

UNIVERSITY OF WAIKATO

**Hamilton
New Zealand**

Understanding Stakeholder Values

Using Cluster Analysis

Pamela Kaval

Department of Economics

Working Paper in Economics 07/16

August 2007

Pamela Kaval

Economics Department
University of Waikato
Private Bag 3105
Hamilton, New Zealand

Tel: +64 (7) 838 4045
Fax: +64 (7) 838 4331
Email: pkaval@mngt.waikato.ac.nz
pam98k@yahoo.com

Abstract

The K-Means and Ward's Clustering procedures were used to categorize value similarities among respondents of a public land management survey. The clustering procedures resulted in two respondent groupings: an anthropocentrically focused group and an ecocentrically focused group. While previous studies have suggested that anthropocentric and ecocentric groups are very different, this study revealed many similarities. Similarities between groups included a strong feeling towards public land and national forest existence as well as the importance of considering both current and future generations when making management decisions for public land. It is recommended that land managers take these similarities into account when making management decisions. It is important to note that using the Ward's procedure for clustering produced more consistent groupings than the K-Means procedure and is therefore recommended when clustering survey data. K-Means only showed consistency with datasets of over 500 observations.

Keywords

Cluster analysis
Ward's hierarchy method
K-Means
public land management
stakeholders
ecocentric
anthropocentric

JEL Classification

C4, Q2

Acknowledgements

I would like to thank the Forest Service for use of the survey data. I would also like to thank Lane Eskew and Dr. Rudy King for their assistance on an earlier version of this paper. Finally, I would like to thank Thomas Wilding for his editing assistance on the current version of the paper.

INTRODUCTION

The National Environmental Policy Act (1969) requires public participation in management decisions; public values are therefore a driving force for public land management decisions (Beckley and Korber, 1995). The values of an individual represent that person's fundamental, biological, and social needs; these needs are used in decision-making. A person's values influence, and are influenced by, the cultural, social, institutional, and economic framework in which they live (Shultz and Zelezny, 1999). Once values are formed, they are difficult, but not impossible to change. Changes can happen through various forms of education such as personal experiences or learning from newspaper and television media sources (Bengston et al., 2001).

The values of a segment or collection of the population is considered to be a social or cultural value. Two general social value categories are commonly observed in natural resource studies: (1) an anthropocentric or instrumental category (ANTHRO) and (2) an ecocentric, biocentric, non-instrumental, or intrinsic category (ECO) (McFarlane and Boxall, 2000; Jacobson and Marynowski, 1997; Steger and Witt, 1990; Kearney et al., 1999).

People placed in an ANTHRO group typically value forests for services or products that satisfy their wants and needs. They generally support issues related to mining, logging, and off-highway vehicle use (Thompson et al., 1994; Bengston, 1994; Manning et al., 1999; Shultz and Zelezny, 1999). The ANTHRO respondents consist of a high percentage of men, have attended some university or obtained a university degree, and have a high percentage of respondents that depend on forests for their livelihood (McFarlane and Boxall, 2000; Jacobson and Marynowski, 1997; Steger and Witt, 1990; Kearney et al., 1999).

People in the ECO group, on the other hand, tend to value forests for the intrinsic value and not necessarily for the value they provide to people. For instance, a member of an ECO group would support protecting an area merely to preserve an endangered species, even if they never saw the species or its habitat (Thompson et al., 1994; Bengston, 1994; Manning et al., 1999; Shultz and Zelezny, 1999). Studies by McFarlane and Boxall (2000), Jacobson and Marynowski (1997), Steger and Witt (1990), and Kearney et al. (1999), show ECO group respondents to generally be younger than those in an ANTHRO group and have either graduated from an undergraduate university or attended some graduate school. They reside

in urban settings, are more politically liberal, and consist of a higher percentage of women than those in an ANTHRO group.

