II

Eight Scales for Leisure Travel Research

**Replicating and Revising
the Lillehammer Scales**

Jo Kleiven

Høgskolen i Lillehammer

Kr. 30,-

Eight Scales for Leisure Travel Research —

Replicating and Revising the Lillehammer Scales

by

Jo Kleiven

Faculty of Health and Social Work, Lillehammer University College

Abstract

Building on previous research in Norway and abroad, Kleiven (2005) developed *'The Lillehammer scales'* of leisure motives to measure nine motive dimensions relevant to Norwegian leisure and vacation choices.

The present study replicates Kleiven's original study, through a representative travel and leisure survey of the Norwegian township of Sandefjord. Eight of the nine motive factors were upheld in a confirmatory factor analysis: The dimensions of Sun/warmth, Mastery, Children/Family, Friends, Culture, Nature, Relaxation and Physical exercise all appear to replicate in a satisfactory manner.

Potential new items had largely negligible effects. Support for a ninth factor (Indulgence/Luxury) was considerable weaker. Cronbach alphas for the nine scales yield comparable results.

Consistent with the previous study, the scales have acceptable validity in terms of predicting several types of vacation behavior, especially in SEM models also including demographic variables. It is argued that the scales may prove useful for a variety of segmentation and marketing purposes, and that a full-scale national standardization of the scales should be undertaken.

Key words: Leisure motives, measurement, scale construction

Sammendrag

Med grunnlag i tidligere norsk og utenlandsk forskning utviklet Kleiven (2005) "Lillehammerskalaene" for ferie- og fritidsmotiver for å måle ni motivdimensjoner som var aktuelle for nordmenns valg av ferie- og fritids-muligheter.

Denne studien etterprøver Kleivens opprinnelige arbeid, i en representativ ferie- og fritidsundersøkelse i Sandefjord kommune. Åtte av de ni motivfaktorene fant støtte i en bekreftende faktoranalyse: Både Sol/varme, Mestring, Barn/Familie, Venner, Kultur, Natur, Fred og ro og Trim synes å være replikert på en tilfredsstilende måte i den nye studien.

Noen mulige nye skaleledd viste seg å gjøre svært liten forskjell. Støtte for en niende faktor (/Luksus) var betydelig svakere enn for de øvrige åtte. Cronbach's alfa-verdier for de ni skalaene var i samme størrelses-orden som i den originale studien.

Det er også i samsvar med den tidligere studien når skalaene viser brukbar validitet for prediksjon av flere former for ferieatferd, særlig i SEM-modeller som også inkluderer demografiske variable. Det synes klart at skalaene kan være nyttige for flere ulike segmenterings- og markedsføringsoppgaver, og at en fullskala nasjonal standardisering av skalaene burde gjennomføres.

Emneord: Fritidsmotiver, måling, skalakonstruksjon

The author: Senior Lecturer Jo Kleiven has been working with vacation and leisure motives over a period of several years. Other interests are Social/Cognitive and Environmental psychology and Organizational Behavior. He is presently attached to the Psychology Unit of the College.

Preface

This report is based on a Travel and Tourism survey of the Norwegian township of Sandefjord. 24 students carried out some of the planning and most of the interviews, since project participation was required in a course in Social Science methods at the Tourism Studies of Lillehammer University College.

The author highly appreciated the pleasant company of the students during a busy (and rainy) two-week period in Sandefjord. The friendly, cheerful and competent efforts of the student majority were truly instrumental to the modest success of the survey. Christer Thrane also made welcome contributions at several stages of the survey process.

At the time of the survey there were no clear plans of including it in a larger context. By and by, however, the author's interest in Vacation and Leisure motives developed into a comprehensive research project, where the present data could achieve even more meaning.

This research has been supported by the *Tourism Research Program* of the *Norwegian Research Council*, through a grant to the project *Motive dimensions as demand predictors in Norwegian Leisure and Tourism*.

Lillehammer, November 2006.

Jo Kleiven

Content:

1. Introduction

Over the years, the question of identifying salient motives for leisure and travel behavior has been addressed by several authors. Early writers like Crompton (1979), Crandall (1980) and Tinsley (1984) reveal no readily apparent consensus in the matter, and widely different theories, approaches and measurements have been used . More recently, attempts have been made to integrate different approaches to the motive problem (Harrill & Potts, 2002; Jamal & Lee, 2003). But still, commonly accepted procedures of motive measurements are hard to find.

It may be argued, however, that certain motives do occur on *several* motive lists, although under different names (Ryan, 1997; Kleiven, 1998b). Among such motives, Beard & Ragheb's (1983) motive dimensions of *"Intellectual"*, *"Social"*, *"Mastery/Competence"* and *"Stimulus Avoidance"* are often cited. The four dimensions have been replicated in the U.S (Lounsbury & Franz, 1990) and in Britain (Ryan, 1993; Ryan, 1994a), and their usefulness in tourism research seems well established (Ryan & Glendon, 1998).

In a recent article, Kleiven (2005) showed that the four Beard & Ragheb (1983) motive dimensions also replicate in a Norwegian sample, using the labels of *"Culture"*, *"Friends"*, *"Accomplishment"* and *"Peace/Quiet"*. The motives of *"Sun/warmth"*, *"Family"*, *"Nature"* and *"Fitness"* may also be added, without altering the factor structure of the original four motives. All eight dimensions were reliably measured through four-item summed scales, and the scales are likely to have some predictive validity for leisure and travel research. Although some of the scales were correlated, CFA seemed not to indicate serious problems with convergent or divergent validity of an eight-factor measurement model.

Adding a ninth dimension (*"Indulgence"*), however, yielded rather unpromising results. For this scale, reliability was low, it explained less than ¼ of the variance, and it was virtually indistinguishable from the *"Peace/Quiet"* scale (r = .93). A measurement model combining *"Indulgence"* and the other eight factors was not supported by the data.

