Psychometric properties, factorial structure, and measurement invariance of the English and French versions of the Medical Outcomes Study social support scale

by Annie Robitaille, Heather Orpana and Cameron N. McIntosh

May 2011
Psychometric properties, factorial structure, and measurement invariance of the English and French versions of the Medical Outcomes Study social support scale

by Annie Robitaille, Heather Orpana and Cameron N. McIntosh

Abstract

Background
The Medical Outcomes Study (MOS) social support scale is a 19-item survey that measures four dimensions of functional support. The current study reports on the psychometric properties, factorial structure, and measurement invariance of the scale for a sample of English- and French-speaking Canadians aged 55 or older.

Data and methods
The internal consistency and composite reliability for a congeneric measurement model of the dimensions of functional social support were examined. A confirmatory factor analysis and test of invariance across language (English = 2,642; French = 469) were also performed.

Results
Across both English- and French-speaking respondents, results indicated good internal consistency (Cronbach’s alpha ranged from .90 to .97) and composite reliability (ranging from .93 to .97) for all dimensions of functional social support. The confirmatory factor analysis revealed acceptable fit indices for the 4-factor structure similar to the original one. The scale appears to function uniformly across both language groups.

Interpretation
The MOS social support scale appears to be a psychometrically sound instrument for use in research on social support with samples of English- and French-speaking older adults.

Keywords
aging; data analysis; data collection; empirical research; factor analysis; geriatrics; language; questionnaires; social environment; statistical models

Authors
Annie Robitaille (1-613-907-1065; annie.g.robitaille@gmail.com) is with the Centre on Aging, University of Victoria, Victoria, British Columbia. Heather Orpana is with the School of Psychology, University of Ottawa, Ottawa, Ontario. Cameron N. McIntosh is with Public Safety Canada, Ottawa, Ontario, Canada.

Social support plays an important role in successful aging, physical health, mortality, and mental health. Very broadly, it can be defined as the help furnished by an individual’s social network, such as the provision of information, financial aid or emotional support. To understand the association between specific types of social support and mental and physical health, it is recommended that researchers use measures that include as many dimensions as possible and focus on types of support that have been related to positive health outcomes.

The Medical Outcomes Study (MOS) social support scale developed by Sherbourne and Stewart consists of 19 items pertaining to the functional aspects of social support, and one item related to structural social support. The original study was conducted on a sample of 2,987 patients aged 18 to 98 who had chronic health conditions. Based on a confirmatory factor analysis (CFA) on the 19 items designed to measure functional support, the authors reported that a four-factor model was a good fit to the data. The four functional dimensions of social support are: tangible support (material aid and assistance); affectionate support (love and affection); positive social interaction (engaging in entertaining activities with others); and emotional/informational support (feedback, guidance and information). Standardized factor loadings were high for items in each dimension.

The factorial validity of the MOS social support scale was later examined by Gjesfjeld et al. on a sample of 330 mothers whose children were receiving mental health treatment. The authors conducted a CFA on an 18-item scale and on 12- and 4-item abbreviated versions. They found a better-fitting model for the 12- and 4-item versions. These results suggest that some uncertainty remains about the fit of the original version of the scale and that a better-fitting model may be attained if certain items are removed.
Thus, despite the use of the MOS social support scale in numerous studies, more research is needed to test its factor structure and psychometric properties on different populations.29

For example, given that the MOS social support scale is brief, easy to understand, and was designed to minimize respondent burden, it is especially suited for older respondents. However, the psychometric soundness and factor structure of the support scale have not been examined with a national sample of older adults.

As well, information about measurement invariance (MI) between English- and French-speaking respondents on the scale is lacking. This is especially important in Canada, where comparisons between these groups are common. According to Statistics Canada and the Canadian Institute for Health Information, rates of emotional and informational support are relatively low in Quebec,28 but more analysis is needed to determine if these are true differences or artifacts of translation. Before studies of social support can report findings from English- and French-speaking respondents, MI must be established to ensure that the constructs have the same meaning for each group.

Statistics Canada’s National Population Health Survey (NPHS), which incorporates the MOS social support scale, is administered in English and French. Two earlier studies that translated the scale from English to French and examined its psychometric properties found good internal consistency, reliability and convergent validity.30 However, the translation used by the NPHS was designed independently from these other versions, so it is unlikely that the psychometric results would be the same.

The primary purpose of this study is to examine the internal consistency of the English and French versions of the MOS social support scale for a sample of older adults. The second objective is to conduct a CFA to assess the factor structure of the English and French versions of the scale. A third purpose is to determine if the items comprising the scale operate in the same way for English- and French-speaking respondents.

Methods

Sample and data

The data are from the household component of the longitudinal National Population Health Survey (NPHS), which has collected data from the same individuals every two years since 1994/1995. The household component covers the population of the ten provinces, excluding full-time members of the Canadian Forces, residents of Indian Reserves and Crown Lands, residents of some remote areas in Ontario and Quebec, and all residents (military and civilian) of Canadian Forces bases.34 The Health Institutions component of the NPHS, which was administered to residents of health care institutions, was ended after cycle five and was not included in this study.

For all provinces except Quebec, a stratified two-stage sample design, where dwellings were selected within clusters, was used. The design was based on the Labour Force Survey (LFS). The Quebec sample was selected based on a two-stage sample design from households participating in the “Enquête sociale et de la santé.”34 The longitudinal sample size was 17,276. In cycle 3 (1998/1999), the response rate for the 17,276 respondents in the longitudinal data file was 88.3%.33 Computer-assisted telephone and personal interviewing was used to collect the data; the majority of interviews were by telephone.33

The present study includes only respondents aged 55 or older at the start of the survey in 1994/1995 and is based on cycle 3 data (1998/1999).

Measures

Social support

The 19-item MOS social support survey measures four dimensions of functional social support.28 (The survey also contains a structural support item not included in this study.) Emotional/Informational support comprises eight items estimating the extent to which interpersonal relationships provide guidance, positive affect, and empathetic understanding. Tangible support comprises four items pertaining to material aid and behavioural assistance. Affectionate support comprises three items that measure expressions of love and affection. Positive social interaction comprises four items that relate to the availability of someone with whom to have fun. Questions are answered on a five-point scale ranging from “none of the time” to “all of the time,” with higher values indicating more support.

Language

NPHS respondents were interviewed in English or French. This variable was used to identify English- and French-speaking respondents for the present study. The vast majority who chose to be interviewed in French lived in the province of Quebec.

Analyses

The descriptive analyses for this study were generated using SAS software (Version 8).35 Because of its capacity to perform CFA of ordered categorical scales, Mplus (version 4.1) was chosen to conduct all other analyses.36,37

The internal consistency of the MOS social support scales was measured with Cronbach’s alpha using polychoric correlations, provided that the data were ordered categorically.38 The use of Cronbach’s alpha as an indication of internal consistency has been criticized,38,41 especially when data are not continuous. Therefore, the composite reliability for congeneric measures model (CRCMM) was also examined.40

Fitting ordered categorical variables to a model for continuous variables can distort the factor structure and fit of the model and affect comparisons between groups.42,44 Consequently, the CFA and test of invariance were conducted with procedures for ordered categorical data. A polychoric correlation matrix was analyzed using a weighted least squares estimator with a mean- and variance-adjusted chi-square (WLSMV).47,43 If the sample size is large enough, this
method performs well for categorical variables, even when modest violations of normality are reported.\textsuperscript{43,44} A two-step approach to testing MI was used.\textsuperscript{37} First, a baseline model where thresholds and factor loadings were free across both groups with reasonable fit to the data was established. Second, the consistency of the social support scale for English- and French-speaking respondents was tested by constraining all factor loadings and thresholds as being equal across both groups.

Survey sampling weights were used to avoid parameter estimate bias.\textsuperscript{33,45} To get a normalized weight, the weighted value was further divided by the average weight.

Because stratified and cluster sampling were used to collect the data, the bootstrap technique was used for the preliminary analyses to adjust for violation of the assumption of independence between observations.

In structural equation modeling (SEM), the use of complex survey design data may underestimate standard errors, and, in turn, affect chi-square values. The linearization (Taylor Series approximation) method with both the cluster and stratum identifiers was used. This method provides robust estimates of parameters and standard errors.\textsuperscript{46}

Use of approximate fit indices (AFIs) in SEM for models that fail the chi-square test rather than reporting the significance of chi-square has been criticized.\textsuperscript{47} However, given the large sample size in this study, it is likely that the chi-square test statistic will be significant, suggesting that the discrepancy between the observed and hypothesized model will be greater than would be expected by chance alone. Therefore, both the significance of the chi-square test and the AFIs are reported. The AFIs are not meant to provide support or lack thereof of a perfectly fitting model, but rather, information about whether the model is acceptable based on the approximate fit.\textsuperscript{48}

Similar to tests of overall model fit, chi-square difference tests used in MI are sensitive to sample size.\textsuperscript{49} Recent work suggests that AFIs in MI research are less sensitive to sample size than are chi-square statistics.\textsuperscript{49-51} Vandenberg and Lance recommend using change in comparative fit index (CFI) with a cutoff value of .02 to detect lack of invariance.\textsuperscript{52} The use of root mean square error of approximation (RMSEA) values is not recommended.\textsuperscript{49}

Unlike other estimation methods where degrees of freedom, and consequently, the chi-square statistic are based on the specification of the model, the degrees of freedom of the WLSMV method are adjusted depending on both sample size and model specification.\textsuperscript{37,43,53} Therefore, only the $p$ values, not the degrees of freedom or chi-square values, are interpretable. For this reason, degrees of freedom and chi-square values are not reported.

The following criteria were used to evaluate model fit: the $p$-values of chi-square; the Comparative Fit Index (CFI); the Tucker–Lewis index (TLI), and RMSEA. According to Hu and Bentler, CFI and TLI values greater than .95 indicate an acceptable fit.\textsuperscript{54} RMSEA values less than .05 represent a good fit, and values up to .08, a reasonable fit.\textsuperscript{54,55} Parameter estimates and standard error estimates are also examined.

**Results**

**Preliminary analysis**

A total of 4,444 NPHS cycle 3 (1998/1999) respondents were aged 55 or older in 1994/1995. The present analyses exclude 8 respondents with missing data on the language of interview variable. A further 525 cases were deleted because they had died, and 111 because they had moved to an institution.

Of the remaining respondents, 669 had at least one missing value on the MOS social support scale and were deleted. Respondents with missing values were similar to those without missing values in gender and language distribution, but they were older and had lower levels of education and income. Because the majority of them (530) were missing data on all MOS social support items, imputations would have been difficult.

**Table 1**

<table>
<thead>
<tr>
<th>Language of interview</th>
<th>Number</th>
<th>Percent$^1$</th>
<th>Number</th>
<th>Percent$^1$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>1,020</td>
<td>43.6</td>
<td>198</td>
<td>40.3</td>
</tr>
<tr>
<td>Women</td>
<td>1,622</td>
<td>56.5</td>
<td>291</td>
<td>59.7</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/Common-law/Living with partner</td>
<td>1,383</td>
<td>65.1</td>
<td>258</td>
<td>58.1</td>
</tr>
<tr>
<td>Single</td>
<td>160</td>
<td>4.4</td>
<td>37</td>
<td>6.5</td>
</tr>
<tr>
<td>Widowed/Separated/Divorced</td>
<td>1,099</td>
<td>30.5</td>
<td>194</td>
<td>35.4</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowest</td>
<td>131</td>
<td>3.7</td>
<td>23</td>
<td>4.4</td>
</tr>
<tr>
<td>Lower-middle</td>
<td>477</td>
<td>12.9</td>
<td>125</td>
<td>23.5</td>
</tr>
<tr>
<td>Middle income</td>
<td>945</td>
<td>25.4</td>
<td>198</td>
<td>41.8</td>
</tr>
<tr>
<td>Upper-middle</td>
<td>677</td>
<td>33.5</td>
<td>88</td>
<td>24.0</td>
</tr>
<tr>
<td>Highest</td>
<td>236</td>
<td>14.6</td>
<td>19</td>
<td>6.3</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than secondary graduation</td>
<td>1,195</td>
<td>39.6</td>
<td>301</td>
<td>59.4</td>
</tr>
<tr>
<td>Secondary graduation</td>
<td>328</td>
<td>13.0</td>
<td>55</td>
<td>11.8</td>
</tr>
<tr>
<td>Some postsecondary</td>
<td>539</td>
<td>21.7</td>
<td>56</td>
<td>11.9</td>
</tr>
<tr>
<td>Postsecondary graduation</td>
<td>578</td>
<td>25.8</td>
<td>77</td>
<td>16.9</td>
</tr>
</tbody>
</table>

$^1$ weighted estimates

* significant differences between English- and French-speaking respondents

Table 2

Means, standard deviations (S.D.), coefficient alpha estimates, and composite reliability estimates (CRCMM) for Medical Outcomes Study social support factors, by language of interview, household population aged 55 or older, Canada excluding territories, 1998/1999

<table>
<thead>
<tr>
<th>Scale</th>
<th>Language of interview</th>
<th></th>
<th></th>
<th>Language of interview</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>English (n = 2,642)</td>
<td>French (n = 489)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>alpha</td>
<td>CRCMM</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>Tangible support</td>
<td>13.54</td>
<td>3.47</td>
<td>0.93</td>
<td>0.93</td>
<td>12.69</td>
<td>4.89</td>
</tr>
<tr>
<td>Affectionate support</td>
<td>10.30</td>
<td>2.58</td>
<td>0.94</td>
<td>0.94</td>
<td>9.55</td>
<td>3.52</td>
</tr>
<tr>
<td>Positive social interaction</td>
<td>13.14</td>
<td>3.52</td>
<td>0.95</td>
<td>0.95</td>
<td>12.62</td>
<td>4.11</td>
</tr>
<tr>
<td>Emotional/Informational support</td>
<td>26.37</td>
<td>6.79</td>
<td>0.97</td>
<td>0.97</td>
<td>24.93</td>
<td>9.12</td>
</tr>
</tbody>
</table>

Note: Normalized survey sampling weights and Taylor linearization method were used.

An option would have been to add covariates to the CFA model and make “missingness” conditional on these, but this approach does not work well with cases that are missing all values on the indicator variables.

The analyses were conducted on the remaining 3,131 adults, who ranged in age from 58 to 99 (M = 69.72, SD = 7.91) (Table 1). English-speaking respondents had significantly higher levels of education and income than did those who were French-speaking. The latter were less likely than the former to be in a relationship. No differences emerged in gender and age.

Reliability
Cronbach’s alphas all exceeded .90 (Table 2). The composite reliability was based on the standardized loadings and standardized measurement error variances of the fully constrained model. These estimates were used given that the model appears to function similarly for both groups. All values exceeded .93.

Confirmatory factor analysis
A CFA with four factors was performed separately on the English- and French-speaking respondents (Table 3). For both groups, the model represented an adequate fit to the data. However, the RMSEA values for the English-speaking respondents were somewhat high. Investigation of the MI indices suggested cross-loading of item 5 (“someone to take you to the doctor if you needed it”) on the affectionate (modification index = 41.30, standardized expected parameter change = .31), positive social interaction (modification index = 41.16, standardized expected parameter change = .29), and emotional/informational (modification index = 51.76, standardized expected parameter change = .33) factors. It appears that item 5 measures not only tangible support, but also affectionate social support, positive social interaction, and emotional/instrumental social support.

For French-speaking respondents, all modification indices were low. After a number of factors were considered, the specification of the model was not changed. The RMSEA was only slightly high and is acceptable; the CFI and TLI values suggested an acceptable fit, and theoretically, it is difficult to explain why these items would be cross-loading on the other factors.

The items of the MOS social support scale all loaded significantly on their respective latent variable (Table 4, Figure 1). (Tables reporting the correlation matrix, including the means and standard deviations for each item from the model, are available from the authors.)

Test of invariance
The unconstrained multigroup model, in which the thresholds and the factor loadings are relaxed, represents an acceptable fit to the data. The second model, in which factor loadings and thresholds are constrained to be equal across both groups, also represents an acceptably fitting model (Table 3). Constraining the factor loadings and thresholds to be equal across both groups resulted in a significant chi-square difference test ($\Delta \chi^2[\Delta df= 21] = 50.84, p = .0003$). However, $\Delta$CFI for testing the invariance of factor loadings and thresholds was .008, suggesting that the weight of the factor loading and the thresholds were invariant across both

Table 3

Goodness-of-fit statistics for Medical Outcomes Study social support scale, household population aged 55 or older, Canada excluding territories, 1998/1999

<table>
<thead>
<tr>
<th>Model</th>
<th>Probability values</th>
<th>Comparative fit index (CFI)</th>
<th>Tucker-Lewis index (TLI)</th>
<th>Root mean square error of approximation (RMSEA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>English-speaking</td>
<td>&lt;0.0001</td>
<td>.96</td>
<td>.99</td>
<td>.076</td>
</tr>
<tr>
<td>French-speaking</td>
<td>&lt;0.0001</td>
<td>.96</td>
<td>.99</td>
<td>.047</td>
</tr>
<tr>
<td>French and English combined</td>
<td>&lt;0.0001</td>
<td>.95</td>
<td>.99</td>
<td>.084</td>
</tr>
<tr>
<td>Test for equality across language</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor loading and thresholds unconstrained</td>
<td>&lt;0.0001</td>
<td>.97</td>
<td>.99</td>
<td>.086</td>
</tr>
<tr>
<td>Factor loading and thresholds constrained</td>
<td>&lt;0.0001</td>
<td>.98</td>
<td>.10</td>
<td>.075</td>
</tr>
</tbody>
</table>

Note: Normalized survey sampling weights and Taylor linearization method were used.
models. Partial measurement invariance was examined by constraining each item one at a time as equal across both groups. No differences were found.

**Discussion**

The primary purpose of the study was to examine the psychometric properties of the English and French versions of the MOS social support scale for a sample of older adults. Overall, the findings are similar to earlier studies reporting good psychometric properties for the MOS social support scale.28,29 The high Cronbach’s alphas and CRCMM values for all subscales suggest good internal consistency.

A second objective was to test the hypothesis of a four-factor structure of the 19-item MOS social support scale. The results suggest that the four-factor model of functional social support is acceptable. While this is consistent with earlier studies,30 some items appeared to be cross-loading on more than one factor, which suggests that a better-fitting model could be obtained if some items were deleted. This aligns with Gjesfjeld and colleagues’ finding.29

A third objective was to examine the invariance of the model between English- and French-speaking older adults. Based on change in CFI, the instrument functions relatively uniformly across both groups. Furthermore, when each item was examined independently, no differences between individual items emerged. Nonetheless, because the chi-square difference test was significant, lack of measurement invariance cannot be ruled out.

The English and French versions of the MOS social support scale used in the NPHS appear to be good measures of older adults’ perception of the availability of social support, which is a predictor of healthy aging.57-61 However, the scale does not measure all dimensions of social support. For example, it does not include reassurance of worth, 27,62-65 nor does it indicate the source of support (relatives, friends, children). The addition of such items is an important area for future research.

The current findings apply only to people aged 55 or older with similar demographic characteristics. Moreover, although differences between those with and without missing values were significant, the former were excluded from the study because there was no valid way to impute their data. Results might have been slightly different had these people been included in the analyses.

As well, the assessment of MI between English- and French-speaking older

---

**Table 4**

Standardized and unstandardized estimates from final models, household population aged 55 or older, Canada excluding territories, 1998/1999

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Language of interview</th>
<th>English (n = 2,642)</th>
<th>French (n = 489)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard-</td>
<td>Unstandardized</td>
<td>Standard-</td>
</tr>
<tr>
<td></td>
<td>ized estimate</td>
<td>estimate (standard</td>
<td>ized estimate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>error)</td>
<td></td>
</tr>
<tr>
<td>Item 2 on tangible support</td>
<td>0.80†</td>
<td>1.00</td>
<td>0.83†</td>
</tr>
<tr>
<td>Item 3 on emotional/informational support</td>
<td>0.85†</td>
<td>1.00</td>
<td>0.82†</td>
</tr>
<tr>
<td>Item 4 on emotional/informational support</td>
<td>0.83</td>
<td>.97 (.012)</td>
<td>0.80</td>
</tr>
<tr>
<td>Item 5 on tangible support</td>
<td>0.85</td>
<td>1.06 (.024)</td>
<td>0.89</td>
</tr>
<tr>
<td>Item 6 on affectionate support</td>
<td>0.90†</td>
<td>1.00</td>
<td>0.86†</td>
</tr>
<tr>
<td>Item 7 on positive social interaction</td>
<td>0.89†</td>
<td>1.00</td>
<td>0.87†</td>
</tr>
<tr>
<td>Item 8 on emotional/informational support</td>
<td>0.88</td>
<td>1.03 (.012)</td>
<td>0.87</td>
</tr>
<tr>
<td>Item 9 on emotional/informational support</td>
<td>0.91</td>
<td>1.07 (.012)</td>
<td>0.90</td>
</tr>
<tr>
<td>Item 10 on affectionate support</td>
<td>0.90</td>
<td>1.00 (.014)</td>
<td>0.90</td>
</tr>
<tr>
<td>Item 11 on positive social interaction</td>
<td>0.90</td>
<td>1.01 (.009)</td>
<td>0.88</td>
</tr>
<tr>
<td>Item 12 on tangible support</td>
<td>0.92</td>
<td>1.14 (.020)</td>
<td>0.91</td>
</tr>
<tr>
<td>Item 13 on emotional/informational support</td>
<td>0.87</td>
<td>1.02 (.012)</td>
<td>0.89</td>
</tr>
<tr>
<td>Item 14 on positive social interaction</td>
<td>0.93</td>
<td>1.05 (.009)</td>
<td>0.83</td>
</tr>
<tr>
<td>Item 15 on tangible support</td>
<td>0.95</td>
<td>1.18 (.021)</td>
<td>0.90</td>
</tr>
<tr>
<td>Item 16 on emotional/informational support</td>
<td>0.94</td>
<td>1.11 (.013)</td>
<td>0.91</td>
</tr>
<tr>
<td>Item 17 on emotional/informational support</td>
<td>0.94</td>
<td>1.10 (.012)</td>
<td>0.92</td>
</tr>
<tr>
<td>Item 18 on positive social interaction</td>
<td>0.95</td>
<td>1.06 (.008)</td>
<td>0.91</td>
</tr>
<tr>
<td>Item 19 on emotional/informational support</td>
<td>0.94</td>
<td>1.10 (.013)</td>
<td>0.88</td>
</tr>
<tr>
<td>Item 20 on affectionate support</td>
<td>0.94</td>
<td>1.04 (.013)</td>
<td>0.85</td>
</tr>
</tbody>
</table>

† Fixed parameter

Note: Normalized survey sampling weights and Taylor linearization method were used.

adulthood is valuable to researchers exploring social support in these populations. In the past, the two groups had been combined or compared although MI between them had not been evaluated. The English- and French-speaking Canadians in the current study do not appear to differ in their interpretation of the meaning of the MOS social support scale items. This suggests that cultural differences between the two groups likely played an important role in differences in levels of social support found by previous studies using the MOS social support scale.

**Acknowledgements**

The authors are grateful to Rochelle Garner for all the help she provided with the paper. This study was supported by the Statistics Canada Tom Symons Research Stipend Program to the first author.
References


47. Vernon T, Eysenck SBG. Special issue on structural equation modeling [Special Issue]. *Personality and Individual Differences* 2007; 42(5).


