

MEASURING LEISURE AND TRAVEL MOTIVES IN NORWAY: REPLICATING AND SUPPLEMENTING THE LEISURE MOTIVATION SCALES

JO KLEIVEN

Lillehammer University College, N-2626 Lillehammer, Norway

Based on previous research, nine central motive dimensions were identified. Four scales (Culture, Friends, Accomplishment, and Peace/Quiet) may be seen as conceptual replications of the four scales of Beard and Ragheb, while five (Sun/Warmth, Family, Nature, Fitness, and Indulgence) were found in other research in Norway and abroad. In a representative survey of the Norwegian inland city of Gjøvik ($N = 401$), four-item rating scales for each of the nine dimensions were tried out. The psychometric properties of most scales are promising, with alpha values comparable to those found in previous work in this field. Judging from a SEM perspective, the replication of the four Beard and Ragheb scales is only moderately successful. While a combined nine-factor model does not appear tenable, an eight-factor model is closer to the data. More work is needed on the issue of validity.

Key words: Leisure; Travel; Motives; Scale construction; Structural equations modeling

The identification of important motives has been a central topic in the field of leisure research. In a recent textbook, Pearce (1995) even believes that “As tourism grows into an increasingly sophisticated consumer industry, ... the motivation of tourists will become a core part of all tourism studies” (p. 178). And Lee, O’Leary, Lee, and Morrison (2002) show motives to be important indeed: motive factors influence destination choice more than even demographics and budget. Although motives are commonly seen as important for understanding the great variation in human leisure activities, rather different conceptual approaches to this question have been offered (Harrill & Potts, 2002; Jamal & Lee, 2003).

Motive Lists

Early writers in the field, like Crompton (1979), Crandall (1980), and Tinsley (1984), revealed no apparent consensus, offering different lists of leisure and travel motives. As Schmidhauser (1989) has pointed out, however, tourists’ needs and motives have changed a great deal over time, being subject to the influence of historical, political, economic, and technological change. A similar point is made by Jamal and Lee (2003), arguing that individual motives should be viewed in their social (macro) context.

Working with different populations and different forms of leisure in quite different situations, then, should lead to substantial variation in the understand-

ing of leisure and travel motives. And Hall and Page (1999) indeed find tourist motivation to be "an incredibly complex area of research" (p. 52) and that "no all-embracing theory of tourist motivation has been developed which has been adapted and legitimized by researchers in other contexts" (p. 52). The changing perspectives in tourism motivation studies are also thoughtfully reviewed by Harrill and Potts (2002).

However different, leisure motive lists may also be seen to overlap substantially. Certain motives do crop up in several lists, but under slightly different labels. As Ryan (1997) puts it, "The adjectives may differ, . . . but recurrent themes emerge" (p. 27)

This should not be interpreted as a general consensus, however. Differences between motive lists are more common than similarities, even if related motivational concepts have been used in several surveys.

Nonetheless, certain researchers have been more influential than others. In a well-known article, "Measuring leisure motivation," Beard and Ragheb (1983) propose four subscales for measuring leisure motivation dimensions. The dimensions were labeled *Intellectual*, *Social*, *Mastery/Competence*, and *Stimulus Avoidance*. Cronbach's alpha values of the 12-item scales ranged from 0.90 to 0.93, indicating high reliability.

The four scales have also been employed by a number of other researchers. In an American replication, alpha values above 0.80 for the four scales were observed (Lounsbury & Franz, 1990). A British replication (Ryan, 1993) also appears to be quite successful, and Ryan (1994) sees the replications as supporting the scales. More recently, the scales were used by Ryan and Glendon (1998) in an article demonstrating the usefulness of a shorter version for tourism research. The four subscales from Beard and Ragheb's (1983) research, then, appear to represent interesting and useful motivational dimensions. Further encouragement may be gained from the fact that similar motives are also found in other studies, often with slightly different names. Consequently, the four Beard and Ragheb scales form a natural starting point for further work on leisure motives.

Research Questions

In a recent note, Kim (1999) reminds us that travel motives should be seen in a *cross-cultural* perspec-

tive. It should not be taken for granted that the interests and needs of one set of people are valid or relevant for other people in a different culture.

In addition, there is the problem of translating similar concepts from one language to another. Frequently, quite different words, phrases, or examples are needed to convey the same general idea in two cultures. Simply translating of the test items, therefore, may not work.

A first aim of this article, therefore, is to attempt a *conceptual replication* of the four leisure motivation subscales in Norway, not a literal translation. If this replication is successful across countries, the scales are more likely to be generally useful in international leisure and tourism research.

It should be noted, however, that Beard and Ragheb (1983) make no claim that their list of factors is complete or exhaustive. Their argument for preferring a smaller number of factors boils down to a wish "to build more general subscales" (p. 220). But no statement is made about the adequate or maximal number of such "general" subscales. In my view, therefore, it is no violation of Beard and Ragheb's assumptions when the present article will suggest a higher number of motive factors.

Several analyses do in fact suggest additional motive types. In the 1992 Norwegian Vacation Survey (Haukeland, 1993), the new motive dimensions of *Family*, *Culture*, and *Indulgence* were found. From Kleiven's (1998a) reanalysis of the same data, the motive factors of *Friends*, *Nature*, *Peace/Quiet*, and *Fitness* were added. In Jamrozny and Uysal (1994) and in Schmidhauser (1989) a *Sun/Warmth* motive was found. A more detailed discussion of this may be found in Kleiven (1998b).

A second purpose of this article, therefore, is to supplement Beard and Ragheb's (1983) four-factor view of leisure motivation, adding other scales known to be relevant to the Norwegian population. The new scales are *Indulgence*, *Nature*, *Fitness*, and *Sun/Warmth*. The *Accomplishment* scale, however, is intended to replicate the original Beard and Ragheb *Mastery/Competence* factor. Likewise, the *Culture* scale should correspond to their *Intellectual* scale, and the *Peace/Quiet* to *Stimulus Avoidance*. Following the advice of Seaton (1994), it would be interesting to see if Beard and Ragheb's general *Social* factor can be replaced by two more specific factors: *Friends* and *Family*. (In Beard and Ragheb's

“Social” scale, no item specifically mentions social relations within the family.) We thus arrive at nine motive dimensions, listed in Table 1.

A third research question concerns the possible intercorrelation of the scales. Beard and Ragheb (1983) used common factor analysis and Varimax rotation in the initial phases of their work. Proceeding to use summed-item scales, however, their dimensions were positively correlated, with four out

of the six intercorrelations, reaching 0.33 or higher. According to Nunnally and Bernstein (1994), orthogonal (Varimax) rotation (independent dimensions model) is not necessarily appropriate in a case like this. It should be checked, therefore, if an independent factors model or a correlated factors model is more appropriate.

