Tourism Destination Marketing
Gaining the Competitive Edge

Edited by
Joseph Ruddy
*Dublin Institute of Technology*
and
Sheila Flanagan
*Dublin Institute of Technology*
LEISURE MOTIVES AS PREDICTORS OF ACTIVITIES — THE LILLEHAMMER SCALES IN A NATIONAL SURVEY

JO KLEIVEN¹

Abstract:
Following their development in surveys of two Norwegian local populations, the eight 'Lillehammer scales' of travel and leisure motivation were used in a nationally representative survey. Nationally valid 'standards' for the scales were established, facilitating the use of scales in later non-representative user group or guest surveys.

Based on a sample of more than 1300 respondents, scales again appeared to have acceptable reliability, as measured by Cronbach's alpha. Confirmatory factor analyses further indicate that the measurement model for each scale fits the data reasonably well. The complete model should probably allow for inter-scale covariation.

Our first attempts to assess the predictive validity of the eight scales yielded quite promising results. SEM models suggest, however, that the predictive effects of some scales on certain leisure behaviors are dependent on the state of other predictor variables (i.e. interaction effects).

Keywords: Leisure, motives, scales, Structural Equations Modeling, prediction

¹ The author is a graduate of Social Psychology from the University of Oslo, now working as a senior lecturer with the Travel & Tourism Unit of Lillehammer College, Norway. His project, Motive dimensions as demand predictors in Norwegian Leisure and Tourism, is supported by grants from the Tourism Research Program of the Norwegian Research Council.
INTRODUCTION

Several authors have tried to measure salient dimensions of leisure and tourism motivation. In the field of tourism, the work of Crompton (1979), Crandall (1977; 1980), and Beard & Ragheb (1983) is well known internationally; while Tinsley (1984; 1981) and Driver (1977; 1987; 1991) are often cited in the leisure literature.

The Beard & Ragheb (1983) scales and Driver's (1991) 'Paragraphs about leisure' have also been used by other authors (Loundsbury & Hoopes, 1988; Loundsbury & Franz, 1990; Ryan, 1993; Ryan, 1994a; Ryan, 1994b; Driver et al., 1991). No general consensus seems to have emerged, however, as to which dimensions or scales are the most important. Nevertheless, two common insights are worth mentioning. Firstly, it has often been pointed out that general psychological theories of motivation may be of limited value to understanding the specific travel and leisure motives (Dann, 1983; Dann, 1981; Pearce, 1993; 1983; 1981). Secondly, multi-motive models are needed, acknowledging the simultaneous influence of more than one motive (Pearce, 1993; Krippendorf, 1987; Schmidhauser, 1989; Witt & Wright, 1992).


Inspired by this research, I have been working to develop a series of scales to fit the realities of Norwegian leisure and tourism, specifically. While some dimensions may be common to most people everywhere, I did suspect that Norwegians' needs and motives for leisure may be somewhat different from those of people living in warmer or more urban parts of the world (Kleiven, 1998b). In two representative surveys of the towns of Gjøvik and Sandefjord, nine motive dimensions were identified (Kleiven, 1999; Kleiven, 1998b). Nine summated scales of four items were developed for measuring these motives. Most scales had acceptable reliability and validity.

There was also some room for improvement, however. One scale (Indulgence) was excluded from further work, due to low reliability and very high inter-correlation with other scales. The eight remaining motive scales were: Sun/warmth, Accomplishment, Family, Friends, Culture, Peace/Quiet, Nature, and Fitness. The items of each scale are shown in table 1. Items 2d, 5d, and 6d are new in the present survey, replacing items that had given minor problems with earlier versions of the scales.

In both previous surveys, the score on most scales varied substantially with demographic differences. Clearly, representative samples were required to avoid misleading results. For quick visitor surveys and other surveys with convenience samples, therefore, we saw the need for scale standards. Comparing the results of non-representative surveys with such standards, they will be easier to interpret correctly. It was decided, therefore, to standardize the scales with data from a representative survey of the Norwegian population.
### Table 1. Motive scales, Cronbach *alphas* and associated items

1. **SUN/WARMTH (SOL/VARME)** Gjøvik *alpha* = .69; Sandefjord = .76  
   a) Feel the heat of the sun (*Føle varme fra solen*)  
   b) Enjoy beach and swimming (*Nyte strand og badeliv*)  
   c) Get a tan (*Bli brun*)  
   d) Swim in clean water (*Bade i rent vann*)

2. **ACCOMPLISHMENT (MESTRING)** Gjøvik *alpha* = .76; Sandefjord = .72  
   a) Using skill and knowledge (*Få brukt ferdigheter og kunnskaper*)  
   b) Developing personal interest/hobby (*Videreutvikle personlig interesse/hobby*)  
   c) Learning something new (*Lære noe nytt*)  
   d) Being with people with same interests (*Omgås folk med samme interesser som deg selv*)

3. **FAMILY (FAMILIE)** Gjøvik *alpha* = .71; Sandefjord = .73  
   a) Having time for the family (*Ha tid til familien*)  
   b) Keeping in touch with family living elsewhere (*Beholde kontakt med familie som bor andre steder*)
   c) Being with children of my relatives (*Være sammen med barn i slekten*)  
   d) See to it that the children have a pleasant vacation (*Sørge for at barna har det bra i ferien*)

4. **FRIENDS (VENNER)** Gjøvik *alpha* = .59; Sandefjord = .60  
   a) Keeping in touch with friends (*Beholde kontakten med venner*)  
   b) Getting to know new people (*Bli kjent med nye mennesker*)  
   c) Eat and drink in good company (*Spise og drikke i godt lag*)  
   d) Not being lonely during the vacation (*Ikke være ensom i ferien*)

5. **CULTURE (KULTUR)** Gjøvik *alpha* = .72; Sandefjord = .75  
   a) Experience art and culture (*Oppleve kunst og kultur*)  
   b) Seeing well-known places or sights (*Oppleve kjente steder eller severdigheter*)  
   c) Getting to know other countries and cultures (*Bli kjent med andre land og kulturer*)  
   d) Satisfying an interest in history (*Tilfredsstille historisk interesse*)

6. **NATURE (NATUR)** Gjøvik *alpha* = .70; Sandefjord = .68  
   a) Experience landscape and nature (*Oppleve landskap og natur*)  
   b) Feeling you belong in nature (*Føle tilhørighet til naturen*)  
   c) Experience the silence of nature (*Oppleve stillheten i naturen*)  
   d) See and experience Norway (*Se og oppleve Norge*)

7. **PEACE/QUIET (FRED/RO)** Gjøvik *alpha* = .64; Sandefjord = .63  
   a) Getting away from push and stress (*Komme bort fra mas og stress*)  
   b) Getting away from noise and pollution (*Komme vekk fra støy og forurensning*)  
   c) Recovering strength (*Hente nye krefter*)  
   d) Avoid the push and stress of travelling (*Unngå mas og stress ved det å reise*)

8. **FITNESS (TRIM)** Gjøvik *alpha* = .81; Sandefjord = .74  
   a) Getting a workout, exercising (*Få trim eller mosjon*)  
   b) Working out, really tiring your body (*Ta deg ut og bli skikkelig sliten i kroppen*)  
   c) Taking care of your health (*Ta vare på egen helse*)  
   d) Getting in shape (*Komme i form*)

The main research questions were:
1. Will the revised scales have acceptable reliability?  
2. What are the test scores for the general Norwegian population?  
3. Will the factor structure be replicated?  
4. Will the scales have predictive validity?
METHOD

The scales were included in an omnibus survey conducted by Statistics Norway (Teigum, 1997). Within our block of questions, items were distributed in an approximately random fashion, keeping same-scale items separated by at least three or four items from other scales. The question given was *Thinking of your leisure and vacation last summer, how important was ...*(item). The range of the response scale ran from 1 (Very important) to 5 (Not important). To make results more easily comprehensible, however, the scale was reversed (i.e. 1 = Not important, 5 = Very important) before the data processing.

Personal interviews were planned for a representative sample of 2000, for a period of five weeks in the autumn of 1996. Due to procedural problems, Statistics Norway had to extend the interviewing period to seven weeks. The final size of the sample was only 1334, (67% of the planned sample), mainly due to refusals. Younger males (ages 16-24) were slightly overrepresented, while older females (ages 67-79) were somewhat underrepresented. Statistics Norway therefore recommended that weighting be considered for variables related to age and sex (Teigum, 1997). In this preliminary account of the results, however, no weighting has been applied.

In addition to the data from our scales, we also had access to extensive data on sociodemographics and on nature-based leisure behavior (Vorkinn et al., 1996).

RESULTS

Reliability
As shown in table 2, most of the revised scales generally have acceptable reliability. The FRIENDS scale, however, only yields an alpha of .62. While this is a lower figure than what I would like to see, it is consistent with previous results. The remaining scale alphas range from .70 to .81. For four-item scales, I view this as satisfactory (DeVellis, 1991; Nunnally & Bernstein, 1994).

<table>
<thead>
<tr>
<th>Scale</th>
<th>Cronbach alpha</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>S.E.of mean</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUN/WARMTH</td>
<td>.71</td>
<td>1327</td>
<td>3.28</td>
<td>0.86</td>
<td>0.02</td>
<td>-0.36</td>
<td>-0.14</td>
</tr>
<tr>
<td>ACCOMPLISHM.</td>
<td>.73</td>
<td>1327</td>
<td>3.10</td>
<td>0.87</td>
<td>0.02</td>
<td>-0.26</td>
<td>-0.31</td>
</tr>
<tr>
<td>FAMILY</td>
<td>.72</td>
<td>1327</td>
<td>3.74</td>
<td>0.92</td>
<td>0.03</td>
<td>-0.80</td>
<td>0.09</td>
</tr>
<tr>
<td>FRIENDS</td>
<td>.62</td>
<td>1328</td>
<td>3.64</td>
<td>0.76</td>
<td>0.02</td>
<td>-0.72</td>
<td>0.56</td>
</tr>
<tr>
<td>CULTURE</td>
<td>.73</td>
<td>1327</td>
<td>2.72</td>
<td>0.88</td>
<td>0.02</td>
<td>0.11</td>
<td>-0.43</td>
</tr>
<tr>
<td>NATURE</td>
<td>.81</td>
<td>1327</td>
<td>3.59</td>
<td>0.89</td>
<td>0.02</td>
<td>-0.66</td>
<td>0.01</td>
</tr>
<tr>
<td>PEACE/QUIET</td>
<td>.70</td>
<td>1327</td>
<td>3.77</td>
<td>0.82</td>
<td>0.02</td>
<td>-0.85</td>
<td>0.67</td>
</tr>
<tr>
<td>FITNESS</td>
<td>.79</td>
<td>1327</td>
<td>3.35</td>
<td>0.90</td>
<td>0.02</td>
<td>-0.29</td>
<td>-0.43</td>
</tr>
</tbody>
</table>

Table 2. Statistics of eight scales
Scale statistics
The scale means range from 2.72 to 3.77, suggesting that neither ceiling nor floor effects are likely. Skewness and kurtosis figures do not indicate normal distributions. It may perhaps also be noted that in the Norwegian sample, 'culture' and 'accomplishment' appears to be less important than, e.g., 'family', 'friends', or 'peace/quiet'.

Gender differences
Table 3 indicates that for all scales but one, ('Accomplishment'), scale means of males are significantly higher than those of females. In terms of the five-point scale, however, the differences may be too small to warrant different scale norms for men and women.

Table 3. Scale means for males and females

<table>
<thead>
<tr>
<th>Scale</th>
<th>Males</th>
<th>Females</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUN/WARMTH</td>
<td>3.17</td>
<td>3.38</td>
<td>-4.246</td>
<td>1325</td>
<td>0.000</td>
</tr>
<tr>
<td>ACCOMPLISHM.</td>
<td>3.11</td>
<td>3.10</td>
<td>0.184</td>
<td>1325</td>
<td>0.854</td>
</tr>
<tr>
<td>FAMILY</td>
<td>3.58</td>
<td>3.92</td>
<td>-6.811</td>
<td>1325</td>
<td>0.000</td>
</tr>
<tr>
<td>FRIENDS</td>
<td>3.58</td>
<td>3.70</td>
<td>-2.933</td>
<td>1326</td>
<td>0.003</td>
</tr>
<tr>
<td>CULTURE</td>
<td>2.67</td>
<td>2.78</td>
<td>-2.254</td>
<td>1325</td>
<td>0.024</td>
</tr>
<tr>
<td>NATURE</td>
<td>3.51</td>
<td>3.67</td>
<td>-3.246</td>
<td>1325</td>
<td>0.001</td>
</tr>
<tr>
<td>PEACE/QUIET</td>
<td>3.72</td>
<td>3.82</td>
<td>-2.205</td>
<td>1325</td>
<td>0.028</td>
</tr>
<tr>
<td>FITNESS</td>
<td>3.28</td>
<td>3.42</td>
<td>-2.737</td>
<td>1325</td>
<td>0.006</td>
</tr>
</tbody>
</table>

It should perhaps be mentioned that numerous other demographic differences have also been found in the scale data, although it is beyond the scope of the present paper.

Factor structure
As shown in table 4, all scale intercorrelations are positive and rather high. They are also significant at .001 or more. Clearly, a measurement model with uncorrelated factors is not likely to fit our data.

Table 4. Scale intercorrelations

<table>
<thead>
<tr>
<th>Scale</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SUN/W.</td>
<td>1.00</td>
<td>.30</td>
<td>.19</td>
<td>.37</td>
<td>.29</td>
<td>.22</td>
<td>.37</td>
<td>.32</td>
</tr>
<tr>
<td>2. ACCOMPL.</td>
<td>1.00</td>
<td>.24</td>
<td>.48</td>
<td>.47</td>
<td>.45</td>
<td>.39</td>
<td>.59</td>
<td></td>
</tr>
<tr>
<td>3. FAMILY</td>
<td>1.00</td>
<td>.23</td>
<td>.22</td>
<td>.42</td>
<td>.36</td>
<td>.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. FRIENDS</td>
<td>1.00</td>
<td>.30</td>
<td>.28</td>
<td>.31</td>
<td>.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. CULTURE</td>
<td>1.00</td>
<td>.40</td>
<td>.26</td>
<td>.33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. NATURE</td>
<td>1.00</td>
<td>.52</td>
<td>.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. PEACE/Q.</td>
<td>1.00</td>
<td>.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. FITNESS</td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
We shall first look at confirmatory factor analyses of the eight scales separately, however. Using LISREL 8.30, and using our data as continuous scales\(^2\), we get quite similar results for all the scales. Following Hoyle & Panter's (1995) recommendations, both 'Stand-alone', 'Type-2' and 'Type-3' indexes are displayed in table 5.

Chi-square values are generally quite high, indicating no good fit between data and model on six scales. Other measures of fit, however, are very encouraging. General Fit Index, Incremental Fit Index and Comparative Fit Index all are at satisfactory levels. Viewed separately, I therefore feel that all eight scales are supported by the CFA.

Table 5. Measures of fit from LISREL CFA of scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>N</th>
<th>Chi-sq.</th>
<th>df</th>
<th>p</th>
<th>GFI</th>
<th>IFI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SUN/W.</td>
<td>1326</td>
<td>57.79</td>
<td>2</td>
<td>0.0000</td>
<td>0.98</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>2. ACCOMPL.</td>
<td>1324</td>
<td>0.58</td>
<td>2</td>
<td>0.7496</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3. FAMILY</td>
<td>1293</td>
<td>117.08</td>
<td>2</td>
<td>0.0000</td>
<td>0.96</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>4. FRIENDS</td>
<td>1323</td>
<td>7.65</td>
<td>2</td>
<td>0.0219</td>
<td>1.00</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>5. CULTURE</td>
<td>1325</td>
<td>24.05</td>
<td>2</td>
<td>0.0000</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>6. NATURE</td>
<td>1327</td>
<td>9.72</td>
<td>2</td>
<td>0.0077</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7. PEACE/Q.</td>
<td>1323</td>
<td>4.67</td>
<td>2</td>
<td>0.0970</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>8. FITNESS</td>
<td>1326</td>
<td>31.15</td>
<td>2</td>
<td>0.0000</td>
<td>0.99</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>

As expected, putting the eight models into a common model with eight uncorrelated factors is not supported. Chi-square then is very high (7864.85; df = 464; p < .0001), and other fit indices are very low (GFI = 0.72; IFI = .63; CFI = .63). Furthermore, LISREL's modification index suggests that factors be allowed to correlate.

Turning the eight models into a model with eight correlated factors, then, the fit is somewhat better. Chi-square is still high (2874.07; df = 436; p < .0001), and other fit indices are still below .90 (GFI = 0.88; IFI = .84; CFI = .84). Although much improved, the fit between data and this model is also not satisfactory.

To achieve an even better fit, however, only ten minor modifications are needed. Figure 1 on the next page shows an example of possible adjustments, resulting in an acceptable model fit (Chi-square = 2079 (df = 426; p < .0001), GFI = 0.91; IFI = .90; CFI = .90).

Looking closely at the adjusted model, it does not seem to invalidate the original. Seven items now load on two factors instead of only one. In most cases, this appears quite meaningful and consistent with the meaning or content of the eight dimensions. Also, the adjusted model allows for three correlated error terms. In all cases, however, the correlations occur between the error terms of items within the same scale.

\(^2\) Most authors would argue that ordinal scales would be more appropriate. The scales for our items have only five values, and do not meet assumptions of normal distribution.
Figure 1. Modified eight-factor measurement model: Standardized loadings.
For readability reasons, LISREL \textit{Phi} values (scale intercorrelations) are not shown in the figure, but will appear in table 6. As we see, most scales are still highly correlated, and the general pattern is rather similar to that of table 4. It may be worth noting, however, that ACCOMPLISHMENT and FITNESS now are very highly correlated, as are NATURE and PEACE/QUIET.

Table 6. LISREL \textit{Phi} values (Scale intercorrelations).

<table>
<thead>
<tr>
<th>Scale</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SUN/W.</td>
<td>1.00</td>
<td>.27</td>
<td>.14</td>
<td>.50</td>
<td>.22</td>
<td>.08</td>
<td>.35</td>
<td>.31</td>
</tr>
<tr>
<td>2. ACCOMPL.</td>
<td></td>
<td>1.00</td>
<td>.25</td>
<td>.59</td>
<td>.53</td>
<td>.54</td>
<td>.53</td>
<td>.73</td>
</tr>
<tr>
<td>3. FAMILY</td>
<td></td>
<td></td>
<td>1.00</td>
<td>.27</td>
<td>.31</td>
<td>.50</td>
<td>.52</td>
<td>.37</td>
</tr>
<tr>
<td>4. FRIENDS</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.40</td>
<td>.36</td>
<td>.46</td>
<td>.55</td>
</tr>
<tr>
<td>5. CULTURE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.43</td>
<td>.36</td>
<td>.41</td>
</tr>
<tr>
<td>6. NATURE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.75</td>
<td>.68</td>
</tr>
<tr>
<td>7. PEACE/Q.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>.62</td>
</tr>
<tr>
<td>8. FITNESS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

Still, the scales appear to have both divergent and convergent validity. Scales do not collapse into a smaller number of scales, and only seven out of 32 items load on more than one factor.

Predictive validity
Since the survey has data on subjects' participation in several outdoor recreation activities, this may be used for a preliminary assessment of the predictive validity of our eight scales. Since the scales are correlated, we turn to multiple regression for this first analysis. In table 7, some sample results are shown.

Table 7. Multiple regression, beta weights for eight scales predicting activities.

<table>
<thead>
<tr>
<th>Scale/Activity</th>
<th>Jogging</th>
<th>Alpine skiing</th>
<th>Sea fishing</th>
<th>Outdoor swimming</th>
<th>Short walk</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. SUN/W.</td>
<td>-.028</td>
<td>.093**</td>
<td>.031</td>
<td>.262***</td>
<td>-.004</td>
</tr>
<tr>
<td>2. ACCOMPL.</td>
<td>-.009</td>
<td>.059</td>
<td>.122**</td>
<td>.029</td>
<td>-.054</td>
</tr>
<tr>
<td>3. FAMILY</td>
<td>-.156***</td>
<td>-.110***</td>
<td>-.019</td>
<td>-.067*</td>
<td>.013</td>
</tr>
<tr>
<td>4. FRIENDS</td>
<td>.025</td>
<td>.091**</td>
<td>-.027</td>
<td>-.007</td>
<td>-.035</td>
</tr>
<tr>
<td>5. CULTURE</td>
<td>.009</td>
<td>-.050</td>
<td>-.098**</td>
<td>.002</td>
<td>-.065*</td>
</tr>
<tr>
<td>6. NATURE</td>
<td>-.032</td>
<td>-.145***</td>
<td>.076*</td>
<td>-.017</td>
<td>.189***</td>
</tr>
<tr>
<td>7. PEACE/Q.</td>
<td>.013</td>
<td>-.009</td>
<td>.091**</td>
<td>-.048</td>
<td>.036</td>
</tr>
<tr>
<td>8. FITNESS</td>
<td>.249***</td>
<td>.059</td>
<td>-.062</td>
<td>-.007</td>
<td>.081*</td>
</tr>
</tbody>
</table>

\[ R^2 \] = .059 .051 .028 .064 .051

= p < .05; ** = p < .01; *** = p < .001
For all activities, two or more scales are significant predictors. It is also worth noting that different scales are relevant for different types of activity, and that even highly correlated scales make different predictions. We also see that negative predictors exist: FAMILY and CULTURE motives seem to be negatively related to some outdoor activities. At any rate, the scales do have some predictive power.

As $R^2$ values indicate, however, not much of the variance is explained. Clearly, there must be other influences on these behaviors. Some additional variables may certainly be simply included in the regression models, allowing us to gauge the effect of our motive variables within a larger context.

I have come to believe, however, that even more complex effect models will be needed to do this in a satisfactory manner. An example will hopefully support my assertion. As shown in table 7, scales' ability to predict Sea fishing was not very strong. This is certainly confirmed in figure 2, where the three 'best' scales have been put into a predictive model for women's participation in this activity.

![Diagram](image)

**Figure 2.** ACCOMPLISHMENT, PEACE/QUIET and CULTURE SCALES as predictors of women's sea fishing. Standardized estimates.

While the *structure* of the model fits our data in a mediocre way (Chi-square = 507.37; df = 60; p < .001; GFI = .89; IFI = .77; CFI = .77), the low coefficients between the factors and 'Fishing' show that it has a very limited predictive ability.

For men, however, the picture is a bit different, as shown in figure 3. The structure of the same model fits the data slightly better in this case (Chi-square = 439.93; df = 60; p < .001; GFI = .91; IFI = .82; CFI = .81). But more importantly, coefficients between factors and 'Fishing' are substantially higher, indicating a fairly useful predictive power.
While not generally very strong, this prediction model looks more appropriate for men than for women. There seems to be an interaction effect between gender and important properties of the model. When testing the predictive validity of scales, therefore, it is essential to distinguish between men and women.

This problem is not unique to the gender variable, however; similar effects may be shown with, e.g. groups in different phases of life, or with different types of education. The predictive validity of our scales, therefore, should be assessed within more complex models than multiple regressions.

**DISCUSSION**

Apparently, the eight revised scales have several desirable properties. Scales generally have an acceptable reliability. The factor structure known from previous studies is replicated in a satisfactory manner, even if high scale intercorrelations are found. Further, there seems to be no reason to expect ceiling or floor problems. The scale results from the national survey, therefore, may probably be useful standards for further work with other Norwegian samples.

Assessing the predictive validity of scales, however, seems to require more work. Multiple regression models do yield encouraging results, showing that scales have some ability to predict certain outdoor recreation activities. It is likely, however, that the predictive validity of the scales should better be assessed within more complex effect models.
REFERENCES