Cluster analysis can be used to categorize people with similar values. Two commonly applied clustering methods include the Ward's minimum variance method (Ward's) and the K-Means procedure (Gil et al., 2001; Aerni, 2005; Podnar and Jancic, 2006; Lee et al, 2000; Needham et al., 2004; Nielsen and Mathiesen, 2006). The Ward's procedure uses a hierarchical clustering method in which individuals are categorized. A hierarchical method begins with each observation in its own group or cluster (i.e., 800 observations, 800 clusters). It then iteratively categorizes similar observations together. Each iteration results in a smaller number of clusters than the last, with each cluster containing more observations. This process continues, as the clusters get larger and larger, until all observations are members of one single cluster. It forms these cluster groups by maximizing the distance between clusters. The distance between clusters is based upon the analysis of variance of the sum of squares of the particular variable of interest between clusters and then adding this over all variables. The result is that in each cluster generation, the sum of squares within each cluster over all possible partitions is minimized (SAS, 1985; Milligan and Cooper, 1985).

The K-Means procedure is a non-hierarchical procedure. As opposed to Ward's, the K-Means model does not form a tree structure. It is a simple non-hierarchical clustering algorithm where Euclidian distances are calculated to locate cluster centers. In the first iteration, the program selects K observations that have large Euclidean distances from each other and designates these as tentative cluster centroids. The remaining data are then classified according to the minimum distance to centroids and centroids are recalculated. This process is repeated and centroids are recalculated until centroids do not change more than a numerical round-off in successive calculations (SAS, 1985).

In this paper, it is hypothesized that Ward's and K-Means models will have consistent results and will result in useful groupings of survey respondents that can be used by public land managers to help them in their decision making processes. To test this hypothesis, value and demographic data from a 1998 South-Central Colorado Survey were analyzed.

METHODS AND RESULTS

The Survey

A 13-page mail survey was designed to examine stakeholders' views of the management of public lands, specifically of national forests in South-Central Colorado. The survey consisted of five types of questions: values, objectives, attitudes, behaviors, and demographic questions. Only responses to the value questions and the demographic information were analyzed in this paper.

To get a representative sample for this study, possible respondents were selected from three mailing lists in South-Central Colorado. The mailing lists included the Colorado Mountain Club (CoMC), the Colorado Off Highway Vehicle Coalition (CoOHV), and a senior citizen group (SeniorCit) (Colorado School of Mines, 1998). The CoMC encourages the preservation of the Colorado Mountains; including aspects such as natural scenery, flowers, and forests (CoMC, 2003). The CoOHV is committed to the protection and promotion of off-highway motorized recreation throughout Colorado; they advocate and promote the responsible use and conservation of our public lands and natural resources to preserve their aesthetic and recreational qualities for future generations (CoOHV, 2003). SeniorCit is a group of citizens over the age of 55 interested in the welfare of public lands in their area, most specifically national forests.

In total, 1,112 people were mailed copies of the survey along with a return addressed stamped envelope. The only contact made with potential respondents was the initial survey mailing; non-respondents did not receive follow-ups or reminder cards. Overall, 556 usable surveys were returned for a response rate of 50%.

Clustering procedures were used to categorize respondent's answers to the 25 value questions. Since clustering procedures are sensitive to missing data (SAS, 1985), 45 survey responses were omitted from the analysis, leaving a sample size of 511 observations.

RESULTS

Atta-Boateng and Moser's (1998) research showed that to determine the number of clusters to use, the combination of the results of four tests: the cubic clustering criterion (CCC), semi-partial R squared (SPRS), Pseudo F (PF), and Pseudo T (PT) statistics, could be used. For this analysis, all four test statistics (CCC, SPRS, PF, PT) suggested there were two clusters of observations. All four of these tests were accomplished using the PROC CLUSTER procedure in SAS Version 8.0 (SAS, 2000).

To test the hypothesis that different clustering procedures produce comparable results, both Ward's and K-Means procedures were used. As suggested by previous studies (McFarlane and Boxall, 2000; Jacobson and Marynowski, 1997; Steger and Witt, 1990; Kearney et al., 1999; Thompson et al., 1994; Bengston, 1994; Manning et al., 1999; Shultz and Zelezny, 1999), cluster groups will typically fall into either an ANTHRO or ECO categorization. Therefore, variables suggested by these studies to be related to ANTHRO and ECO categorizations were reviewed to see if clusters fell into one of these groupings. Variables reviewed included, but were not limited to: sex, education, using the forest for their livelihood, supporting environmental causes, and believing a forest should exist for its own sake. If one group had a higher percentage of females, respondents were more educated, they did not depend on the forest for their livelihood, supported environmental causes and believed in forests existing for their own sake, they would be placed in an ECO group. If the reverse were true, they would be placed in an ANTHRO group.