A fourth and final question to be addressed by the present article is the potential *validity* of the

Table 1
Motive Items and Scales of the Survey

1. Sun/warmth (<i>Sol/varme</i>)
a) Feel the heat of the sun (<i>Føle varme fra solen</i>) (1)
b) Enjoy beach and swimming (<i>Nyte strand og badeliv</i>) (6)
c) Get a tan (<i>Bli brun</i>) (21)
d) Swim in clean water (<i>Bade i rent vann</i>) (36)
2. Accomplishment (<i>Mestring</i>)
a) Using skill and knowledge (<i>Få brukt ferdigheter og kunnskaper</i>) (7)
b) Exposing your skills (<i>Vise dine ferdigheter</i>) (15)
c) Developing personal interest/hobby (<i>Videreutvikle personlig interesse/hobby</i>) (22)
d) Learning something new (<i>Lære noe nytt</i>) (31)
3. Family (<i>Familie</i>)
a) Having time for the family (<i>Ha tid til familien</i>) (8)
b) Keeping in touch with family living elsewhere (<i>Beholde kontakt med familie som bor andre steder</i>) (16)
c) Being with children of my relatives (<i>Være sammen med barn i slekten</i>) (24)
d) See to it that the children have a pleasant vacation (<i>Sørge for at barna har det bra i feriene</i>) (28)
4. Friends (<i>Venner</i>)
a) Keeping in touch with friends (<i>Beholde kontakten med venner</i>) (2)
b) Getting to know new people (<i>Bli kjent med nye mennesker</i>) (9)
c) Eat and drink in good company (<i>Spise og drikke i godt lag</i>) (17)
d) Not being lonely during the vacation (<i>Ikke være ensom i feriene</i>) (25)
5. Culture (<i>Kultur</i>)
a) Experience art and culture (<i>Opplive kunst og kultur</i>) (14)
b) Using your language skills (<i>Bruke de språkkunnskapene du har</i>) (18)
c) Seeing well-known places or sights (<i>Opplive kjente steder eller severdigheter</i>) (29)
d) Getting to know other countries and cultures (<i>Bli kjent med andre land og kulturer</i>) (34)
6. Nature (<i>Natur</i>)
a) Experience landscape and nature (<i>Opplive landskap og nature</i>) (3)
b) Feeling you belong in nature (<i>Føle tilhørighet til naturen</i>) (10)
c) Feeling the smell of the salty sea (<i>Kjenne lukten av salt sjø</i>) (19)
d) Experience the silence of nature (<i>Opplive stillheten i naturen</i>) (26)
7. Peace/quiet (<i>Fred/ro</i>)
a) Getting away from push and stress (<i>Komme bort fra mas og stress</i>) (4)
b) Getting away from noise and pollution (<i>Komme vekk fra støy og forurensning</i>) (11)
c) Recovering strength (<i>Hente nye krefter</i>) (27)
d) Avoid the push and stress of traveling (<i>Unngå mas og stress ved det å reise</i>) (33)
8. Fitness (<i>Trim</i>)
a) Getting a workout, exercising (<i>Få trim eller mosjon</i>) (5)
b) Working out, really tiring your body (<i>Ta deg ut og bli skikkelig sliten i kroppen</i>) (12)
c) Taking care of your health (<i>Ta vare på egen helse</i>) (20)
d) Getting in shape (<i>Komme i form</i>) (32)
9. Indulgence (<i>Nytelse</i>)
a) Being in romantic company (<i>Ha romantisk samvær</i>) (13)
b) Experiencing the special atmosphere of the resort (<i>Opplive den spesielle atmosfæren på feriestedet</i>) (23)
c) Having plenty of time/time to do what you please (<i>Ha god tid/ha tid til å gjøre det du har lyst til</i>) (30)
d) Traveling about/being on the move (<i>Reise omkring /være på farten</i>) (35)

The numbers in parentheses indicate each item's position on the list used for data collection.

scales. To a Norwegian reader, the name of the scaled dimensions will correspond quite well to the factual content of the items included in each scale. And, hopefully, the translation to English will not impede this immediate "understanding" of what the scales are covering. Neither face validity nor content validity, however, guarantees the usefulness of the scales in leisure and travel research. Therefore, a preliminary check will be made on the scales' *concurrent* validity. A more comprehensive discussion of this question must be left for future work, however.

A common problem in leisure and travel motive research is sample dependency. Frequently, simple *convenience samples* have been used for investigating leisure and travel motives. Crompton (1979), for example, used a convenience sample of only 39 informants, while Fontaine (1993) had 133 respondents and Lounsbury and Franz (1990) had 186 in their convenience samples. Also, some rather *special groups* have been studied. Dunn Ross and Iso-Ahola (1991) looked at visitors riding sightseeing buses in Washington DC; Graefe, Ditton, Roggenbuck, and Schreyer (1981) studied river floaters in national parks; Loker-Murphy (1996) gathered information from young budget travelers on extended holidays, while Tinsley and Kass (1979) investigated undergraduate college students. While the authors of this research probably are well aware of the limitations of their samples, not all their readers may perhaps be trusted to understand the obstacles to generalization inherent in such samples.

It is well known, however, that important demographic differences exist within and between several forms of leisure and travel behavior (Kelly, 1990; Middleton, 1994; Thrane, 1996, 1997).

In order to arrive at travel and leisure motives with some *general* validity, therefore, samples that are representative of larger populations should be preferred. Of course, such samples are also found in the travel motive literature. Schmidhauser (1989), Cha, McCleary, and Uysal (1995), and Ryan and Glendon (1998) may serve as examples of this, as well as several Norwegian studies (Aasetre, Kleiven, & Kaltenborn, 1994; Haldorsen, 1981; Haukeland, 1993). Partly continuing this line of research, the present article will also use a representative sample.

A final problem to be mentioned is also a methodological one. Beard and Ragheb's (1983) subscales

each contain 12 items. For use in omnibus surveys, where motive scales may only be a small part of a comprehensive questionnaire, this may be too extensive. Even the shorter eight-item version of Beard and Ragheb's four scales is likely to stress the space and time available in such interviews. In order to optimize the scales for use within more comprehensive instruments, therefore, the number of items in the scales will be kept to an absolute minimum.

It may not be self-evident, however, exactly how small the number of items may be. On this point, a rather practical consideration has been guiding my choice. Confirmatory Factor Analysis (Jöreskog & Sörbom, 1993; Nunnally & Bernstein, 1994) is likely to be useful in this type of work, and I would like to have access to this method. Because four items is recommended as a minimum in that context (Kline, 1998), an attempt will be made to arrive at four-item motive scales.