The eight-factor motive measurement model is interesting for several reasons. Firstly, the model is based on the well-known "Leisure Motivation" scales (Beard & Ragheb, 1983), allowing comparisons with previous research. The extended model should be able to cope with greater motivational complexity, however. Secondly, the eight scales appear to have acceptable psychometric properties. Thirdly, the test items may not only be useful in conventional summed-item scales, but also as observable variables in Structural Equation Models, viewing motive factors as latent variables.

Findings of this nature, however, may be dependent on the particular sample studied. There is always a risk of capitalizing on chance relations in the data that do not exist outside that specific data set. It is recommended in the SEM literature, therefore, that interesting findings be replicated in a new sample (MacCallum, Roznowski, & Necowitz, 1992; MacCallum, 1995; Arbuckle, 1997).

Another reason for replicating Kleiven's (2005) findings is the possibility of improving the scales. Three items in his 36-item motive inventory did not to contribute very much towards the measurement of the latent motive factors, as shown by both reliability analyses and CFA. Replacing these items may yield more reliable scales.

It is also worth noting that two of three the "weak" items (*"Feeling the smell of the salty sea"* and *"Swim in clean water"*) are clearly related to the sea. In the inland community sample employed, these items may have caused problems. The author therefore recommends that "...both replacing the items and replicating the study in a coastal town should be considered" (Kleiven, 2005).

Following this suggestion, the present study was designed to replicate the inland study in a coastal town. Thus, the central research question is whether the scales will replicate in a new setting or not. Five subordinate research questions were:

1. May a similar factor structure be shown, with both divergent and convergent validity?
2. Are scale means and reliabilities comparable?
3. Will scales have predictive validity – as independent predictors or as part of more complex models?
4. May the scales be improved by replacing certain items?
5. Is there, e.g., an inland/coast problem with two items?

To improve comparability, a representative sample was used, as was most motive inventory items and the response format from the Kleiven (2005) study.

2. Method

An omnibus survey was planned and carried out by the author and a colleague, to investigate the "Vacation Habits" of the population in the town of Sandefjord at the South coast of Norway. Most of the practical data collection was done by a group of Travel and Tourism students at Lillehammer College, who had gone through a social science methods course and brief interviewer training.

A 1.3% intended sample was drawn from the Sandefjord census, controlling for gender, age group and electoral district within the town. Individuals included in the sample were approached in their private homes by one of our students, who asked for an interview. The sample thus obtained, however, was only about 0.9% of the population. High refusal rates resulted in 261 personal interviews, where women and young people (age 18-24) turned out to be slightly underrepresented. While cautioning against analyses based on small subgroups, the preliminary report (Kleiven & Thrane, 1996) nevertheless views the data set as a whole as satisfactory.

Table 1; Motive items in the Sandefjord survey (* = items not used by Kleiven (2005))

1	Feel the heat of the sun *(Føle varme fra solen)*	24	Taking care of your health *(Ta vare på egen helse)*
2	* Find your roots at home *(Søke tilbake til egne røtter på hjemstedet)*	25	Get a tan *(Bli brun)*
3	Keeping in touch with friends *(Beholde kontakten med venner)*	26	Developing personal interest/hobby *(Videreutvikle personlig interesse/hobby)*
4	Experience landscape and nature *(Oppleve landskap og natur)*	27	Being with children of my relatives *(Være sammen med barn i slekten)*
5	Getting away from push and stress *(Komme bort fra mas og stress)*	28	Not being lonely during the vacation *(Ikke være ensom i ferien)*
6	Getting a workout, exercising *(Få trim eller mosjon)*	29	Experience the silence of nature *(Oppleve stillheten i naturen)*
7	* Being free from children for a while *(Ha fri fra barna en tid)*	30	Recovering strength *(Hente nye krefter)*
8	Enjoy beach and swimming *(Nyte strand og badeliv)*	31	Getting in shape *(Komme i form)*
9	Using skill and knowledge *(Få brukt ferdigheter og kunnskaper)*	32	* Relate to people with similar interests *(Omgås folk med samme interesser som deg selv)*
10	Having time for the family *(Ha tid til familien)*	33	See to it that the children have a pleasant vacation *(Sørge for at barna har det bra i ferien)*
11	Getting to know new people *(Bli kjent med nye mennesker)*	34	* Satisfying an interest in history *(Tilfredsstille historisk interesse)*
12	*Experience fun and exitement *(Oppleve fart og spenning)*	35	* Relax without doing much *(Slappe av og ikke gjøre noe særlig)*
13	Feeling you belong in nature *(Føle tilhørighet til naturen)*	36	* Enjoy luxury *(Nyte luksus)*
14	Getting away from noise and pollution (Komme vekk fra støy og forurensning)	37	* Participate in family events *(Delta i familiære begivenheter)*
15	Working out, really tiring your body *(Ta deg ut og bli skikkelig sliten i kroppen)*	38	Seeing well-known places or sights *(Oppleve kjente steder eller severdigheter)*
16	Being in romantic company *(Ha romantisk samvær)*	39	Having plenty of time/time to do what you please *(Ha god tid/ha tid til å gjøre det du har lyst til)*
17	Experience art and culture *(Oppleve kunst og kultur)*	40	* See and experience Norway *(Se og oppleve Norge)*
18	Exposing your skills *(Vise dine ferdigheter)*	41	Experiencing the special atmosphere of the resort *(Oppleve den spesielle atmosfæren på feriestedet)*
19	Keeping in touch with family living elsewhere *(Beholde kontakt med familie som bor andre steder)*	42	Travelling about/being on the move *(Reise omkring /være på farten)*
20	Eat and drink in good company *(Spise og drikke i godt lag)*	43	Learning something new *(Lære noe nytt)*
21	Using your language skills *(Bruke de språkkunnskapene du har)*	44	Avoid the push and stress of travelling *(Unngå mas og stress ved det å reise)*
22	Feeling the smell of the salty sea *(Kjenne lukten av salt sjø)*	45	Getting to know other countries and cultures *(Bli kjent med andre land og kulturer)*
23	* Being alone *(Være alene)*	46	Swim in clean water *(Bade i rent vann)*

Within a comprehensive questionnaire also covering other travel and leisure topics, a motive checklist was to be filled in by the respondent. In this list, all 36 motive items used in Leaven's (2005) original study were included. Ten additional items were also constructed, with a view to replacing items and possibly extending the scales. The resulting 46 motive items were arranged in an approximate random order, with the intent of spacing items expected to belong to the same scale. The item list is displayed in table 1.