When analyzing the entire dataset (25 value variables) from all usable surveys, two groupings resulted for both the Ward's procedure and the K-Means procedure. Variable averages for Cluster 1 and Cluster 2 were then reviewed. It was discovered that Cluster 1 for both Wards and K-Means had a higher number of females, were more educated, supported environmental causes, and felt that the forest should exist just for the sake of existing. There were also less dependent on the forest for their livelihood in Cluster 1. Cluster 1 was then categorized as the ECO group and Cluster 2 as the ANTHRO group.

Next, data was clustered in a different way. The Ward's procedure was used to cluster the CoMC variables. This resulted in Cluster 1 and Cluster 2. ANTHRO and ECO variables were then reviewed and it was determined that Cluster 1 was an ECO and Cluster 2 was an ANTHRO. Next, Ward's was used to cluster the CoOHV group. This also resulted in an

ECO and an ANTHRO grouping. Finally, Ward's was used to cluster the SeniorCit group. Again, this resulted in an ECO and ANTHRO grouping. The ECO groupings from CoMC, CoOHV, and SeniorCit were then combined to make one ECO group and the same was done to make one ANTHRO group.

Finally, this process was repeated using the K-Means procedure. Two clusters were formed for the CoMC, CoOHV, and SeniorCit groups. When the ANTHRO and ECO variables were reviewed, it seemed that both Cluster 1 and Cluster 2 from the CoMC, CoOHV, and SeniorCit groups revealed preferences from ECO groups. They had similar percentages of females, were higher educated, believed forests should exist for their own sake, did not have a high percentage of people dependent on the forest for their livelihood and had a higher percentage of people that were interested in supporting environmental causes. Clustering the smaller stakeholder groups in this manner using K-Means seemed to produce two similar ECO groups instead of two groups that were more distinct.

Because of the discrepancy with the K-Means grouping of ECO-ECO, percent agreement and kappa statistics were used to assess inter-rater reliability. Percent agreement represents the number of agreements divided by the total number of observations. The kappa statistic takes into account not only percent agreement, but also the agreement percentage that would have occurred by chance (Cohen, 1960).

The kappa statistic can be calculated by using the following equation:

$$Kappa = \frac{(p_{11} + p_{22}) - (p_{1B} * p_{1a} + p_{2B} * p_{2a})}{1 - (p_{1B} * p_{1a} + p_{2B} * p_{2a})}$$

where:

p_{11} = the percentage of respondents that are in the ECO group for Ward's and K-Means.

p_{22} = the percentage of respondents that are in the ANTHRO group for Ward's and k-Means.

p_{1a} = the percentage of respondents that are in the ECO group for K-Means.

p_{1B} = the percentage of respondents that are in the ECO group for Ward's.

p_{2a} = the percentage of respondents that are in the ANTHRO group for K-Means.

p_{2B} = the percentage of respondents that are in the ANTHRO group for Wards.

The output from percent agreement and kappa tests is a rating scale between 0 through 1, where 1 indicates complete agreement and 0 indicates no agreement (Cohen, 1960; Fleiss, 1981).

Table 1 presents the results from the ECO and ANTHRO group comparisons described previously using the two statistical methods. Results are sorted with the highest level of agreement at the top of the table and lowest level of agreement at the bottom. The highest level of agreement was achieved between Ward’s procedure, separated by stakeholder group, and Ward’s procedure using the entire dataset. The next highest agreement was Wards procedure when stakeholder groups were clustered separately and K-Means for the entire dataset. The third highest agreement was when both Ward’s and K-Means clustered on the entire dataset. All three comparisons that included K-Means clustering by each stakeholder group (CoMC, CoOHV, SeniorCit) resulted in a kappa statistic that was less than 0.2 and a general agreement statistic of less than 0.6. Therefore, the results shown in the rest of this paper refer only to averages of the ANTHRO and ECO groups. The ECO-ECO group created during the K-Means clustering by stakeholder groups was dropped from this analysis due to the inconsistency with the other groupings.