Summing it all up, then, the present article has four aims:

1. A *conceptual replication* of the four Beard and Ragheb (1983) leisure motivation subscales will be undertaken.
2. Based on other research, five more motive dimensions will be added, resulting in a *nine-factor leisure motive model*.
3. The data will be examined to see if *independent factors* may be assumed, or if a *correlated-factors* model is more appropriate.
4. Preliminary analyses of the scales' *concurrent validity* will be made.

The scales are based on data from a representative sample, and scale length is limited to four items.

Method

Data for this study were collected as part of an omnibus survey, investigating the "Vacation Habits" of the population in the inland town of Gjøvik. The survey was planned and carried out by the author, one colleague, and a very cooperative group of students at Lillehammer College. [A closer account of the practicalities of the survey is given in Kleiven & Thrane (1994).]

Building on the nine motive dimensions selected, a large number of potential items were used in ex-

tensive pretesting. Some were taken from related Norwegian vacation and leisure surveys (Haukeland, Nymoen, & Rideng, 1991; Vaagbø, 1993), and some were translated and adapted from international reports (Beard & Ragheb, 1983; Crompton, 1979; Driver, 1977; Tinsley & Kass, 1979). A small number of entirely new items were also developed. The first versions of a preliminary item list were administered informally to small convenience samples, and a number of initial corrections and improvements were made during this process.

In Norway, about 25% of the population do not travel much, their longest vacation trip lasting less than 4 days (Haukeland et al., 1991). To obtain motive information from all parts of the population, therefore, we used the question format "During your vacation and leisure time this summer, how important were these following issues to you?" The response alternatives given were "Not important," "A little important," "Important," and "Very important," and answers were coded as corresponding numbers 1 through 4.

The last version of the preliminary list of items contained 7 to 9 items for each scale. This list was used on a small convenience sample, consisting of 25 respondents found in suburban homes close to the dormitories of the college. Factor analysis and reliability analysis on the resulting material indicated that most items performed rather well. For each scale, the four items with the highest reliability were selected for inclusion in the final list of items, shown in Table 1.

The 36 remaining motive items were arranged in a quasi-random order, making sure that same-scale

items were placed at some distance, and that items from all nine scales were found throughout the list. Further details may be found in Kleiven's (1998b) preliminary report.

Data for the survey were collected through personal interviews with people in their private homes. The motive list, however, was administered as a separate check list within this procedure, and was filled in by the respondent. The 2% representative sample ($N = 401$) was randomly drawn from the census of the town of Gjøvik, stratified on the factors gender, age group, and electoral district within the town. Both the questionnaire and the interview procedures proved satisfactory, and no major problems were reported. Practical details of the survey are given by Kleiven and Thrane (1994).

Results

It may be argued that data from each item should be seen as ordinal level measurements. I have chosen, however, to treat the data as interval measurements, in line with well-established psychometric traditions (Nunnally & Bernstein, 1994).

Means

To preserve the original item metric, the mean of the completed items in each scale is used as the scale score for each person. The basic statistics of the scales are shown in Table 2.

Statistics for individual items were also computed. Means ranged from 1.54 ("Exposing your skills") to 3.42 ("Having plenty of time/time to do what you

Table 2
Statistics for Nine Leisure and Travel Motive Scales ($N = 398$)

	Cronbach's Alpha	Mean	SEM	SD	Skewness	SE of Skewness	Kurtosis	SE of Kurtosis
Sun/warmth	0.69	2.47	0.03	0.68	-0.01	0.12	-0.63	0.24
Accomplishment ^a	0.76	2.00	0.03	0.68	0.45	0.12	-0.46	0.24
Family	0.71	2.96	0.04	0.73	-0.60	0.12	-0.27	0.24
Friends ^a	0.59	2.78	0.03	0.60	-0.27	0.12	-0.25	0.24
Culture ^a	0.72	2.10	0.04	0.70	0.39	0.12	-0.52	0.24
Nature	0.70	2.59	0.03	0.69	-0.23	0.12	-0.54	0.24
Peace/quiet ^a	0.64	3.00	0.03	0.66	-0.59	0.12	0.30	0.24
Fitness	0.81	2.39	0.04	0.73	0.14	0.12	-0.67	0.24
Indulgence	0.51	2.61	0.03	0.57	-0.20	0.12	0.04	0.24

^a"Replicates" of the four Beard and Ragheb (1983) scales.

please”). The highest skewness (1.39) was found with “Exposing your skills.” Two more items had a skew larger than 1.0: “Having time for the family” with -1.19, and “Having plenty of time/time to do what you please” with -1.16. Seven items had a kurtosis above 1.0, with 1.17 being highest (“Exposing your skills”).

It may also be noted that small, but significant, gender differences existed on four of the scales, with females scoring higher. The dimensions more important to women were Family [$t(396) = 7.56, p < 0.001$], Friends [$t(396) = 2.90, p < 0.001$], Culture [$t(396) = 2.40, p < 0.02$], and Fitness [$t(396) = 2.72, p < 0.01$].

Reliability

Seven of the nine scales had a fairly good internal consistency, yielding Cronbach’s alphas above 0.60. For five of the scales, alpha was even at 0.70 or higher. Two out of the nine scales, however, (Friends and Indulgence) had an alpha of less than 0.60. The alpha values are included in Table 2.

For the Friends and Indulgence scales, no single item was apparently causing the problem; alpha will not increase upon the removal of any one item. In the Nature scale, however, item 19 (“Feeling the smell of the salty sea”) was a problem. If this item was removed, the alpha of the Nature increased from 0.70 to 0.84.

It should also be noted that the Sun/Warmth scale contains one item (item 36: “Swim in clean water”), which made little difference to the alpha score. Al-

pha was much the same whether or not the item was included in the scale. The Fitness scale also had a similar weak point. Here, alpha remained largely unchanged upon removal of item 20, “Taking care of your health.”

Confirmatory Factor Analysis of Individual Scales

The measurement models for each of the nine scales will first be examined separately. The LISREL 8.30 program was used throughout the analyses, with the “standard” Maximum Likelihood procedure (Jöreskog & Sörbom, 1993). “Congeneric” measurement models were employed, without constraints on lambda or delta components (Pedhazur & Schmelkin, 1991).

The central results, summarized in Table 3, are somewhat complex. While more than half of the 36 factor loadings were statistically significant, some scales have four significant loadings and others have none. Also, the proportion of explained variance ranged from 22% to 52%.

