SWAP DIFFERENTIAL ITEM FUNCTIONING

Assessment of English-French differential item functioning of the Satisfaction with Appearance Scale (SWAP) in systemic sclerosis

Lisa R. Jewett¹, Linda Kwakkenbos²*, Marie Hudson³, Murray Baron⁴, Brett D. Thombs⁶, and the Canadian Scleroderma Research Group⁷

¹Department of Educational and Counselling Psychology, McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada.
²Department of Psychiatry, McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; Behavioural Science Institute, Clinical Psychology, Radboud University, Nijmegen, the Netherlands.
³Department of Medicine, McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada.
⁴McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada.
⁵Departments of Educational and Counselling Psychology; Psychiatry; Medicine; Epidemiology, Biostatistics, and Occupational Health; and Psychology, McGill University and Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada.
⁷Canadian Scleroderma Research Group Recruiting Rheumatologists: G. Gyger, Montréal, Québec; J. Pope, London, Ontario; M. Larché, Hamilton, Ontario; N. Khalidi, Hamilton, Ontario; A. Masetto, Sherbrooke, Québec; E. Sutton, Halifax, Nova Scotia; D. Robinson, Winnipeg, Manitoba; T.S. Rodriguez-Reyna, Mexico City, Mexico; D. Smith, Ottawa, Ontario;
C. Thorne, Newmarket, Ontario; P.R. Fortin, Québec, Québec; M. Fritzler, Mitogen Advanced Diagnostics Laboratory, Cumming School of Medicine, Calgary, Alberta.

*Corresponding author:* Linda Kwakkenbos, PhD; Jewish General Hospital; 4333 Cote Ste Catherine Road; Montréal, Québec H3T 1E4; Telephone (514) 340-8222 ext. 6811; E-mail: kwakkenbosl@gmail.com.

**Acknowledgements:** Not applicable.

**Compliance with ethical standards**

**Disclosure of potential conflicts of interest:** Ms. Jewett and Drs. Kwakkenbos, Hudson, Baron, and Thombs declare that they have no conflicts of interest.

**Funding:** The Canadian Scleroderma Research Group (CSRG) has been funded by the CIHR (#FRN-83518), the Scleroderma Society of Canada and its provincial chapters, Scleroderma Society of Ontario, Scleroderma Society of Saskatchewan, Sclérodermie Québec, Cure Scleroderma Foundation, INOVA Diagnostics Inc. (San Diego, CA), Dr. Fooke Laboratorien GmbH (Neuss, Germany), Euroimmun (Lubeck, Germany), Mikrogen GmbH (Neuried, Germany), Fonds de recherche du Québec - Santé (FRQS), the Canadian Arthritis Network, and the Lady Davis Institute of Medical Research of the Jewish General Hospital, Montreal, Quebec. The CSRG has also received educational grants from Pfizer and Actelion pharmaceuticals. Ms. Jewett was supported by a CIHR Doctoral Research Award. Dr. Kwakkenbos was supported by a CIHR Banting Postdoctoral Fellowship. Dr. Thombs was supported by an Investigator Salary Award from the Arthritis Society.
Authors’ contributions: Ms. Jewett, Dr. Kwakkenbos, and Dr. Thombs made substantial contributions to the conception and design, analysis and interpretation of data, and were involved in the drafting of the final manuscript. Drs. Hudson and Baron and the Canadian Scleroderma Research Group made substantial contributions to the acquisition of data and provided feedback on the final draft of the manuscript. All authors read and approved the final manuscript. Dr. Thombs is the overall guarantor.

Research Involving Human Participants: “All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.”

Informed Consent: “Informed consent was obtained from all individual participants included in the study.”
Highlights

- Significant body image concerns are common in systemic sclerosis (SSc).
- As a rare disease, studies in SSc often include patients from multiple countries.
- The Satisfaction with Appearance Scale (SWAP) lacks cross-linguistic validation.
- SWAP scores can be combined for English- and French-speaking SSc patients.
Abstract

The Satisfaction with Appearance Scale (SWAP) has been used to assess body image distress among people with the rare and disfiguring disease systemic sclerosis (SSc); however, it has not been validated across different languages groups. The objective was to examine differential item functioning of the SWAP among 856 Canadian English- or French-speaking SSc patients. Confirmatory factor analysis was used to evaluate the SWAP two-factor structure (Dissatisfaction with Appearance and Social Discomfort). The Multiple-Indicator Multiple-Cause model was utilized to assess differential item functioning. Results revealed that the established two-factor model of the SWAP demonstrated relatively good fit. Statistically significant, but small-magnitude differential item functioning was found for three SWAP items based on language; however, the cumulative effect on SWAP scores was negligible. Findings provided empirical evidence that SWAP scores from Canadian English- and French-speaking patients can be compared and pooled without concern that measurement differences may substantially influence results.