Table 1. Kappa Statistics of Agreement and General Percent Agreement for the Ward’s and K-Means Procedures.

Comparison of:		Kappa	General Agreement
Wards - Entire Dataset	Wards – Dataset Separated by Stakeholder Group	0.590	0.793
K-Means - Entire Dataset	Wards – Dataset Separated by Stakeholder Group	0.500	0.775
K-Means - Entire Dataset	Wards - Entire Dataset	0.370	0.669
K-Means – Dataset Separated by Stakeholder Group	Wards – Dataset Separated by Stakeholder Group	0.193	0.593
K-Means - Entire Dataset	K-Means – Dataset Separated by Stakeholder Group	0.117	0.544
K-Means – Dataset Separated by Stakeholder Group	Wards - Entire Dataset	0.109	0.448

Note: The data was clustered as one entire dataset and also separated by each of the three stakeholder groups (0=no agreement, 1=complete agreement).

Average demographic statistics for the ANTHRO and ECO groups were calculated (Table 2). It was found that the ecocentric group had a higher percentage of females and had typically graduated from university. The anthropocentric group consisted of a higher proportion of males and has some university coursework, but did not get a university degree. All other demographic variables were found to be similar between groups.

Table 2. Average Demographic Responses Reported for the Overall Dataset as well as for the Ecocentric (ECO) and Anthropocentric (ANTHRO) Groups

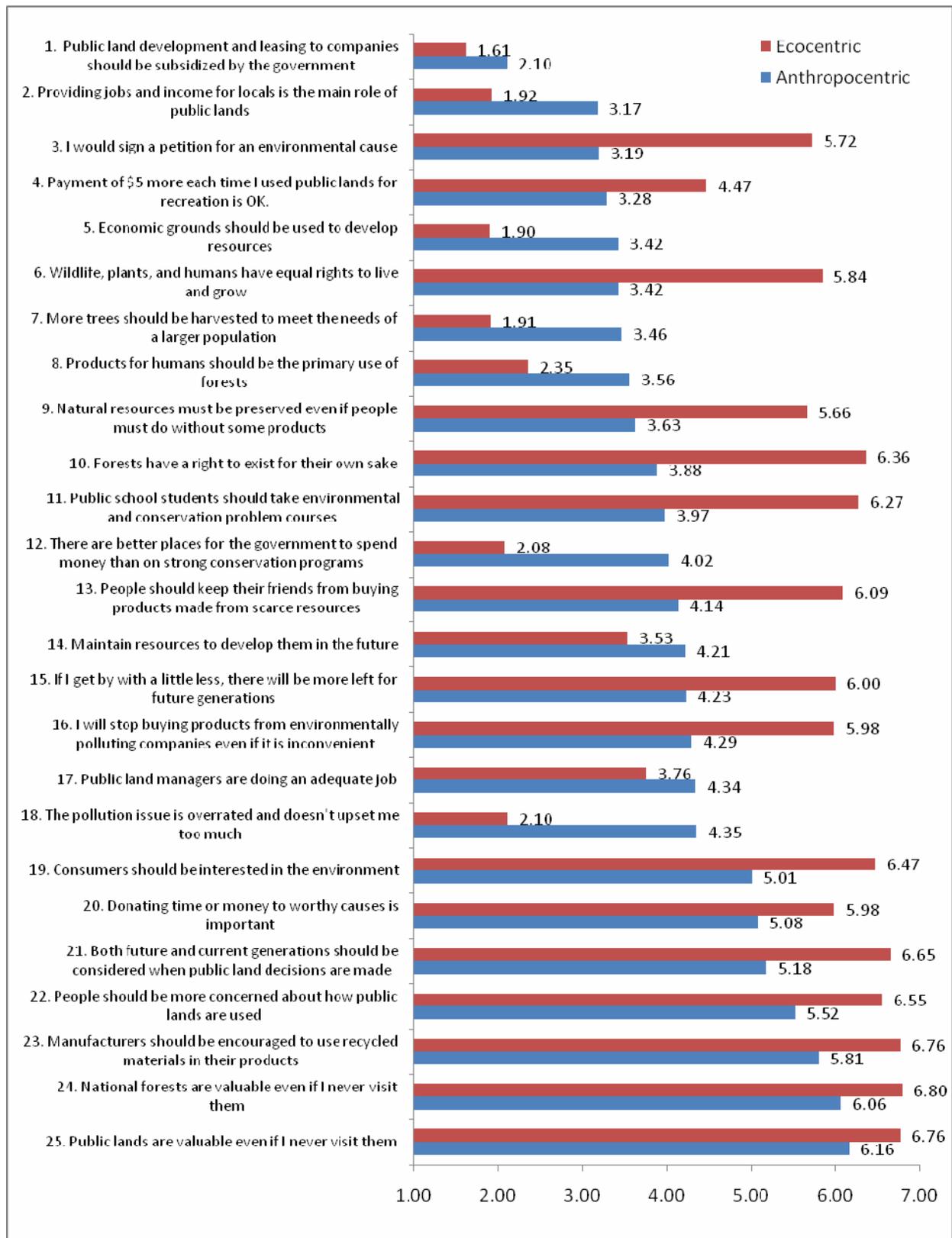
	Ecocentric	Anthropocentric	Entire Dataset
Average Age	46	45	45
Gender (Percent Female)*	37%	17%	28%
Average Education*	University Grad	Some University	University Grad
Occupation (Percentage)			
Professional/Technical	47%	45%	46%
Retired	29%	28%	29%
Self Employed	11%	13%	12%
Other	13%	14%	13%
Average Income	\$37,467	\$40,133	\$38,100
Proximity to Public Lands (Percentage)			
Less than 10 miles	65%	59%	64%
More than 10 miles	35%	41%	36%
Average number of years of residence in the area	7.67	9.17	8.00
Number of Respondents in Each Cluster	309	202	511