Among the “stand-alone” indices recommended by Hoyle and Panter (1995), chi-square suggested inadequate model fit ($p < 0.05$) on six scales. However, the Goodness of Fit Index (GFI) was high (>0.90). The “Type-2” index NNFI (Nonnormed Fit Index) was also strong (>0.90) for all scales except Family and Friends. The only type-3 index produced by LISREL, the Comparative Fit Index (CFI), was also generally positive (>0.90), also here with the Family scale as an exception.

Table 3
Completely Standardized Factor Loadings (λ), Variance Explained, and Fit Indices in Nine Separate Factor Models

	Item No.				V.E.	χ^2	df	p	GFI	NNFI	CFI
	a	b	c	d							
Sun/warmth (N = 392)	0.65*	0.66	0.65*	0.46	37%	12.25	2	0.00	0.98	0.88	0.96
Accomplishment (N = 390)	0.75*	0.62*	0.62*	0.67*	44%	15.71	2	0.00	0.98	0.89	0.96
Family (N = 378)	0.57*	0.61*	0.78	0.53	40%	64.92	2	0.00	0.92	0.46	0.82
Friends (N = 393)	0.54	0.47	0.47	0.59	27%	10.13	2	0.01	0.99	0.82	0.94
Culture (N = 393)	0.58*	0.56*	0.62*	0.74*	40%	11.68	2	0.00	0.99	0.90	0.97
Nature (N = 392)	0.72*	0.86*	0.15	0.81*	49%	0.51	2	0.78	1.00	1.01	1.00
Peace/quiet (N = 388)	0.61	0.56	0.61*	0.47	32%	1.89	2	0.39	1.00	1.00	1.00
Fitness (N = 392)	0.76*	0.72*	0.59*	0.80*	52%	20.01	2	0.00	0.98	0.90	0.97
Indulgence (N = 391)	0.46	0.48	0.54*	0.38	22%	0.33	2	0.85	1.00	1.06	1.00

* $p < 0.05$

Clearly, the fit between models and data is far from perfect. In particular, the generally high chi-square values suggest that improvements may be possible.

Also, the Indulgence and Friends scales both cover less than 30% of the variance, perhaps consistent with the reliability problems seen in Table 3.

The Family scale yielded particularly unconvincing fit measures, even if the Goodness of Fit Index was acceptable (0.92). The chi-square value [$\chi^2(2) = 64.92, p = 0.00$] indicated a poor fit, as did the NNFI (0.46) and the CFI (0.82). Nevertheless, the model explained 40% of the variance.

It should be noted, however, that no major modification was needed to achieve a better fit with the Family scale. By allowing the error terms of two items ("Having time for the family" and "See to it that the children have a pleasant vacation") to correlate, the GFI jumped to 0.99. The NNFI also became clearly acceptable (0.94), as did the CFI (0.99). Most indices thus suggested a quite good fit between the revised measurement model and the data. The chi-square statistic was still significant, however [$\chi^2(1) = 4.25, p < 0.04$], and the variance explained showed no appreciable change (39%).

In all the analyses reported in this section, data were treated as continuous measurements. However, additional analyses were also performed. By pre-processing the data with PRELIS (du Toit, du Toit, Jöreskog, & Sörbom, 1999; Jöreskog & Sörbom, 1988), and using the (default) ordinal level measurement, a polychoric covariance matrix is obtained. CFA analyses on this matrix yield similar results, but with consistently "weaker" model fit than what is reported above.

Taken as a whole, then, this first attempt at identifying measurement models for nine independent dimensions of leisure motivation through CFA gave rather encouraging results. For the three scales of Nature, Peace/Quite, and Indulgence, the simple "congeneric" model fit the data extremely well. For five scales (Sun/Warmth, Accomplishment, Friends, Culture, and Fitness), most "fit indices" showed an acceptable fit between model and data. For these scales, however, significant chi-square values indicated a need for further improvement. Only for the one remaining scale (Family), the simple measurement model apparently did not give an adequate representation of the data.

CFA of Combined Four-Scale Model (Beard & Ragheb Conceptual Replication)

We now turn to the task of combining four scales. Do the data from the four scales of Accomplishment, Friends, Culture, and Peace/Quiet support a *combined* measurement model, replicating Beard and Ragheb's (1983) four dimensions of leisure motivation?

Keeping in mind that Beard and Ragheb (1983) reported four highly correlated factors, factor intercorrelations were also allowed in the present confirmatory factor analysis. The resulting fit was not perfect. Chi-square was large [$\chi^2(98) = 308.53, p = 0.00$], but the GFI was at 0.90. The NNFI was at 0.81, the Incremental Fit Index (IFI) at 0.85, and the CFI at 0.84. Clearly, there is more room for improvement. {A combined measurement model with four *independent* (uncorrelated) factors had even less support in our data. Chi-square was very large [$\chi^2(104) = 531.85, p = 0.00$]. The GFI was 0.84, NNFI 0.62, the IFI 0.68, and the CFI 0.67.}

Only seven minor model modifications were needed, however, to achieve a better fit with the data. In the modified model shown in Figure 1, additions to the original measurement model are shown with dotted lines. Four error covariations were allowed, and three items were allowed to have loadings on two factors.

With these changes, Chi-square was still high [$\chi^2(91) = 181.07, p = 0.00$]. The other fit indices, however, suggested an acceptable fit between model and data (GFI = 0.94, NNFI = 0.90, IFI = 0.93, and CFI = 0.93).

Combined Nine-Scale Model

Initial efforts to combine all nine scales into one model were not very successful, and no acceptable model was found within a reasonable range of modifications.

Trying to find a better model, it eventually became clear that the Indulgence scale did not go well with the other eight scales. Firstly, its high correlation with the Peace/Quiet scale (0.93) indicated that it made little sense to distinguish between the two. Secondly, two items of this scale (items 23 and 35) actually loaded higher on Culture than on Indulgence.

Bearing in mind that the Indulgence scale also had a rather low reliability and explained less than one quarter of the variance, the scale and all its items were dropped from further analyses.

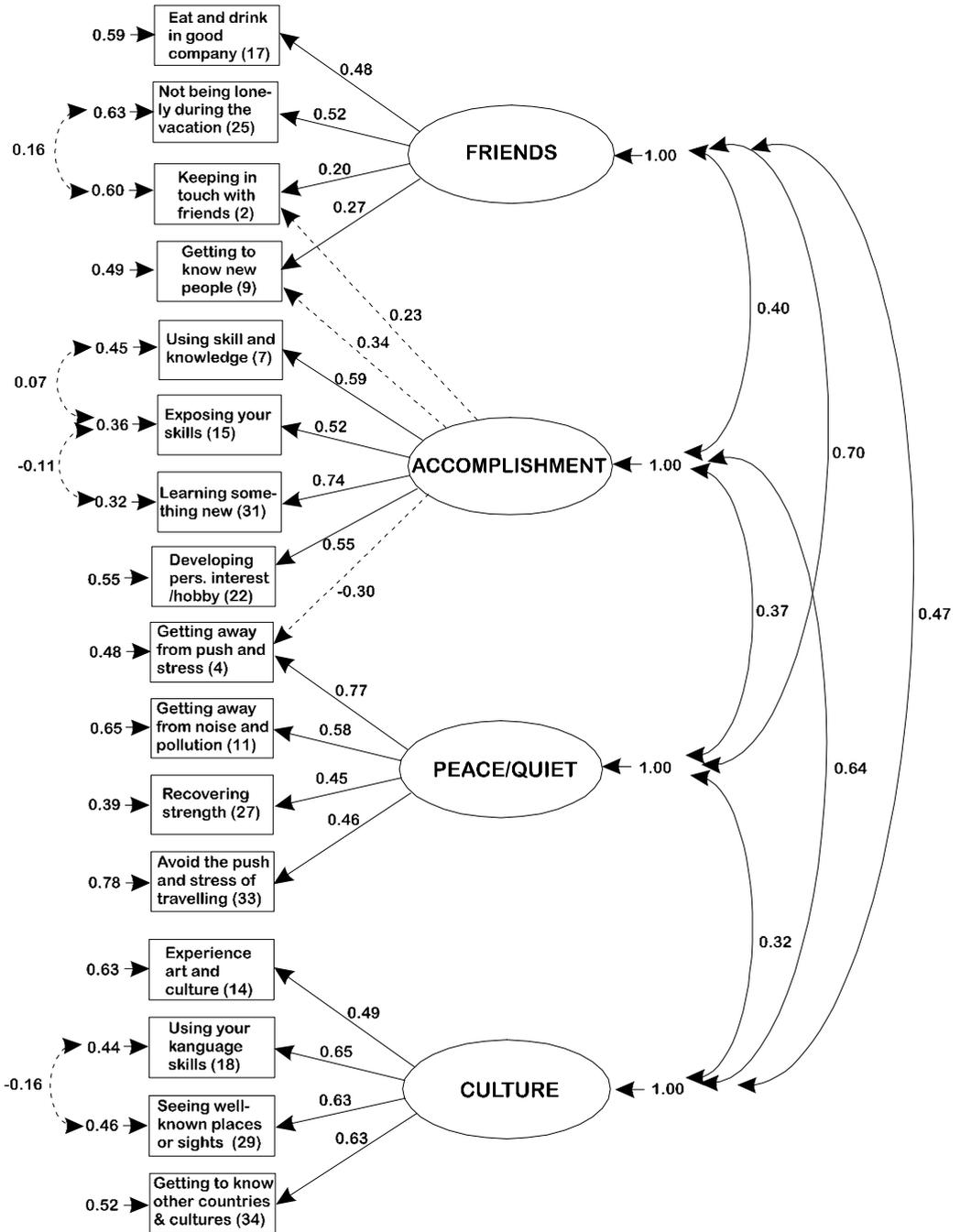


Figure 1. Improved measurement model for four scales, replicating Beard and Ragheb (1983). Standardized estimates (N = 357).

Combined Eight-Scale Model

Working with the eight remaining scales, however, modeling results looked somewhat more interesting. Combining the eight four-item scales into a model with correlated factors, the fit was still not good. It was close enough, however, for accepting the model as a starting point for a more detailed evaluation. While chi-square was very high [$\chi^2(436) = 1364.39, p = 0.00$], other fit indices looked a bit more promising (GFI = 0.81, NNFI = 0.75, IFI = 0.78, and CFI = 0.78).

If assuming noncorrelated factors, however, the fit was considerably worse [$\chi^2(464) = 2505.75, p = 0.00$; GFI = 0.69, NNFI = 0.57, IFI = 0.60, and CFI = 0.60]. The difference between the two models was substantial [$\chi^2(28) = 1141.36, p = 0.00$], indicating that the model with correlated factors gave a much better fit with the data. Still, this model clearly was no adequate representation of the data, and a closer look was needed.

Because items 19, 20, and 36 were shown not to contribute much to scale reliabilities, they were first excluded from the analysis. This gave some improvement [$\chi^2(349) = 1041.37, p = 0.00$; GFI = 0.83, NNFI = 0.79, IFI = 0.82, and CFI = 0.82], but the resulting model still was not a satisfactory representation of the data.

It took only a few additions to the model, however, to have a better fit. The improved model in Figure 2 is an example of this. Of course, chi-square for this model was too high [$\chi^2(337) = 637.42, p = 0.00$]. But the other indices were acceptable or very close to acceptable (GFI = 0.89, NNFI = 0.89, IFI = 0.91, and CFI = 0.91). Additions to the original measurement model are shown by dotted lines. For reasons of readability, theta coefficients are displayed in Table 4, not in the figure.

Looking at Table 4, we see that the magnitude of theta coefficients (interfactor correlations) varied a great deal. Some were very high, like the correlations Accomplishment/Fitness, and Family/Friends. Several others, however, were very close to 0.0.

Validity of Scales

In Table 5, seven types of leisure and travel behavior are shown, selected to illustrate the range of behavior differences in this field.

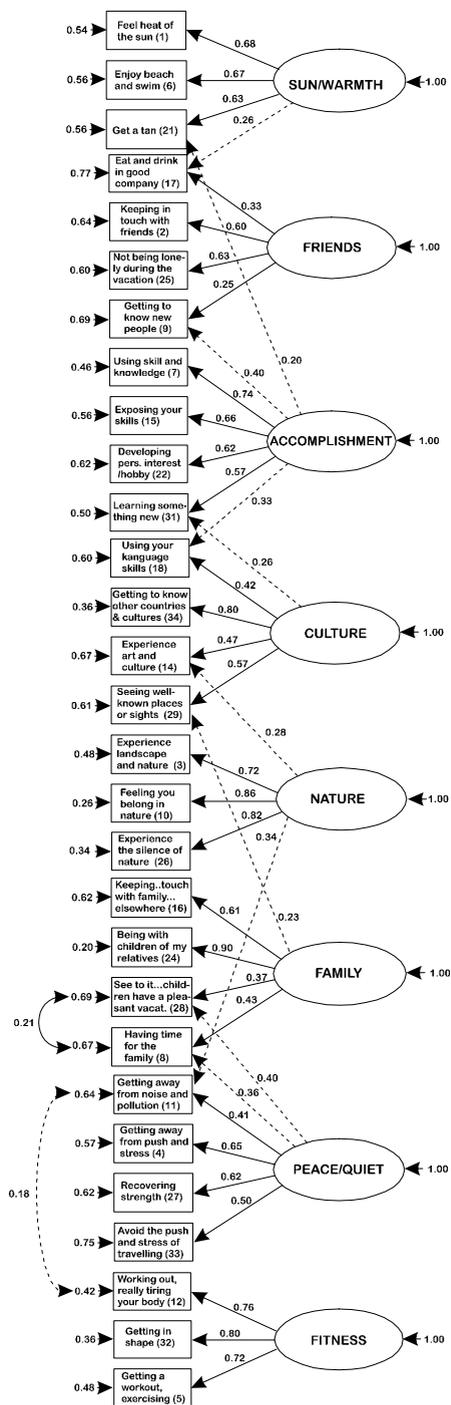


Figure 2. Improved measurement model for eight scales. Standardized estimates (N = 357).

Table 4
Interfactor Correlations (ϕ Coefficients) in Eight-Factor Model ($N = 357$)

	Sun/Warmth	Accomplishment	Family	Friends	Culture	Nature	Peace/Quiet	Fitness
Sun/warmth	1.00							
Accomplishment	0.05	1.00						
Family	-0.04	0.36	1.00					
Friends	0.30	0.47	0.54	1.00				
Culture	0.32	0.40	0.06	0.32	1.00			
Nature	-0.06	0.48	0.33	0.17	0.10	1.00		
Peace/quiet	0.41	0.11	0.06	0.52	0.20	0.31	1.00	
Fitness	-0.02	0.70	0.41	0.42	0.20	0.65	0.09	1.00

Clearly, most of these behaviors are correlated with more than one scale. Because some of the scales are also highly correlated, there may be a collinearity problem here. Furthermore, there is reason to believe that even more influences may exist; we have already seen that the scale scores are affected by gender.

Therefore, the effect of the scales should probably not be estimated separately from that of other predictors. This calls for the use of multiple regression models for assessing the validity of the scales, including demographic variables among the predictors.

Some examples of this are given in Table 6. In these regressions, gender is a dummy variable, using the score of 0 for men and 1 for women. Positive betas for gender on an activity thus indicate a relatively higher participation by women.

Although the resulting picture is fairly complex, one central observation should be made. On all behaviors, one or more scales do contribute significantly to the regression. Even in the face of other potent factors, the scales do retain some predictive

power. The strength of this relationship varies considerably between the different activities, however.

Discussion

Scale Properties

With the present sample, seven of the four-item scales had alpha values within an acceptable range, which indicated a satisfactory internal consistency. For two scales, however, alpha scores were below 0.60, suggesting weak internal consistency.

Here one must keep in mind that the alpha statistic is sensitive to the scale's number of items. Given the same average item intercorrelation, longer scales will have a higher alpha. Using the Spearman-Brown prophecy formula (Nunnally & Bernstein, 1994), we find, for example, that an alpha value of 0.57 in the four-item scales indicates an average item intercorrelation of 0.25, which would have given an alpha value of 0.80 on Beard and Ragheb's (1983) 12-item scales.

Table 5
Correlations Between Scales and Selected Leisure Behaviors ($N = 398$)

	Taking a Walk	Running, Jogging	Renting a Video	Going to Art Exhibition	Fish or Hunt in Season	Games and Play at Beach/Pool	Short Drive With Family
Sun/warmth	-0.04	0.02	0.15**	0.07	-0.03	0.31***	-0.01
Accomplishment	0.01	0.03	0.09	0.11*	0.05	-0.12*	0.08
Family	0.08	-0.17***	-0.23***	0.08	-0.10*	0.03	0.28***
Friends	0.03	0.02	0.21***	0.01	-0.07	0.12	0.09
Culture	0.07	-0.04	-0.02	0.37***	-0.10*	0.04	0.07
Nature	0.28***	-0.07	-0.25***	0.25***	0.16***	-0.05	0.15**
Peace/quiet	0.12*	-0.05	-0.02	0.007	0.11*	0.16***	0.09
Fitness	0.24***	0.08	-0.19***	-0.01	0.01	-0.12*	0.09

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6

Multiple Regression by the ENTER Method. Beta Weights of Motive Scales and Other Predictors of Leisure Behaviors ($N = 326-327$)

	Taking a Walk	Running, Jogging	Renting a Video	Going to Art Exhibition	Fish or Hunt in Season	Games and Play at Beach/Pool	Short Drive With Family
Sun/warmth	-0.08	-0.04	0.05	0.00	-0.05	0.28***	-0.03
Accomplishment	-0.19**	-0.00	0.14**	0.01	0.10	-0.19**	0.07
Family	-0.06	-0.07	-0.15**	0.07	0.04	0.18**	0.17**
Friends	0.03	0.01	0.20**	-0.04	-0.10	-0.01	-0.01
Culture	0.03	-0.02	-0.09	0.31***	-0.10	0.08	0.03
Nature	0.23***	-0.07	-0.06	0.26***	0.27***	-0.03	0.08
Peace/quiet	0.01	-0.03	-0.14**	-0.17**	0.00	0.00	0.09
Fitness	0.21**	0.36***	0.01	-0.20**	-0.02	0.06	-0.12
Gender	0.24***	-0.14**	-0.08	0.10*	-0.31***	-0.12*	0.11
Age group	0.05	-0.33***	-0.59***	0.15**	-0.21***	-0.41***	0.15*
Education level	0.10	0.23***	0.04	0.27***	-0.05	0.12*	-0.04
Household income	0.03	0.08	-0.07	0.20***	0.07	0.07	-0.02
Daily TV hours	0.03	-0.00	0.12**	0.03	0.13	-0.17***	-0.01
Adjusted R²	0.18	0.28	0.51	0.35	0.19	0.34	0.12

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Using only four response categories (“Not important,” “A little important,” “Important,” and “Very important”) was probably not a wise decision. With this response format, the variability on each item is of course limited, and variance of the summed scales will suffer (DeVellis, 1991). Both reliability and validity may of course be adversely influenced by this.

The Indulgence scale was dropped during the analysis, partly because of its low reliability. Taking the limited scale length into consideration, however, all eight remaining scales had alpha scores that were comparable to scales that have previously been used in this field. I view this as a satisfactory first result.

The summed scales also have other desirable properties. Clearly, the scale means were not too close to the ends of the scale. This increases the scales' chance of functioning properly, without ceiling or floor effects.

Using LISREL models, we gain more detailed information on the eight individual scales. Seven of the eight single-scale measurement models may be seen as satisfactory. On the Family scale, however, some interitem covariance was not accounted for.

Four-Scale Combined Model

Combining four of the scales into a conceptual replication of Beard and Ragheb's four scales, some

modifications were needed to obtain a satisfactory measurement model. Four within-factor item covariances had to be allowed, and three items were found to load on more than one factor.

It should also be noted that two factor intercorrelations (Culture*Accomplishment and Friends*Peace/Quiet) were noticeably higher than others. This may suggest a model with two higher-order factors: one implying a more relaxing and unpretentious group of motives than the other.

In my view, however, the modified model should not be viewed as a new, *alternate* model. Rather, it should be read as a suggestion that the original model really is not too far from its target. The modifications applied do not invalidate the central properties of the measurement model; the main factor structure is upheld, and the added covariances are all *within* factors.

Of course, neither the original model nor the modified one may be proven right or wrong. The modified model may serve, however, as an example of the *type of model* needed to combine the four Beard and Ragheb-inspired scales. Correlated factors should be assumed, possibly with higher-order factors, and some of the items will have substantial loadings on more than one factor. And, quite likely, improvements may be made in the choice of items in some of the factors.

Eight-Scale Combined Model

Much the same conclusions may apply to the combined eight-factor model. The original measurement model was not clearly supported by the data. Using an improved model as an example, however, we see that freeing a limited number of constraints helped. Allowing one new interitem covariance and letting 10 out of 32 items load on two factors, acceptable fit between model and data was obtained.

The unequal interscale correlations may suggest that some factor correlations should *not* have been included in the model. And the combined pattern may indicate that models with second-order factors should be considered, replacing the present model that has free covariation between all factors.

Like the four-factor improved model, the eight-factor one is not proven right or wrong. The main factor structure is upheld, however. And, more importantly, most modifications are meaningful or understandable additions that do not invalidate or change the *meaning* of the eight scales. In my view, this suggests that the original measurement model is not far from its target. The question of higher-order factors is an interesting challenge for future work, but is not likely to disturb the underlying distinction between eight dimensions.

The most important aspect of the eight-factor model, however, may be its *inclusion* of the four motive dimensions used by Beard and Ragheb. Using these four scales as part of a more comprehensive motive model, some comparability with previous work is maintained, without excluding the addition of new motive dimensions. The fact that four of the “new” dimensions may coexist with the Beard and Ragheb dimensions is therefore very encouraging.

Scale Validity

Interesting correlations do exist between every scale and some type of behavior. Correlations are not by far as high as one would traditionally like to see for test validity figures, but do suggest that some scales may be useful in predicting certain behaviors.

Further work on scale validity is needed, but the tables with selected leisure activities may serve to illustrate some points that repeatedly emerged during analyses of this material. The first point may indeed

be trivial; different scales are relevant predictors for different activities. No one scale is *generally* predictive across a range of activities; each scale predicts a limited set of activities. The same seems to hold for the demographic predictor variables; while a variable may be relevant to some activities, no one variable is a generally valid predictor.

And, conversely, different activities have different sets of predictors. While one set of predictor variables significantly contributes to the regression equation for one activity, other sets will be found in the equations for others.

In view of this complexity, it should be borne in mind that tourists commonly engage in several activities at the same time. Acting multioptionally, a person may swim, be social, play games, and go fishing within the same vacation. Given that these activities have different predictors, a single consistent motive model may prove insufficient for predicting the person's *set of choices or activities*.

Nonetheless, the motive scales do have some predictive power when seen in conjunction with other predictors. Because the power of the scales often is dependent on the state of other variables, rather complex methods may be needed to adequately assess their validity.

Challenges for Further Work

Sociodemographic differences are common in this kind of research. The observed gender differences in scale means should be seen as an example of this, and are only to be expected. They do support, however, the decision to use a representative sample for this work.

But this sample, representative for a small town, may nevertheless prove inadequate in a larger context. There is no reason to believe that sociodemographic differences are limited to gender. To have an adequate base for conclusions about leisure and travel motives in Norway, it is therefore highly desirable to replicate the present study in other communities. And, of course, a nationally representative survey would be very interesting.

At any rate, certain modifications are likely to be needed. Increasing the number of response categories on each item is one possible improvement. Replacing or removing items is another; one full scale and a few items seem not to function adequately.

More specifically, two out of three “weak” items (“Feeling the smell of the salty sea” and “Swim in clean water”) were clearly related to the sea. Having done this first survey in an inland town, insufficient experience with the seaside in the sample may be a problem for the use of these items. Therefore, both replacing the items and replicating the study in a coastal town should be considered.

Summing Up

A small but definite first step towards measuring Norwegian leisure and travel motives has been made. Psychometric properties of eight scales are promising, LISREL models seem not to contradict the basic dimensionality of the implied measurement model, and scales do have some power in predicting leisure activities.

Most importantly, however, challenges for future work have been identified. To have a broader population represented in the data, replications should be made. Some individual scale items may be changed or rewritten, and measurement models should be replicated. Last but not least, more work on validity is called for.

Looking back at the four research questions, then, some tentative answers may be offered:

1. The four leisure motivation scales did replicate in Norway.
2. Adding five more factors yielded a plausible eight-factor motives model. However, a nine-factor model was not supported by the data.
3. Motive factors were correlated in both four- and eight-factor models.
4. The scales did have some power in predicting leisure behavior. Rather complex relationships with other predictors existed, however, and more advanced analyses are needed to correctly assess scale validity.

For analyses of this type, our representative sample is an advantage. Limiting the scale length to four items is possible, but this of course attenuates alpha values.

Acknowledgments

This research was supported by the Tourism Research Program of the Norwegian Research Coun-

cil, through a grant to the author’s project Motive Dimensions as Demand Predictors in Norwegian Leisure and Tourism. In the early stages of this work, Senior Lecturer Christer Thrane was a very constructive critic.

References

- Aasetre, J., Kleiven, J., & Kaltenborn, B. (1994). *Friluftsliv i Norge—Motivasjon og atferd [Outdoor recreation in Norway—motivation and behaviour]* (Oppdragsmelding No. 309). Trondheim: Norsk institutt for naturforskning.
- Beard, J. G., & Ragheb, M. G. (1983). Measuring leisure motivation. *Journal of Leisure Research*, 15(3), 219–228.
- Cha, S., McCleary, K. W., & Uysal, M. (1995). Travel motivations of Japanese overseas travelers: A factor-cluster segmentation approach. *Journal of Travel Research*, 33–39.
- Crandall, R. (1980). Motivations for leisure. *Journal of Leisure Research*, 12(1), 45–55.
- Crompton, J. (1979). Motivations for pleasure vacation. *Annals of Tourism Research*, 6(4), 408–424.
- DeVellis, R. F. (1991). *Scale development. Theory and applications*. Newbury Park, CA: Sage.
- Driver, B. L. (1977). *Item pool for scales designed to quantify the psychological outcomes desired and expected from recreation participation* (unpublished paper).
- du Toit, S., du Toit, M., Jöreskog, K. G., & Sörbom, D. (1999). *Interactive LISREL*. Chicago: Scientific Software International.
- Dunn Ross, E. L., & Iso-Ahola, S. E. (1991). Sightseeing tourists’ motivation and satisfaction. *Annals of Tourism Research*, 18, 226–237.
- Fontaine, G. (1993). Motivational factors of international travelers. *Psychological Reports*, 72, 1106.
- Graefe, A. R., Ditton, R. B., Roggenbuck, J. W., & Schreyer, R. (1981). Notes on the stability of the factor structure of leisure meanings. *Leisure Sciences*, 4(1), 51–65.
- Haldorsen, T. (1981). *Norske ferieformer [Norwegian vacation types]* (Report No. 81/25). Oslo: Statistisk sentralbyrå.
- Hall, C. M., & Page, S. J. (1999). *The geography of tourism and recreation*. London and New York: Routledge.
- Harrill, R., & Potts, T. D. (2002). Social psychological theories of tourist motivation: Exploration, debate, and transition. *Tourism Analysis*, 7, 105–114.
- Haukeland, J. V. (1993). *Den norske feriedrømmen [The Norwegian vacation dream]* (TØI report 194). Oslo: Transportøkonomisk institutt.
- Haukeland, J. V., Nymoen, M., & Rideng, A. (1991). *Ferieundersøkelsen 1991 [The vacation survey 1991]* (TØI report 105/91). Oslo: Transportøkonomisk institutt.
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling. Concepts, issues, and applications* (pp. 158–176). Thousand Oaks, CA: Sage Publications.
- Jamal, T., & Lee, J. (2003). Integrating micro and macro approaches to tourist motivations: Toward an interdisci-

- plinary theory. *Tourism Analysis*, 8, 47–59.
- Jamroz, U., & Uysal, M. (1994). Travel motivation variations of overseas German visitors. *Journal of International Consumer Marketing*, 6(3/4), 135–160.
- Jöreskog, K. G., & Sörbom, D. (1988). *PRELIS. A program for multivariate data screening and data summarization* (2nd ed.). Chicago: Scientific Software.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8 user's reference guide*. Chicago: Scientific Software International.
- Kelly, J. R. (1990). *Leisure*. Englewood Cliffs, NJ: Prentice Hall.
- Kim, C. (1999). Cross-cultural perspectives on motivation. *Annals of Tourism Research*, 26(1), 201–204.
- Kleiven, J. (1998a). *En vanskelig forenkling—den komplekse norske feriedrømmen [A difficult simplification—the complex Norwegian vacation dream]* (Research Report No. 34/1998). Lillehammer: Lillehammer College.
- Kleiven, J. (1998b). *Skalaer for måling av aktivitets-og motivmønstre i en norsk lokalbefolknings ferie og fritid [Scales for measuring activity and motive patterns in a Norwegian local population's leisure and vacations]* (Working paper No. 77/1998). Lillehammer: Lillehammer College.
- Kleiven, J. & Thrane, C. (1994). *Reiselivsundersøkelsen i Gjøvik 1994—en metoderapport [The vacation travel survey in Gjøvik 1994—a report of methods]* (Working paper No. 28/1994). Lillehammer: Lillehammer College.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Lee, G., O'Leary, J.T., Lee, S. H., & Morrison, A. (2002). Comparison and contrast of push and pull motivational effects on trip behavior: An application of a multinomial logistic regression model. *Tourism Analysis*, 7, 89–104.
- Loker-Murphy, L. (1996). Backpackers in Australia: A motivation-based segmentation study. *Journal of Travel and Tourism Marketing*, 5(4), 23–45.
- Lounsbury, J. W., & Franz, C. P. G. (1990). Vacation discrepancy: A leisure motivation approach. *Psychological Reports*, 66, 699–702.
- Middleton, V. T. C. (1994). *Marketing in travel and tourism*. Oxford: Butterworth-Heinemann.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Pearce, P. L. (1995). Pleasure travel motivation. In R. W. McIntosh, C. R. Goeldner, & J. R. B. Ritchie (Eds.), *Tourism: Principles, practices, philosophies* (pp. 167–190). New York: John Wiley & Sons.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis*. New Jersey: Erlbaum.
- Ryan, C. (1993). Correlations between clusters based on the Ragheb and Beard Leisure Motivation Scale and the attributes of desired holiday locations. In: *Proceedings of Third International Conference of the Leisure Studies Association*. UK: Loughborough University.
- Ryan, C. (1994). Leisure and tourism—the application of leisure concepts to tourist behavior—a proposed model. In: A. V. Seaton (Ed.), *Tourism: The state of the art* (pp. 294–307). Chichester: John Wiley & Sons.
- Ryan, C. (1997). Similar motivations—diverse behaviours. In C. Ryan (Ed.), *The tourist experience* (pp. 25–47). London: Cassell.
- Ryan, C., & Glendon, I. (1998). Application of leisure motivation scale to tourism. *Annals of Tourism Research*, 25(1), 169–184.
- Schmidhauser, H. (1989). Tourist needs and motivation. In: S. F. Witt, & L. Moutinho (Eds.), *Tourism marketing and management* (pp. 569–572). Hemel Hempstead, UK: Prentice-Hall.
- Seaton, A. V. (1994). Are relatives friends? Reassessing the VFR category in segmenting tourism markets. In A. V. Seaton (Ed.), *Tourism: The state of the art* (pp. 316–321). Chichester: John Wiley & Sons.
- Thrane, C. (1996). *Fritiden og dens differensiering [Leisure and its differentiation]* (Research report No. 10). Lillehammer: Lillehammer College.
- Thrane, C. (1997). Norsk Monitors segmenteringsprosjekt: Et alternativ til sosio-økonomisk segmentering? [Norwegian Monitor's segmentation project: An alternative to socio-economic segmentation?]. *Sosiologisk Tidsskrift*, 3, 217–234.
- Tinsley, H. E. A. (1984). The psychological benefits of leisure counseling. *Society and Leisure*, 7, 125–140.
- Tinsley, H. E. A., & Kass, R. A. (1979). The latent structure of the need satisfying properties of leisure activities. *Journal of Leisure Research*, 11(4), 278–291.
- Vaagbø, O. (1993). *Den norske turkulturen [The Norwegian hiking culture]*. Oslo: Friluftslivets fellesorganisasjon.