*Keywords*: Satisfaction with Appearance Scale; systemic sclerosis; body image; measurement; differential item functioning
Systemic sclerosis (SSc), or scleroderma, is a rare, chronic autoimmune connective tissue disease. SSc is characterized by abnormal fibrotic processes and excessive production of collagen, which commonly manifest in thickening and hardening of the skin (Boin & Wigley, 2012; Mayes, 2008). As a result, people with SSc often develop visible differences in the texture of the skin, in addition to other disfiguring appearance changes including telangiectasia (visible dilation of blood vessels beneath the skin), hand contractures, skin pigmentation changes, and altered facial features (Mayes, 2008). These appearance changes commonly affect body parts that are highly visible and play a central role in social interactions, such as the face, mouth, and hands (Boin & Wigley, 2012; Rumsey & Harcourt, 2005) and are not alleviated by treatments.

Disfiguring aspects of SSc have been associated with concerns about appearance, attempts to conceal affected body parts, social discomfort, and poor psychological functioning (Jewett, Hudson, Malcarne, Baron, & Thombs, 2012; Jewett et al., 2015, 2016; Malcarne, Hansdottir, Greenberg, Clements, & Weisman, 1999; van Lankveld, Vonk, Teunissen, & van den Hoogen, 2007). Increasingly, research examining body image in SSc has focused on the development and validation of measures that accurately reflect disease-specific body image concerns (Jewett et al., 2010, 2015; Mills et al., 2015). One such measure, the Satisfaction with Appearance Scale (SWAP), which was originally developed with burn patients (Lawrence et al., 1998), but has been adapted and validated in SSc, is commonly used as an outcome to assess dissatisfaction with appearance and social discomfort relative to acquired disfigurements from the disease (Jewett et al., 2010, 2012, 2015; Mills et al., 2015).

The measurement properties of the SWAP have been examined for SSc patients in several studies (e.g., Heinberg et al., 2007; Jewett et al., 2010, 2012; Mills et al., 2015), including those using the same cohort of patients as the present study (Jewett et al., 2010; 2012).
which have collectively established strong evidence of internal consistency reliability, construct validity, and a two-factor structure with Social Discomfort and Dissatisfaction with Appearance factors. However, the degree to which the SWAP generates scores that are equivalent across linguistic or cultural groups has yet to be explored. This is an important consideration because patients who complete the SWAP in countries with more than one common language, such as Canada, or as part of an international cohort (Kwakkenbos et al., 2013), which is often the case in rare diseases like SSc, may do so in different languages (e.g., French, English, Spanish).

Importantly, results pooled across language groups are only valid to the extent that the measurement metric is equivalent, meaning that scores are not affected by linguistic or cultural differences, beyond actual differences in the construct being measured (Mokkink et al., 2010). When the measurement metric is equivalent, individuals from different linguistic groups with similar levels of the construct being measured (e.g., social discomfort) should obtain similar scores on the measure (e.g., SWAP) and respond similarly to individual items of the measure. Differential item functioning (DIF), on the other hand, is said to occur 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). When DIF is present, the scores on the item are likely influenced by group characteristics (e.g., language) that are not directly related to the construct being measured (Mokkink et al., 2010). DIF in cross-linguistic comparisons may occur, for instance, because translations shift meanings, formats, or severity of items used in measures (Zumbo, 1999). As such, it is possible that individuals’ endorsement of items reflects these changes, or some other characteristic of the test item or testing situation (Zumbo, 2007), beyond actual levels of the construct being measured (e.g., social discomfort due to appearance changes). Evaluation of DIF is essential to disambiguate group comparisons and to determine if scoring
differences are actual functions of the outcome being measured, versus an artifact of the measurement process, such as interpretation of item meaning (Teresi & Fleishman, 2007). To date, no studies have assessed DIF for the SWAP in SSc for patients from different linguistic groups. As such, the objective of the present study was to assess the equivalence of SWAP item scores across English- and French-speaking SSc patients in a large Canadian cohort.

Method

Patients and Procedure

The study sample consisted of SSc patients recruited from 15 Canadian Scleroderma Research Group (CSRG) Registry sites. To be eligible for the Registry, patients must be at least 18 years of age, classified as having SSc by a CSRG rheumatologist, and fluent in English or French. Patients completed the SWAP in their preferred language. Over 98% of patients in the Registry meet the 2013 ACR/EULAR classification criteria for SSc (Alhajeri et al., 2015). Each year at their annual Registry visit, patients complete clinical evaluations and fill out a series of self-report questionnaires, including the SWAP. Only patients with complete data for all SWAP items were included in this study. The first available assessment with complete SWAP data was included for patients who had filled out the SWAP on more than one occasion. All patients provided informed consent, and the research ethics board of each participating center approved the data collection protocol.

Measures

Socio-demographic and Disease Characteristics. Patients enrolled in the CSRG Registry provided socio-demographic data, including age, sex, race/ethnicity, education level (greater than a high school level of education versus high school or less), marital status (married or living as married versus unmarried). CSRG rheumatologists provided medical information, including
disease duration based on time since onset of a patient’s first symptom (calculated for both first Raynaud’s or non-Raynaud’s symptoms) and disease subtype (limited or diffuse cutaneous SSc).

Limited cutaneous SSc was defined as skin involvement distal to the elbows and knees only (LeRoy et al., 1988). Patients with sine SSc, which is SSc without skin involvement, were included in the same group as patients with limited cutaneous SSc for the purposes of analysis (Hachulla & Launay, 2011). Diffuse cutaneous SSc was defined as skin involvement proximal to the elbows and knees or the trunk (LeRoy et al., 1988).

_The Satisfaction with Appearance Scale (SWAP)._ The 14-item SWAP was originally designed to measure non-weight related body image concerns among individuals with burn injuries (Lawrence et al., 1998). Items pertain to social discomfort relative to disfigurement (Social Discomfort subscale) and satisfaction with the appearance of particular body parts (Dissatisfaction with Appearance subscale). Respondents rate the degree to which they feel each item reflects their thoughts and feelings about their appearance on a 7-point scale ranging from 0 (strongly disagree) to 6 (strongly agree). Item scores are summed to calculate total subscale scores, reverse-scoring items pertaining to satisfaction with appearance. Higher scores indicate greater social discomfort and dissatisfaction with appearance (Lawrence et al., 1998). The original SWAP was adapted for SSc by replacing the word “burn” with the word “scleroderma” as indicated (Heinberg et al., 2007). The SWAP has consistently demonstrated strong psychometric properties, including internal consistency reliability estimates (i.e., Cronbach’s alphas ≥ .88 for both subscales) and construct validity (i.e., positive moderate correlations with measures of depressive symptoms, quality of life, physical functioning, and disease severity; and low negative correlations with measures of various dimensions of pain) (Jewett et al., 2010, 2012; Mills et al., 2015). Previous studies of the SWAP in SSc samples have consistently found
that a two-factor structure (Social Discomfort and Dissatisfaction with Appearance), but not a one-factor structure, fit the data well (Heinberg et al., 2007; Jewett et al., 2010, 2012).

**Data Analysis**

Descriptive statistics were calculated for all socio-demographic and disease variables, including means and standard deviations (SDs) for continuous variables. Socio-demographic and disease-related variables were compared between English- and French-speaking patients using chi-square tests for categorical variables and t-tests for continuous variables.

The factor structure of the SWAP was evaluated first in the total sample using confirmatory factor analysis (CFA). Ideally for DIF assessment, the simplest structure with reasonable fit is used. Given the previously established two-factor structure of Social Discomfort and Dissatisfaction with Appearance for the SWAP (Heinberg et al., 2007; Jewett et al., 2010, 2012), it was expected that a two-factor structure would fit the data reasonably well in the present study.

Item responses for the SWAP are ordinal Likert data, so the weighted least squares estimator with a diagonal weight matrix, robust standard errors, and a mean- and variance-adjusted chi-square statistic was used with delta parameterization (Muthén & Muthén, 1998-2012). Modification indices were used to identify pairs of items within scales for which model fit would improve if error estimates were freed to covary and for which there appeared to be theoretically justifiable shared method effects (e.g., similar wording; McDonald & Ringo Ho, 2002). To assess model fit, the chi-square test, the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), the Comparative Fit Index (CFI; Bentler, 1990), and the Root Mean Square Error of Approximation (RMSEA; Steiger, 1990) were used. Since the chi-square test is highly sensitive to sample size, it can lead to the rejection of well-fitting models (Reise, Widaman, & Pugh,
1993); therefore, the TLI, CFI, and RMSEA fit indices were emphasized. Good fitting models are indicated by a TLI and CFI ≥ .95 and RMSEA ≤ .06 (Hu & Bentler, 1999).

A general method for assessing DIF involves identifying differences in individual item scores across groups that remain present after controlling for levels of the overall construct being measured (Mokkink et al., 2010). In order to assess if SWAP items exhibited DIF for English-versus French-speaking SSc 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 (i.e., language) is added to the basic CFA model as an observed variable. As such, 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).

To assess for potential DIF, the direct effect of group on SWAP items was assessed for each item separately, by regressing the items, one at a time, on language (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 between language and the item, after controlling for any differences in the overall level of the latent factor between language 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 language 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.

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 CFA model 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). Both CFA and DIF analyses were conducted using MPlus Version 7 (Muthén & Muthén, 1998-2012), and all other analyses were conducted using SPSS Statistics, Version 22.

Results

Sample Characteristics

In total, 856 SSc patients had complete data for all SWAP items and the language variable. Mean age in the total sample was 56.8 (SD = 11.5) years. The majority of patients were female (88%), English-speaking (76%), White (90%), married or living as married (70%), and had limited cutaneous SSc (N = 69%). The mean time since onset of the first disease symptom was 16.3 (SD = 12.2) years, and the mean time since onset of the first non-Raynaud’s disease symptom was 12.2 (SD = 9.4) years. The mean SWAP Social Discomfort subscale score was 8.6 (SD = 8.8), and the mean Dissatisfaction with Appearance subscale score was 19.1 (SD = 12.6). Cronbach’s alpha for the SWAP Social Discomfort subscale was .89 and was .92 for the Dissatisfaction with Appearance subscale. There were no statistically significant differences between English- and French-speaking patients on socio-demographic and disease characteristics, or on SWAP subscale scores (see Table 1).

Confirmatory Factor Analysis (CFA)

For the previously established two-factor structure (Social Discomfort and Dissatisfaction with Appearance), model fit was acceptable based on the CFI and TLI indices, but less than acceptable based on the RMSEA, $\chi^2(76) = 1169.68, p < .001$, CFI = .96, TLI = .96, RMSEA = .13. Inspection of the modification indices indicated that freeing error terms to covary for SWAP
Items 1 and 2 (Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my family and Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my friends) would improve model fit, and there was clear overlap in these items’ content. Modification indices also revealed that freeing error terms to covary for SWAP Items 5 and 6 (I feel that my scleroderma is unattractive to others and I don’t think people would want to touch me) would improve model fit as well. Given these modification indices, and the shared content and meaning of the items (McDonald & Ringo Ho, 2002), the model was refitted to the data, allowing the error terms for these items to be correlated. This change resulted in a model with improved fit, $\chi^2(74) = 846.76, p < .001$, CFI = .97, TLI = .97, RMSEA = .11; however, it was still slightly less than acceptable based on the RMSEA, although sufficiently good-fitting to establish dimensionality for evaluation of DIF (Mokkink et al., 2010). There were no additional modification indices that would have improved model fit substantively. The correlation between the Social Discomfort and Dissatisfaction with Appearance latent factors was .58.

**Differential Item Functioning (DIF)**

*MIMIC Base Model.* To assess possible DIF, the two-factor model was extended to include direct effects of the Social Discomfort and Dissatisfaction with Appearance latent factors on language (English versus French), which demonstrated acceptable fit, $\chi^2(86) = 952.83, p < .001$, CFI = .97, TLI = .97, RMSEA = .11. Table 2 shows the baseline CFA model parameters, before assessing for DIF. Prior to accounting for DIF, there were no statistically significant differences between English- and French-speaking patients for either the Social Discomfort factor (95% Confidence Interval (CI) -.10 to .27) or the Dissatisfaction with Appearance factor (95% CI -.12 to .20).
DIF Assessment. Three items were identified with statistically significant DIF based on language: Items 1 (Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my family), 2 (Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my friends), and 5 (I feel that my scleroderma is unattractive to others). Compared to English-speaking patients, French-speaking patients had higher scores than would be expected on item 1 ($z = 5.60, p < .001$) and item 2 ($z = 2.92, p = .003$), and lower scores on item 5 ($z = -6.94, p < .001$), based on their latent levels of the Social Discomfort factor.

As shown in Table 2, after correcting for DIF, the difference on the Social Discomfort latent factor between English- and French-speaking patients was non-significant (.09 SD, 95% CI -.10 to .27 to .06 SD, 95% CI -.13 to .24). As all three DIF items load onto the Social Discomfort factor, there was no difference on the Dissatisfaction with Appearance latent factor after correcting for DIF. Thus, despite the fact that there was statistically significant DIF for three SWAP items, the cumulative DIF effects across these items did not result in substantive differences in estimates of latent factor levels between English- and French-speaking patients.

Discussion

The main finding of this study was that, although three items demonstrated statistically significant DIF for SSc patients who completed the SWAP in English versus French, the cumulative amount of DIF across these items was small, and the effect on overall SWAP Social Discomfort scores was negligible. Given the large number of tests conducted, interpretations for the differences found at the item-level cannot be made with confidence, and replication of item-specific DIF results is necessary. Taken altogether, however, current findings suggest that SWAP scores can be validly compared and summed across Canadian English-and French-speaking
patients without concern that outcomes will be influenced substantively by differences in scoring metrics.

In DIF assessments, many analyses are conducted, and, particularly when sample size is large, as in the present study, statistically significant findings are not uncommon. In the present study, because the magnitude of the DIF identified was small and the impact on measurement of the SWAP latent constructs was minimal, we did not seek to identify linguistic differences in items identified with DIF.

As large-scale, national and international collaborations in rare chronic diseases like SSc become increasingly common, researchers are often required to integrate data across multiple language groups in order to attain adequate sample sizes. The SWAP is currently used as a key patient-reported outcome measure, for example, in the Scleroderma Patient-centered Intervention Network (SPIN), which collects data from SSc patients in Canada, the USA, and Europe (Kwakkenbos et al., 2013). The SWAP will serve as a primary outcome in a planned SPIN trial to evaluate an online intervention to help people with SSc cope with appearance concerns and body image distress (Kwakkenbos et al., 2013). Results from the present study provide evidence that SWAP scores in English and French can be reasonably combined. Furthermore, it is reasonable to assume, based on the present findings, that items on the recently developed 6-item Brief-SWAP (Jewett et al., 2010) would function similarly in terms of DIF, therefore providing evidence that scores in English and French can be similarly combined for this version of the measure as well.

The present study has limitations that should be considered when interpreting results. First, the CSRG cohort constitutes a convenience sample of SSc patients receiving treatment at one of the sites associated with the Registry. Thus, it was not possible to compare data for patients with
complete SWAP results to all eligible patients. Also, results may depart to some degree from what might be found in a community setting and for different SSc patient groups. However, it has been documented that SSc patients included in the CSRG Registry are similar to patients in other large cohorts in terms of features including age and sex distribution (Hudson et al., 2009). Furthermore, the Registry includes SSc patients across the spectrum of disease subtypes and severity levels, thus, there is evidence supporting that patients sampled in the current study are representative of the general population of SSc patients and those seen in rheumatology clinics more broadly (Hudson et al., 2009). An additional limitation of the present study is that the SSc sample was from Canada only, and, as such, the present findings require replication in other samples of French-speaking patients. Relatedly, DIF was only assessed for English- versus French-speaking patients, and not any for other language groups. We used a MIMIC model to assess DIF. This model implicitly assumes configural invariance, which is a strong assumption (Bauer, 2016). In the case of the SWAP, which was designed to reflect two underlying factors and scored as two subscales, this is likely to be a reasonable assumption, particularly since all previous studies have replicated the two-factor structure. Compared to a multi-group CFA model, which can be used to fully test measurement invariance, including configural invariance, the MIMIC model can provide greater power to detect DIF, if present (Bauer, 2016). This was an important consideration in the present study, since there were only 208 French-speaking SSc patients, and coverage of some parts of the item response spectrum would have been sparse if modelled separately.

In sum, the present study provided evidence that the SWAP is a valid measure of Social Discomfort and Dissatisfaction with Appearance and that items do not function differently across English- and French-speaking SSc patients. There were some SWAP items that exhibited DIF
based on language; however, none of the resulting differences influenced overall scores substantively. Thus, scores on the Social Discomfort and Dissatisfaction with Appearance subscales of the SWAP can be compared and pooled across these languages groups. Replications of the current study with other samples of French-speaking patients are needed.
References


sclerosis: Analysis from the UCLA Scleroderma Quality of Life Study. *Journal of Rheumatology, 42*, 1624-1630. doi: 10.3899/jrheum.141482


Zumbo, B.D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language Assessment Quarterly, 4*, 223-233.
Figure 1. The MIMIC Model for the Satisfaction with Appearance Scale (SWAP)

**SWAP Items**
- Item 1: Family
- Item 2: Friends
- Item 3: Strangers
- Item 4: Relationships
- Item 5: Unattractive
- Item 6: Touch
- Item 7: Overall
- Item 8: Scalp
- Item 9: Face
- Item 10: Neck
- Item 11: Hands
- Item 12: Arms
- Item 13: Legs
- Item 14: Chest

**Latent Factors**
- Social Discomfort
- Dissatisfaction with Appearance

**Grouping Variable**
- Language

*Note.* Dotted line represents the potential DIF effect of language on SWAP items.
Table 1
Socio-demographic and Disease Characteristics by Total and Language Group Samples

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total</th>
<th>English</th>
<th>French</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 856</td>
<td>N = 648</td>
<td>N = 208</td>
<td></td>
</tr>
<tr>
<td>Female n, (%)</td>
<td>751 (87.7)</td>
<td>566 (87.3)</td>
<td>185 (88.9)</td>
<td>.541</td>
</tr>
<tr>
<td>White n, (%)</td>
<td>762 (90.1)</td>
<td>570 (89.2)</td>
<td>192 (98.2)</td>
<td>.138</td>
</tr>
<tr>
<td>Limited cutaneous SSc n, (%)</td>
<td>594 (69.4)</td>
<td>457 (70.5)</td>
<td>137 (65.9)</td>
<td>.205</td>
</tr>
<tr>
<td>Age (years) mean ± SD</td>
<td>56.8 ± 11.5</td>
<td>56.4 ± 11.9</td>
<td>57.9 ± 10.2</td>
<td>.105</td>
</tr>
<tr>
<td>Education (&gt; high school), n, (%)</td>
<td>406 (47.7)</td>
<td>313 (48.6)</td>
<td>93 (44.7)</td>
<td>.329</td>
</tr>
<tr>
<td>Married/living as married n, (%)</td>
<td>594 (69.5)</td>
<td>454 (70.2)</td>
<td>140 (67.3)</td>
<td>.436</td>
</tr>
<tr>
<td>Disease duration (years), mean ± SD</td>
<td>16.3 ± 12.2</td>
<td>16.7 ± 12.4</td>
<td>15.0 ± 11.3</td>
<td>.080</td>
</tr>
<tr>
<td>SWAP Social mean ± SD</td>
<td>8.6 ± 8.8</td>
<td>8.4 ± 8.8</td>
<td>8.9 ± 8.9</td>
<td>.478</td>
</tr>
<tr>
<td>SWAP Dissat mean ± SD</td>
<td>19.1 ± 12.6</td>
<td>19.0 ± 12.9</td>
<td>19.3 ± 11.5</td>
<td>.756</td>
</tr>
</tbody>
</table>

Note. SSc = systemic sclerosis; SD = Standard Deviation; Social = SWAP Social Discomfort subscale; Dissat = SWAP Dissatisfaction with Appearance subscale. Disease duration = time since onset of first symptoms (either Raynaud’s nor non-Raynaud’s). Due to missing values: a N = 846; b N = 852; c N = 855; d N = 639; e N = 644; f N = 647; g N = 207.
Table 2.
Factor Loadings of the SWAP Social Discomfort and Dissatisfaction with Appearance Latent Factors and Influence on the Overall Estimates of Social Discomfort and Dissatisfaction with Appearance Latent Factor Scores