Note: *Indicates a statistically significant difference between the ECO and ANTHRO groups.

The average value response to survey questions was then calculated for the ECO and ANTHRO groups (Figure 1). All value questions were based on a 7-point Likert scale with 1 being “not important at all” and 7 being “very important.” Low values were considered to be from 1 to 3, neutral from 3 to 5 and important from 5 to 7.

Results show many similarities between the two groups. Both groups placed a low value on government subsidies for the leasing and development of public land (Question 1). They were neutral on issues like paying more money to use public lands (Question 4) and maintaining resources to develop public lands in the future (Question 14). Both groups felt strongly about the existence of public lands (Questions 24 and 25), encouraging manufacturers to use recycled materials in their products (Question 23), and that both future and current generations be considered in public land management decisions (Question 2). Overall, similar responses were recorded for 11 of the 25 value questions.

Figure 1. Average Ecocentric and Anthropocentric Group Responses to Value Questions
 (Based on a 7-point Likert scale with 1 = 'not important at all' and 7 = 'very important')



The greatest differences occurred among six of the value questions. The ecocentric group felt much more strongly about signing a petition for environmental causes (Question 3), that wildlife, plants, and humans have equal rights to live and grow (Question 6), that natural resources should be preserved even if people must do without some products (Question 9), that forests have a right to exist for their own sake (Question 10), and that public school students should take environmental and conservation courses (Question 11). There was one question that ECO people did not feel strongly about, but the ANTHRO group ranked as neutral; this was the statement that the pollution issue is overrated and does not upset them much (Question 18).