The question format was "What was important to you during your leisure time this summer? On your vacation trip this summer, **how important** were these issues to you?" [Hva var viktig for deg når du hadde fri i sommer? På feriereisen i sommer, **hvor viktig** var disse forholdene for deg?] Consistent with this format, the list of motive items was administered only to respondents who had been away on a summer vacation lasting four days or more (N=154). The response alternatives given were "Not important", "A little important", "Important", and "Very important"; and answers were coded as corresponding numbers 1 through 4.

Pre-testing indicated no problems with the motive list or with other parts of the interview procedure. Further information on the practicalities of the survey is available in Kleiven & Thrane (1996). A preliminary account of the motive part of the data was given by Kleiven (1999).

3. Results

3.1 Properties of nine summed scales

In order to keep the basic metrics of the response format, the mean score on the four items was used as each person's score on the scales. Only 104 of the 154 respondents completed all 46 items.

3.1.1 Reliabilities

Reliability scores of several versions of Kleiven's (2005) scales are shown in the left-hand column of table 2, with comparable scores from the previous study in parentheses. The 'original' scales contain the original four items for each scale. In the 'revised' scales, a 'new' one replaces one item. A closer look at the reliability analyses is called for, however.

Cronbach's alpha for the original INDULGENCE scale is very low at .38. Since this scale also did not perform well in the previous study, it is dropped from further analysis. For the remaining scales, however, alphas range from .60 to .76.

In the previous study, the SUN/WARMTH scale includes the sea-related item 46 (*Swim in clean water*), which did not contribute much to the scale in that study. In the present study, alpha will increase from .76 to .78 if the item is deleted. However, none of our "new" items may replace this item to yield a higher scale alpha.

The other "maritime" item is item 22 (*Feeling the smell of the salty sea*) in the NATURE scale. Reliability analysis of the four original items indicates that alpha will rise from .68 to .83 upon the removal of this item. Item 40 (*See and experience Norway*) appears to be a good substitute, yielding an alpha of .80.

Table 2: Confirmatory Factor Analysis for revised single scales, standardized estimates

Scale and included items		Item a	Item b	Item c	Item d	χ^2 (df=2)	RMSEA
SUN/WARMTH (Items 1, 8, 25 and 46) $\alpha = .76$ **(.69)**	Factor loading	.73	.74	.75	.47	0.70 (p= .70)	.00
	Residual	.46	.46	.44	.78		
ACCOMPLISHMENT (ORIG.) (Items 9, *18*, 26 and 43) $\alpha = .72$ **(.76)**	Factor loading	.75	.40	.77	.56	2.518 (p= .28)	.043
	Residual	.43	.84	.41	.68		
ACCOMPLISHMENT (REV.) (Items 9, 26, *32* and 43) $\alpha = .75$	Factor loading	.77	.77	.58	.54	0.99 (p= .61)	.00
	Residual	.41	.41	.67	.71		
FAMILY (Items 10, 19, 27, 33) $\alpha = .74$ **(.71)**	Factor loading	.68	.42	.60	.84	28.885 (p= .00)	.328
	Residual	.54	.82	.64	.30		
FRIENDS (Items 3, 11, 20, 28) $\alpha = .60$ **(.59)**	Factor loading	.53	.49	.56	.52	3.975 (p= .13)	.084
	Residual	.72	.76	.69	.73		
CULTURE (ORIG.) (Items 17, *21*, 38, 45) $\alpha = .75$ **(.72)**	Factor loading	.66	.50	.79	.69	0.78 (p= .68)	.00
	Residual	.56	.75	.38	.53		
CULTURE (REV.) (Items 17, *34*, 38, 45) $\alpha = .80$	Factor loading	.76	.74	.75	.58	15.329 (p= .00)	.22
	Residual	.42	.45	.43	.66		
NATURE (ORIG.) (Items 4, 13, *22* and 29) $\alpha = .68$ **(.70)**	Factor loading	.73	.93	.16	.72	3.634 (p= .16)	.076
	Residual	.46	.14	.97	.48		
NATURE (REV.) (Items 4, 13, 29 and *40*) $\alpha = .83$	Factor loading	.72	.92	.73	.50	3.871 (p= .14)	.082
	Residual	.48	.16	.47	.75		
PEACE/QUIET (Items 5, 14, 30, 44) $\alpha = .63$ **(.64)**	Factor loading	.66	.53	.65	.39	.019 (p= .99)	.00
	Residual	.57	.72	.56	.85		
FITNESS (Items 6, 15, 24, 31) $\alpha = .74$ **(.81)**	Factor loading	.63	.52	.65	.78	9.285 (p= .01)	.159
	Residual	.60	.73	.58	.39		

Items 46 and 22, then, do not function too well in the present sample. This is consistent with their performance in the previous study (Kleiven, 2005).

The third problem item in Kleiven's study was item 24 (*Taking care of your health*) in the FITNESS scale. In the present study, however, it appears to function rather well, and removing the item will cause alpha to drop from .75 to 69.

In most other scales, reliability analyses indicated no chance of improvements through deletions or substitutions. In the ACCOMPLISHMENT scale, however, replacing item 18 (*Exposing your skills*) by item 32 (*Relate to people with similar interests*) may be considered. With this replacement, alpha will be .75 instead of .72. Similarly, replacing item 21 (*Using your language skills*) in the CULTURE scale by item 34 (*Satisfying an interest in history*) will give in alpha increase from .75 to .80.

Not having access to Structural Equations Programs or Confirmatory Factor Analysis at the time[1], we used the improved reliability scores as the only basis for considering item replacements. Consequently, item 32 replaced item 18 in the ACCOMPLISHMENT scale; item 34 replaced item 21 in CULTURE, and item 22 was replaced by item 40 in NATURE.

3.1.2 Single-scale CFA

Later, confirmatory factor analyses has been done on both original scales and on the revised scales suggested by the reliability analyses. The *Mplus* program (Muthén & Muthén, 1998) was used, with ML estimation. In this analysis, a 'congeneric' measurement model (Pedhazur & Schmelkin, 1991) was assumed, allowing both factor loadings and error terms for the four items of each scale to be unequal. The results are shown in table 2.

For six of the eight original scales, results clearly support the measurement model (χ^2 p value > .05). This is not the case, however, with the FAMILY scale. Here, both Chi-square and the RMSEA index indicate a bad fit. The FITNESS scale also has problems, but apparently less severe.

For the three revised scales, the picture is mixed. For the revised ACCOMPLISH-MENT scale, the model is even closer to the data than the original. For the NATURE scale, the two versions yield very similar and acceptable results. The factor loading of item 22 in the original scale (.16) is rather low, however, while the loading of the replacement item in the revised scale (.50) looks better. The measurement model for the revised CULTURE scale appears not to fit the data very well (χ^2 p value = .00; RMSEA = .22).

Not only a 'congeneric' measurement model may be relevant, however. Pedhazur & Schmelkin (1991) point out that 'tau-equivalent' measurement models assume equal factor loadings, while the congeneric model does not. 'Parallel-measures' models assume (like the Cronbach's alpha statistic) that all items have *both* equal factor loadings *and* equal error terms.

Results of calculating the 'tau-equivalent' and 'parallel measures' models are shown in table 3. Generally, the congeneric models appear to fit the data better. For five out of eight scales, the fit of the congeneric model is better than with the tau-equivalent model. The parallel measures model fares even worse. Here, data favor the congeneric model for seven of the scales, and even the last scale (FRIENDS) comes very close (Difference chi-square = 10.422; p < .10).

---

[1] The survey was planned in the spring of 1995, and was carried out the following summer.

Table 3: Fit indices for alternate measurement models of eight revised scales.

Scale	Tau- equivalent Models		Parallel measures models	
	RMSEA	χ^2 (df=4)	RMSEA	χ^2 (df=7)
SUN/WARMTH	.074	7.032 (p= .14)*	.150	28.819 (p= .00)**
ACCOMPLISHMENT	.051	5.458 (p= .24)*	.085	14.071 (p= .05)**
FAMILY	.252	35.632 (p= .00)*	.199	41.771 (p=.00)**
FRIENDS	.025	4.364 (p= .36)	.087	14.397 (p= .05)
CULTURE	.153	16.888 (p= .00)	.143	26.713 (p= .00)**
NATURE	.127	12.908 (p= .01)*	.201	46.424 (p=.00)**
PEACE/QUIET	.000	2.665 (p= .61)	.071	11.908 (p= .10)**
FITNESS	.162	19.105 (p= .00)*	.124	22.542 (p= .00)**

  * Fits data less well than congeneric model (difference df =2; p< .05).
** Fits data less well than congeneric model (difference df =5; p< .05).

3.1.3 Scale revisions

The scale revisions may now also be considered in light of the CFA results. In retrospect, the replacements made in the ACCOMPLISHMENT and NATURE scales are acceptable. Not only does Cronbach's alpha (related to simple additive index and 'parallel-measures' model) improve with the substitutions; also CFA (related to a more general, 'unconstrained' or 'congeneric' measurement model) indicates that the revised models fit the data equally well or better.

For the revision of the CULTURE scale, however, Cronbach's alpha and the CFA yield contradictory results. The substitution of item 34 for item 21, therefore, has less support than the two other replacements.

It should also be noted that neither the FAMILY nor the FITNESS scales showed acceptable fits in the CFA. Besides, the alpha value of the FRIENDS scale is rather low (.60). None of our 'new' items could be used to improve these scales, however, so no revision was made here. The data do suggest, however, that the three scales have some room for improvement.

3.1.4 Scale statistics

The basic statistics of the eight resulting scales are displayed in table 4. First, none of the means is close to the scale ends of 1 and 4, indicating that ceiling or floor effects are not likely. Secondly, the distibutions appear fairly normal, containing no major problems with skew or kurtosis.

Table 4: Statistics for revised leisure and travel motive scales (N=261).

Scale	Valid N	Mean	S.E. of Mean	Standard Dev.	Skew-ness	S.E. of Skew-Ness	Kurt-osis	S.E. of Kurtosis
SUN/WARMTH	154	2,35	0,07	0,82	0,29	0,20	-0,88	0,39
ACCOMPLISHMENT	154	2,33	0,06	0,78	0,18	0,20	-0,52	0,39
FAMILY	154	2,78	0,07	0,86	-0,46	0,20	-0,69	0,39
FRIENDS	154	2,68	0,05	0,68	-0,29	0,20	-0,41	0,39
CULTURE	154	2,30	0,07	0,86	0,06	0,20	-1,05	0,39
NATURE	153	2,80	0,06	0,78	-0,48	0,20	-0,48	0,39
PEACE/QUIET	154	2,86	0,06	0,70	-0,51	0,20	-0,16	0,39
FITNESS	153	2,25	0,06	0,73	0,07	0,20	-0,56	0,39

It may also be worth noting that the relative sizes of the means come rather close to the results of the previous study. In figure 1, scale means of the two studies are plotted. All means lie between the values of 2 (*"A little important"*) and 3 (*"Important"*), and the two curves have a rather similar pattern. However, the difference between the two estimated population means is statistically significant on the five scales ACCOMPLISHMENT (t= -4,9210, df= 550), FAMILY (t= 2,3696), CULTURE (t= -2,7734), NATURE (t= -3,0139) and PEACE/QUIET (2,2261).

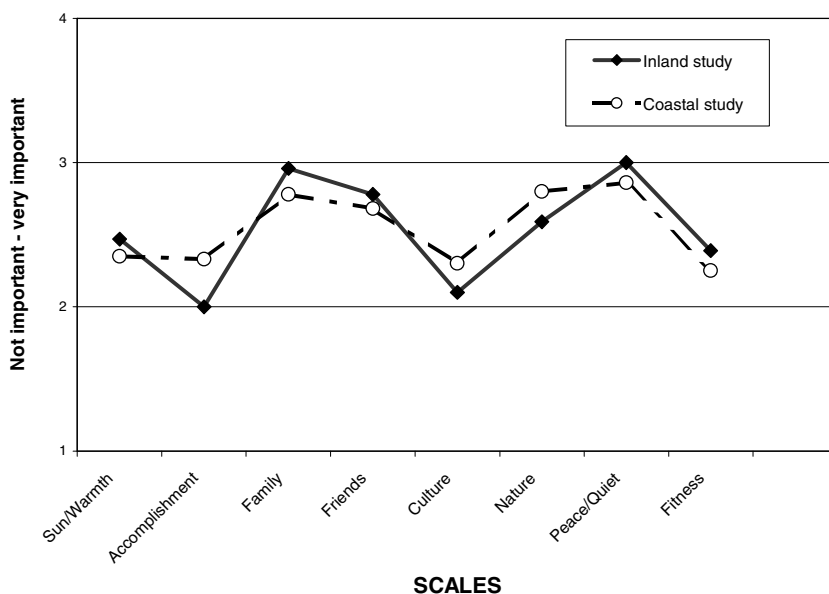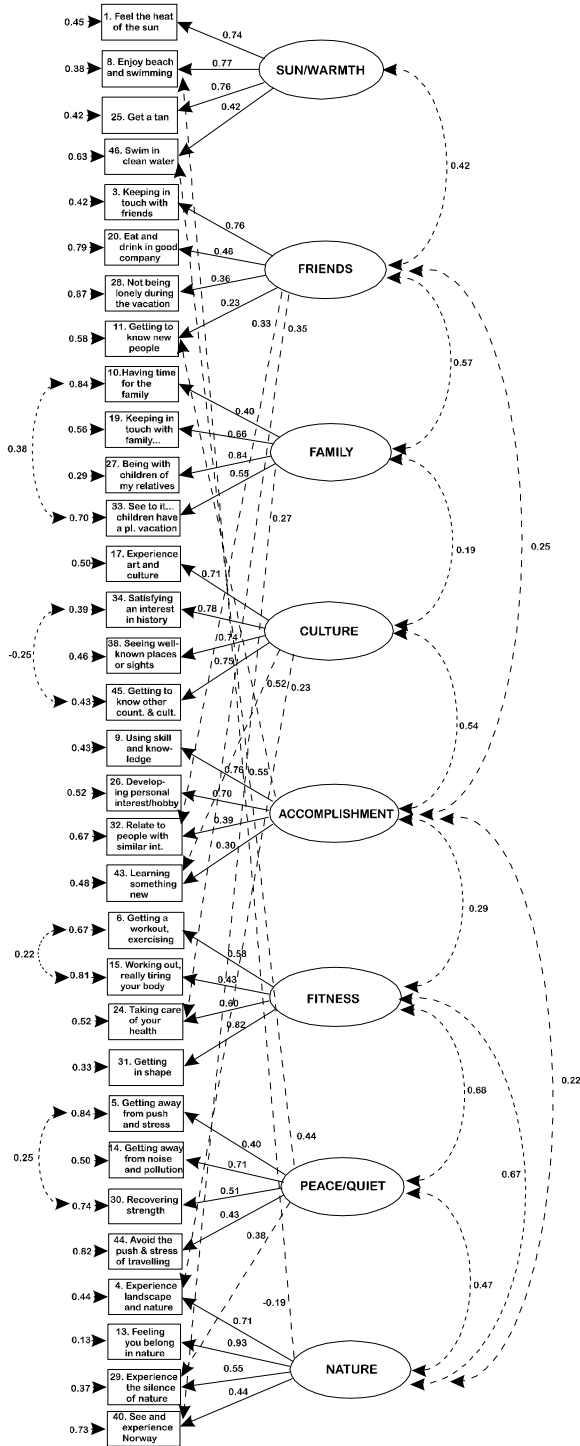**Figure 1; Scale means from two studies**

Figure 2; Modified combined measurement model (N=154). Dotted lines represent modifications to the original independent-factors model.

3.2 Factor structure

A confirmatory factor analysis suggests that an independent factors measurement model combining the eight factors is not tenable (χ^2 = 1095.12; df = 464; p = .00. RMSEA = .117). A fully correlated factors model fares slightly better (χ^2 = 843.21; df = 436; p= .00; RMSEA = .092), but also fits our data badly.

Attempts at modifying the combined model have not been very successful, and acceptable fit measures have not been obtained. Models with a limited number of factors present fewer problems, however, showing acceptable fit with the data after minor modifications.

In spite of fit problems, the differences between our model and the data do not necessarily invalidate the general factor structure of the combined model. A preliminary impression of these differences may be gained from figure 2, which shows some of the modifications required to obtain a better fit. All arrows in the resulting figure represent statistically significant relationships (p < .05). Three types of modifications have been done to this model, all improving the fit considerably. First, nine items have been allowed to load on more than one factor. Second, four within-factor error terms are freed to correlate. And third, only ten between-factor correlations are allowed out of the 28 possible in the eight-factor model.

Due to convergence problems associated with estimating the complex model with a limited data set, the Mplus missing data option (Muthén & Muthén, 1998) was used. Here, missing data is imputed for cases where more than 10% of the data is present, yielding an effective N of 154. Even with the above modifications, however, the fit is not quite acceptable (χ^2 = 743.83; df = 441; p= .00; RMSEA = .067).

3.3 Validity/prediction models

In the omnibus study (Kleiven & Thrane, 1996), questions were asked about the amount of time respondents had spent on six types of activity during their vacation trip last summer. The data from these questions may be used for gaining an impression of the scales' predictive potential.

Table 5: Correlations between summed scales' and time spent on six activities during vacation trip. (N=151-153).

| Scale/Activity	Boating or fishing trips at sea	Physical training or sports	Sun-bathing and swimming	Going for a walk in nature	Go to concert or theatre	Visit relatives or friends
SUN/WARMTH	.18*	.05	.42***	-.04	-.04	.09
ACCOMPLISHMENT	.11	-.03	-.08	-.01	.29***	.12
FAMILY	.16	-.15	-.07	-.03	.05	.41***
FRIENDS	.22**	-.04	.11	-.05	.12	.29***
CULTURE	.10	.00	-.13	.02	.28***	.03
NATURE	.17*	-.03	-.20**	.32***	.03	.00
PEACE/QUIET	.11	.09	.03	.03	.04	.02
FITNESS	-.04	.32***	-.10	.28***	.11	.09

*p< .05; **p< .01; ***p < .001

In table 5, we see some high (and significant) correlations, and most of them do appear in the cells one would expect out of common sense. The SUN/WARMTH motivation is highly correlated with the activities of sunbathing and swimming; and there are strong ties between the FAMILY and FRIEND motives on one hand and visiting friends and relatives on the other. The CULTURE motivation is related to visiting concerts and theatre, while the FITNESS motivation, is associated with sports and with nature walks.

We also find less trivial associations, however. While it may not be surprising to find the NATURE motive to correlate highly and positively with nature walks, its strongly *negative* association with sunbathing/swimming is not self-evident. Also, the fact that both the SUN/WARMTH, the FRIENDS and THE NATURE scales are correlated with boating and fishing may be worth noting.

Closer analyses of the data indicate, however, that role of the motive scales may more correctly be understood in a broader context, also considering the influence of the socio-demographic factors of age, gender, income and education on vacation behavior. Through the use of MIMIC (multiple indicator multiple cause) models (Muthén & Muthén, 1998), we also gain the additional advantage of using 'parallel-measures' models, known to be closer to our motives data. Since the activities data are heavily skewed towards the left of the distribution, they are recoded into the two categories "participating" and "non-participating" for each activity type. This dependent variable being categorical, the Mplus default WLSMW estimator was used (Muthén, 1984; Muthén & Muthén, 1998), with a robust goodness-of-fit test (Satorra, 1992; Satorra & Bentler, 1994; Satorra & Bentler, 1988).

First, let us consider a prediction model for outdoor walks, shown in figure 3. Among the motive scales, only NATURE proves to be a significant predictor of walks (Est./S.E. = 4.153). While no demographic variable is related to walks directly, the path from AGE to NATURE scale is significant (Est./S.E. = 3.631). All factor loadings are significant. This simple model fits the data very well (χ^2 = 3.59; df = 7; p = .83).

Figure 3; Prediction model for outdoor walks during vacation. Completely standardized estimates. N = 135.
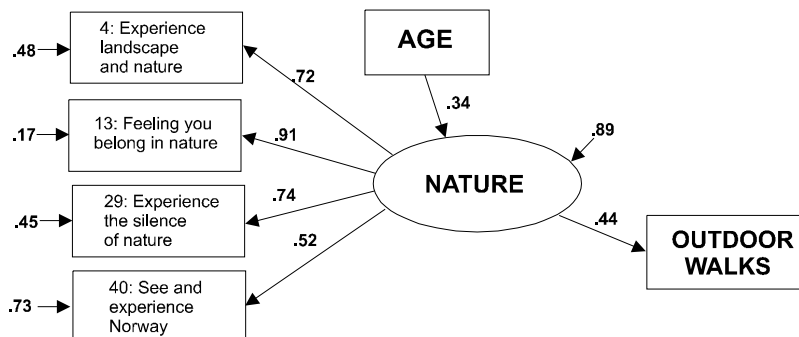
Figure 4; Prediction model for sunbathing/swimming during vacation.
    Completely standardized estimates. N = 130.
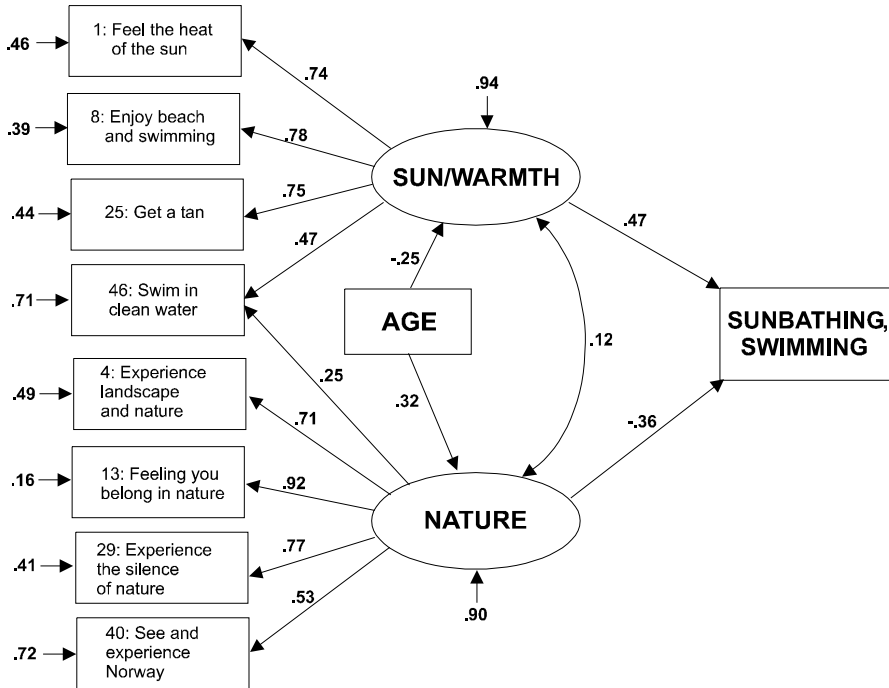
For other activities, more complex prediction models may be identified. An example of this is offered in figure 4. Here, both SUN/WARMTH (Est./S.E. = 4.22) and NATURE (Est./S.E. = -3.574) are significant predictors, and all factor loadings are significant. Please note that item 46 is allowed to load on both factors, since this proved necessary for an acceptable fit of the measurement part of the model. While no demographic variable relates to the activity measure directly, AGE is significantly related to both SUN/WARMTH (Est./S.E. = -2.752) and to NATURE (Est./S.E. = 3.387). The model fit is clearly satisfactory (χ^2 = 18.514; df= 16; p = .29).

Finally, consider the model in figure 5. While all factor loadings on CULTURE are significant, the fit of the measurement part of the model is acceptable only when the error terms of items 34 and 45 were allowed to correlate. While CULTURE is a significant predictor of the CONCERT/THEATRE activities (Est./S.E. = 2.56), several demographic variables also play a part. Both GENDER (Est./S.E. = 2.80) and LENGTH OF EDUCATION (Est./S.E. = 3.56) are significantly related to this activity measure. In addition, AGE is significantly related to both CULTURE (Est./S.E. = 2.459) and to the activity measure (Est./S.E. = -2.896). The model appears to be rather consistent with our data (χ^2 = 14.737; df = 11; p = .19).

Figure 5; Prediction model for concert/theatre during vacation.
      Completely standardized estimates. N = 133.
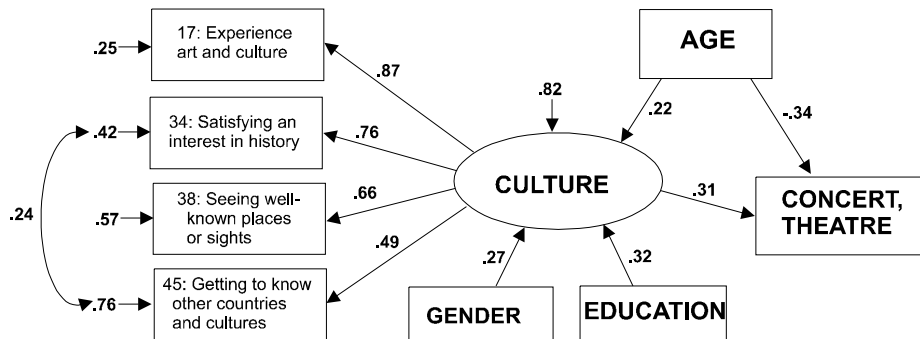
Comments on the *factual content* of these models are not relevant to the present context. The three examples of MIMIC models show, however, that the motive scales do predict interesting vacation behavior types, also when a parallel measures model is assumed. The scales also hold their own in the presence of central demographic variables, but the interplay between motive measures and demographics is not simple.

4. Discussion

Eight leisure motive scales do seem to replicate rather well in our sample, answering the central research question in a fairly clear manner. With Cronbach alpha values ranging from .60 to .76, reliabilities are comparable to the results of the original study (Kleiven, 2005). Summed-scale means are also very close to the original, without important score distribution problems. The relative size of the scale means also come close to previous results. Significant differences between the two samples were observed on most scale means, however.

Yielding a very low reliability score, a ninth scale (INDULGENCE) was dropped from the analysis. Since there also were problems with this scale in the previous study, this is consistent with earlier results.

Two items known to give problems also did not perform well in the present study. Items 46 (*Swim in clean water*) and 22 (*Feeling the smell of the salty sea*) contributed efficiently to the summed scales in neither study. In view of comparable results in both the inland and the coastal sample, the problem thus is not likely to stem from the first sample's limited experience with the sea. A third 'problem' item (24: *Taking care of your health*) appears to function adequately in the new sample. On the basis of improved reliability scores, three scales had one item replaced.

Confirmatory factor analyses show that 'congeneric' measurement models fit our data better than the 'parallel measures' model implied by simple summed scales. This may indicate

that a SEM approach should be preferred to simple summed scales when using data from the eight scales, probably exploiting more of the common scale variance.

Unfortunately, one of the three item replacements (in the CULTURE scale) is not supported by the CFA. In view of the apparent superiority of 'congeneric' models, then, this revision may not have been a wise one.

At any rate, the 'congeneric' measurement model fits our data well on most single scales. This supports the idea that the four items do in fact measure something in common. It does not imply, however, that the four items contribute *evenly* to the factor or that they have similar residuals.

In spite of satisfactory single-scale measurement models, the complete eight-factor measurement model appears not to fit our data very well. Even after some modifications, there are patterns in the data that are not accounted for by the model. Some likely model improvements may be identified. First, most inter-factor correlations should be left out of the model, while a smaller number of correlated factors should be allowed. Apparently, neither an uncorrelated factors model nor a 'fully' correlated factors model is appropriate.[2]

Second, several items do appear to load on two factors. Even in the absence of proper theory for each factor, however, it should be noted that the new item/factor links do not seem to change the immediate 'meaning' of the factors very much. In other words, the face validity of the scales is not really challenged by the items loading on two scales.

Third, the correlated error terms allowed also does not change the basic model very much. All correlated errors occur within a scale, not affecting the basic dimensionality of the factors.

It is also quite likely, however, that the large and complex measurement model may be overly ambitious. Large models with a large number of covariance items will frequently not be supported by the data, and limiting the number of variables and relations is often recommended in SEM work. Our small sample may well be adding to this problem, with a low estimates/observations ratio likely to increase the risk of unstable estimates.

Also, a basic assumption of the CFA may be unrealistically strong. With a multiple factors model, items are typically expected to have zero loadings on all factors but one. This assumption should perhaps be replaced with something closer to what is typically seen in EFA; while factor loadings on the 'proper' factor should be clearly higher than others, small or even medium loadings on other factors are not seen as a threat to the basic model.

All in all, then, I do not see the complications with the eight-factor model as contradicting the notion of eight independent, but partly correlated scales. Convergent validity is shown by the four or more items that do load significantly on each scale. Divergent validity is demonstrated by the fact that most scales are not correlated, and that no scale intercorrelation is high enough to rule out scale independence.

But even if the joint eight-factor measurement model may not be confirmed, the scales may safely be used individually. As indicated by the MIMIC model examples, the scales may contribute to the prediction of interesting forms of leisure behavior, adding to the predictive power of central demographic variables. Even the simple summed scales appear to have some predictive validity, as shown by their correlations with leisure behavior measures.

---

[2] In our modest sample, higher-order factors could not be identified.

All is not well, however. In spite of the generally positive results, there is certainly room for improvements. The internal consistency of scales may probably be increased by new item replacements. Also, having scale data only from two small samples, results must still be viewed as open to sample-specific variations. There is a need, therefore, to see how the scales perform in larger samples. A nationally representative sample would perhaps be ideal, offering the possibility of establishing useful standards for the eight scales.

5. References

1. Arbuckle, J. L. (1997). <u>Amos Users' Guide Version 3.6</u>. Chicago: SmallWaters Corporation.

2. Beard, J. G., & Ragheb, M. G. (1983). Measuring Leisure Motivation. <u>Journal of Leisure Research, 15</u>(3), 219-228.

3. Crandall, R. (1980). Motivations for Leisure. <u>Journal of Leisure Research, 12</u>(1), 45-55.

4. Crompton, J. (1979). Motivations for pleasure vacation. <u>Annals of Tourism Research, 6</u>(4), 408-424.

5. Harrill, R., & Potts, T. D. (2002). Social Psychological Theories of Tourist Motivation: Exploration, Debate, and Transition. <u>Tourism Analysis, 7</u>, 105-114.

6. Jamal, T., & Lee, J.-H. (2003). Integrating Micro and Macro Approaches to Tourist Motivations: Toward an Interdisciplinary Theory. <u>Tourism Analysis, 8</u>, 47-59.

7. Kleiven, J. (1998b). <u>Skalaer for måling av aktivitets- og motivmønstre i en norsk lokalbefolknings ferie og fritid</u> . (Report No. Working paper no. 77/1998). Lillehammer: Lillehammer College.

8. Kleiven, J. (1999). The Lillehammer Scales: Eight Dimensions of Leisure and Travel Motivation (Abstract). In <u>Konferensrapport 7:e Nordiska Forskarsymposiet i Turism 1998.</u> Östersund: ETOUR.

9. Kleiven, J. (2005). Measuring Leisure and Travel Motives in Norway – Replicating and Supplementing the Leisure Motivation Scales . <u>Tourism Analysis, 10</u>(2), 109-122.

10. Kleiven, J., & Thrane, C. (1996). <u>Reiselivsundersøkelsen i Sandefjord 1995 - en metoderapport.</u> (Report No. Working paper no. 16). Lillehammer: Lillehammer College.

11. Lounsbury, J. W., & Franz, C. P. G. (1990). Vacation Discrepancy: A Leisure Motivation Approach. <u>Psychological Reports, 66</u>, 699-702.

12. MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model Modifications in Covariance Structure Analysis: The Problem of Capitalization on Chance. <u>Psychological Bulletin, 111</u>(3), 490-504.

13. MacCallum, R. C. (1995). Model Specification: Procedures, Strategies, and Related Issues. In R. H. Hoyle (Ed.), <u>Structural Equation Modeling: Concepts, Issues, and Applicatons</u> . Thousand Oaks, California: SAGE.

14. Muthén, L. K., & Muthén, B. O. (1998). <u>Mplus: The Comprehensive Modeling Program for Applied Researchers</u>. Los Angeles: Muthén & Muthén.

15. Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. <u>Psychometrika, 49</u>(1), 115-132.

16. Pedhazur, E. J., & Schmelkin, L. P. (1991). <u>Measurement, Design, and Analysis</u>. New Jersey: Erlbaum.

17. Ryan, C. (1993). Correlations between clusters based on the Ragheb and Beard Leisure Motivation Scale and the attributes of desired holiday locations. In: <u>Proceedings of Third International Conference of the Leisure Studies Association</u> UK: Loughborough University.

18. Ryan, C. (1994a). Leisure and tourism - the application of leisure concepts to tourist behavior - a proposed model. In: A. V. Seaton (Ed.), <u>Tourism: The State of the Art</u> (pp. 294-307). Chichester: John Wiley & Sons.

19. Ryan, C. (1997). Similar Motivations - Diverse Behaviours. In C. Ryan (Ed.), <u>The Tourist Experience</u> (pp. 25-47). London: Cassell.

20. Ryan, C., & Glendon, I. (1998). Application of Leisure Motivation Scale to Tourism. <u>Annals of Tourism Research, 25</u>(1), 169-184.

21. Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. In P. V. Marsden (Ed.), <u>Sociological Methodology, Vol. 22</u> . Oxford: Basil Blackwell.

22. Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. <u>1988 Proceedings of the Business and Economic Statistics Section of the American Statistical Association</u> (pp. 308-313.).

23. Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye, & C. C. Clogg (Eds.), <u>Latent Variable Analysis: Applications to Developmental Research</u> (pp. 399-419). Newbury Park: SAGE.

24. Tinsley, H. E. A. (1984). The psychological benefits of leisure counseling. <u>Society and Leisure, 7</u>, 125-140.

Lillehammer University College

Høgskolen i Lillehammer/Lillehammer University College

2626 Lillehammer, Norway

Telefon/phone: (+47) 61 28 80 00

Fax: (+47) 61 26 07 50

www.hil.no