<table>
<thead>
<tr>
<th>Item</th>
<th>Social Discomfort Latent Factor</th>
<th>Dissatisfaction with Appearance Latent Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my family.</td>
<td>.84 (.81, .87)</td>
<td>.84 (.81, .87)</td>
</tr>
<tr>
<td>2. Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of my friends.</td>
<td>.93 (.91, .94)</td>
<td>.93 (.91, .94)</td>
</tr>
<tr>
<td>3. Because of changes in my appearance caused by my scleroderma, I am uncomfortable in the presence of strangers.</td>
<td>.95 (.93, .97)</td>
<td>.95 (.93, .97)</td>
</tr>
<tr>
<td>4. Changes in my appearance have interfered with my relationships.</td>
<td>.76 (.73, .80)</td>
<td>.76 (.73, .80)</td>
</tr>
<tr>
<td>5. I feel that my scleroderma is unattractive to others.</td>
<td>.78 (.75, .81)</td>
<td>.78 (.75, .81)</td>
</tr>
<tr>
<td>6. I don’t think people would want to touch me.</td>
<td>.70 (.66, .74)</td>
<td>.70 (.66, .74)</td>
</tr>
<tr>
<td>7. I am satisfied with my overall appearance.</td>
<td>.86 (.84, .88)</td>
<td>.86 (.84, .88)</td>
</tr>
<tr>
<td>8. I am satisfied with the appearance of my scalp.</td>
<td>.67 (.64, .71)</td>
<td>.67 (.64, .71)</td>
</tr>
<tr>
<td>9. I am satisfied with the appearance of my face.</td>
<td>.89 (.88, .91)</td>
<td>.89 (.88, .91)</td>
</tr>
<tr>
<td>10. I am satisfied with the appearance of my neck.</td>
<td>.85 (.83, .87)</td>
<td>.85 (.83, .87)</td>
</tr>
<tr>
<td>11. I am satisfied with the appearance of my hands.</td>
<td>.72 (.68, .75)</td>
<td>.72 (.68, .75)</td>
</tr>
<tr>
<td>12. I am satisfied with the appearance of my arms.</td>
<td>.86 (.85, .88)</td>
<td>.86 (.85, .88)</td>
</tr>
<tr>
<td>13. I am satisfied with the appearance of my legs.</td>
<td>.79 (.76, .81)</td>
<td>.79 (.76, .81)</td>
</tr>
<tr>
<td>14. I am satisfied with the appearance of my chest.</td>
<td>.87 (.86, .89)</td>
<td>.87 (.86, .89)</td>
</tr>
</tbody>
</table>
### Correlation of Social Discomfort and Dissatisfaction with Appearance Latent Factors

<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation of Social Discomfort and Dissatisfaction with Appearance Latent Factors</td>
<td>.58 (.53, .62)</td>
<td>.58 (.53, .62)</td>
</tr>
</tbody>
</table>

### Direct Effects on Items Attributable to French Language

<table>
<thead>
<tr>
<th>Item</th>
<th>Beta Value</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1. Family</td>
<td>---</td>
<td>.36 (.23, .49)</td>
</tr>
<tr>
<td>Item 2. Friends</td>
<td>---</td>
<td>.17 (.06, .28)</td>
</tr>
<tr>
<td>Item 5. Unattractive</td>
<td>---</td>
<td>-.40 (-.52, -.29)</td>
</tr>
</tbody>
</table>

### Structural Effect of French Language on Latent Factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Value</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language on Social Discomfort factor</td>
<td>.09 (-.10, .27)</td>
<td>.06 (-.13, .24)</td>
</tr>
<tr>
<td>Language on Dissatisfaction with Appearance factor</td>
<td>.04 (-.12, .20)</td>
<td>.04 (-.12, .20)</td>
</tr>
</tbody>
</table>

*Note. Factor loadings represent raw model results. CI = Confidence Interval. *a*Not corrected for Differential Item Functioning (DIF); *b*Corrected for DIF on Items 1, 2, 5.*