DISCUSSION AND CONCLUSIONS

In this paper, the hypothesis that various clustering models and methods will have consistent results was tested. To accomplish this, two clustering procedures were used: Ward's hierarchy method and the K-Means procedure. Data to be clustered was organized in two different ways: the first, where all the data was treated as one set, and the second, where data from the stakeholder groups (CoMC, CoOHV, and SeniorCit) was clustered separately. This study revealed that the Ward's procedure can be used with various data groupings and the results will be similar. The Ward's method consistently resulted in an ANTHRO group and an ECO group. K-Means clustering, however, resulted in an ANTHRO group and an ECO group when the entire dataset of 511 observations was clustered, but when stakeholder group answers were clustered separately, this yielded two ECO oriented clusters. After reviewing models and methods, tests reveal that Ward's is a better method to use when clustering stakeholder values. This result has been confirmed by three other studies: Milligan and Cooper (1985), Atta-Boateng and Moser (1998), and Hix and Percy (1997). However, this study also shows that if K-Means were to be used as a clustering procedure, it should only be used when the number of observations exceeds 500.

Results from this study show many similarities between the ECO and ANTHRO groups. Both groups feel very strongly about the existence value of our natural resources, they feel that bequest value is important, that the government should not provide subsidies to companies for our public lands, that the public should be concerned with how public lands are used, and that consumers should be interested in the environmental consequences of products they purchase. In the past, it was perceived that environmentalists were a radical group that believed only in preserving forests, while non-environmentalists did not focus on

the issue of preserving forests. These findings disprove these perceptions, as it shows that the public, as a whole, is interested in these issues.

While there are many similarities between the ECO and ANTHRO groups, there are also differences. Differences found focused how the groups could do their share (sign petitions for environmental causes) to help the environment (more ECO) and how ECO groups believe that wildlife, plants, and forests have a right to exist for their own sake.

Overall, policy implications include recognizing that while stakeholders generally fall into one of two groups (ANTHRO or ECO), they have many similar values. These similarities should be the focus when making decisions about public lands where ECO, ANTHRO, or both have an interest.

REFERENCES

- Aerni, P. 2005. Stakeholder attitudes towards the risks and benefits of genetically modified crops in South Africa. *Environmental Science and Policy*. 8(5):464-476.
- Atta-Boateng, J. and J. W. Jr. Moser. 1998. A method for classifying commercial tree species of an uneven-aged mixed species tropical forest for growth and yield model construction. *Forest Ecology and Management*. 104(1/3):89-99.
- Beckley, T. M. and D. Korber. 1995. Sociology's potential to improve forest management and inform forest policy. *The Forestry Chronicle*. 71(6):712-719.
- Bengston, D. N. 1994. Changing forest value and ecosystem management. *Society and Natural Resources*. 7(6):515-533.
- Bengston, D. N., G. Xu, and D. P. Fan. 2001. Attitudes toward ecosystem management in the United States, 1992-1998. *Society and Natural Resources*. 14(6):471-487.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 20:37-46.
- Colorado Mountain Club (CoMC). 2003. www.cmc.org.
- Colorado Off-Highway Vehicle Coalition (CoOHV). 2003. www.cohvco.org.
- Colorado School of Mines. 1998. The Pike-San Isabel National Forest Survey: A survey on the use and management of public lands. Golden, CO: The Colorado School of Mines Environmental Policy Center.
- Fleiss, J.L. 1981. *Statistical methods for rates and proportions*. New York, NY: Wiley.
- Gil, M.J.A., J.B. Jimenez, J.J.C. Lorente. 2001. An analysis of environmental management, organizational context and performance of Spanish hotels. *Omega*. 29:457-471.

- Hix, D. M. and J. N. Pearcy. 1997. Forest ecosystems of the Marietta Unit, Wayne National Forest, southeastern Ohio: Multifactor classification and analysis. *Canadian Journal of Forest Research*. 27(7):1117-1131.
- Jacobson, S. K. and S. B. Marynowski. 1997. Public attitudes and knowledge about ecosystem management on Department of Defense land in Florida. *Conservation Biology*. 11(3):770-781.
- Kearney, A.R., G. Bradley, R. Kaplan, and S. Kaplan. 1999. Stakeholder perspectives on appropriate forest management in the Pacific Northwest. *Forest Science*. 45(1):62-73.
- Lee, T.T., J.K. Ziegler, R. Sommi, C. Sugar, R. Mahmond, L.A. