7*, e33281. doi:10.1371/journal.pone.0033281 [doi]

Jewett, L. R., Malcarne, V. L., Kwakkenbos, L., Harcourt, D., Rumsey, N., Korner, A., . . . Canadian Scleroderma Research Group CSRG. (2015). Development and validation of the Body Concealment Scale for Scleroderma (BCSS). *Arthritis Care & Research,* doi:10.1002/acr.22819 [doi]

Kwakkenbos, L., Jewett, L. R., Baron, M., Bartlett, S. J., Furst, D., Gottesman, K., . . . Thombs, B. D. (2013). The Scleroderma Patient-centered Intervention Network (SPIN) Cohort: Protocol for a cohort multiple randomised controlled trial (cmRCT) design to support trials of psychosocial and rehabilitation interventions in a rare disease context. *BMJ Open, 3*, 10.1136/bmjopen-2013-003563. doi:10.1136/bmjopen-2013-003563 [doi]

Mayes, M. (2003). Scleroderma epidemiology. *Rheumatic Diseases Clinics of North America, 29*, 239-254.

Mayes, M. (2005). *The scleroderma book: A guide for patients and families, 2nd edition.* New York, New York: Oxford University Press.

McDonald, R. P., & Ho, M. H. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*, 64-82.

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., ... de Vet. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19,* 539-549. doi: 10.1007/s11136-010-9606-8 [doi]

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide version 7.* Los Angeles: California.

Newell, R. J. (1999). Altered body image: A fear-avoidance model of psycho-social difficulties following disfigurement. *Journal of Advanced Nursing, 30,* 1230-1238. doi:jan1185 [pii]

Peters, L., Sunderland, M., Andrews, G., Rapee, R. M., & Mattlick, R. P. (2012). Development of a short form Social Interaction Anxiety (SIAS) and Social Phobia Scale (SPS) using nonparametric item response theory: The SIAS-6 and the SPS-6. *Psychological Assessment, 24*, 66-76.

Reise, S. P. (2012). Invited paper: The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47,* 667-696. doi:10.1080/00273171.2012.715555 [doi]

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92,* 544-559. doi:10.1080/00223891.2010.496477 [doi]

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*, 552-566.

Rumsey, N., Clarke, A., White, P., Wyn-Williams, M., & Garlick, W. (2004). Altered body image: Appearance-related concerns of people with visible disfigurement. *Journal of Advanced Nursing, 48*, 443-453. doi: 10.1111/j.1365-2648.2004.03227.x [doi]

Rumsey, N., & Harcourt, D. (2004). Body image and disfigurement: Issues and interventions. *Body Image, 1*, 83-97. doi:S1740-1445(03)00005-6 [pii]

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173-180. doi:10.1207/s15327906mbr2502_4 [doi]

Tiggeman, M. (2004). Body image across the adult life span: Stability and change. *Body Image, 1*, 29-41. doi: 10.1016/S1740-1445(03)00002-0 [doi]

Thompson, B. (2004). Internal replicability analyses. *Exploratory and confirmatory factor analysis: Understanding concepts and applications* (pp. 99-108). Washington, D.C.: American Psychological Association. doi:10.1037/10694-009

Thompson, A., & Kent, G. (2001). Adjusting to disfigurement: Processes involved in dealing with being visibly different. *Clinical Psychology Review, 21*, 663-682.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

van den Hoogen, F., Khanna, D., Fransen, J., Johnson, S. R., Baron, M., Tyndall, A., . . . Pope, J. E. (2013). 2013 classification criteria for systemic sclerosis: An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. *Arthritis and Rheumatism, 65*, 2737-2747. doi:10.1002/art.38098 [doi] van Lankveld, W. G., Vonk, M. C., Teunissen, H., & van den Hoogen, F. H. (2007). Appearance self-esteem in systemic sclerosis-subjective experience of skin deformity and its relationship with physician-assessed skin involvement, disease status and psychological variables. *Rheumatology, 46*, 872-876. doi:kem008 [pii]

Table 1.
Sociodemographic and Disease Variables for Scleroderma Patients

Sociodemographic Variables	
Age (*M, SD)*	55.4 (12.1)[a]
Female *(n, %)*	782 (87.2)
White *(n, %)*	774 (86.3)
Years of Formal Education (*M, SD)*	15.3 (3.2)
Employed Full or Part-time *(n, %)*	357 (64.4)[b]
Married or Living as Married *(n, %)*	654 (72.9)
Disease Variables	
Diffuse Scleroderma *(n, %)*	366 (40.8)
Disease Duration in Years (*M, SD)*	11.6 (8.8)[c]

Note. N = 897. Due to missing values: [a]N = 894; [b]N = 554; [c]N = 824.

Table 2.
Confirmatory Factor Analysis Models and Parameter Estimates for BCSS[a]

		Single-Factor Model	Two-Factor Model		Bifactor Model			
Item	*M* (*SD*)	Factor Loading[b]	Factor Loading:[b] Clothing[c]	Factor Loading:[b] Hands[d]	Factor Loading:[b] General	Factor Loading:[b] Clothing[c]	Factor Loading:[b] Hands[d]	θ[e]
1. I wear clothes I do not like	.72 (1.11)	.69	.72		.75	-.21		.62
2. I wear long sleeves to hide skin changes	.80 (1.33)	.86	.88		.84	.23		.50
3. I avoid wearing revealing clothes (e.g., bathing suits, tank tops, or shorts)	1.63 (1.75)	.79	.82		.80	.16		.58
4. I wear clothes that hide the changes to my skin	1.06 (1.46)	.92	.95		.92	.38		.40
5. I wear clothes that will divert attention from my appearance	.80 (1.25)	.82	.86		.89	-.12		.46
6. I wear gloves to hide my hands	.81 (1.28)	.72		.77	.59		.50	.64
7. I avoid shaking hands with people	.92 (1.37)	.78		.83	.60		.61	.52
8. I hide my hands so that people don't see them	.98 (1.32)	.83		.89	.68		.56	.47
9. I avoid directly giving change or other items to people	.69 (1.22)	.80		.87	.67		.52	.53

Note. *N* = 897. [a]BCSS responses are scored on a 6-point scale ranging from 0-5 (0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = usually, 5 = always). [b]Factor loadings are unstandardized, raw factor loadings. [c]Concealment with Clothing factor. [d]Concealment of Hands factor. [e]θ = Square root of the error variance. Sum of error variances = 6.50. ECV (explained common variance) = .79. Coefficient ω (omega) = .89. Coefficient $\omega_{H\text{-}GEN}$ (omega hierarchical – general = .80. $\omega_{H\text{-}Clothing}$ = .01. $\omega_{H\text{-}Hands}$ = .34. PUC (percent of contaminated correlations) = 56%.

Table 3.

Factor Loadings of the BCSS Concealment with Clothing and Concealment of Hands Latent Factors and Influence on the Overall Estimates of Concealment with Clothing and Concealment of Hands Latent Factor Scores for Age

Item	Base Model[a]	DIF Corrected Model[b]
	Factor Loading (95% CI)	Factor Loading (95% CI)
Concealment with Clothing Latent Factor		
1. I wear clothes I do not like.	.72 (.68, .76)	.72 (.68, .76)
2. I wear long sleeves to hide skin changes.	.88 (.86, .90)	.88 (.86, .90)
3. I avoid wearing revealing clothes (e.g., bathing suits, tank tops, or shorts).	.82 (.79, .85)	.83 (.80, .86)
4. I wear clothes that hide the changes to my skin.	.95 (.93, .97)	.94 (.92, .96)
5. I wear clothes that will divert attention from my appearance.	.85 (.82, .88)	.85 (.82, .88)
Concealment of Hands Latent Factor		
6. I wear gloves to hide my hands.	.77 (.73, .81)	.77 (.73, .81)
7. I avoid shaking hands with people.	.82 (.79, .85)	.82 (.79, .85)
8. I hide my hands so that people don't see them.	.89 (.87, .91)	.89 (.87, .91)
9. I avoid directly giving change or other items to people.	.86 (.83, .89)	.86 (.83, .89)
Correlation of Concealment with Clothing and Concealment of Hands Latent Factors	.75 (.71, .79)	.75 (.71, .79)
Direct Effects on Items Attributable to Age		
Item 2. Long sleeves to hide skin changes	---	.17 (.05, .29)
Item 3. Avoid wearing revealing clothes	---	.21 (.10, .32)
Item 8. Hide my hands	---	-.17 (-.29, -.05)
Structural Effect of Age on Latent Factors		
Age on Concealment with Clothing factor	-.22 (-.38, .-06)	-.32 (-.49, -.15)
Age on Concealment of Hands factor	-.41 (-.57, -.25)	-.35 (-.52, -.18)

Note. CI = Confidence Interval. [a]Not corrected for Differential Item Functioning (DIF); [b]Corrected for DIF on items 2, 3, 8.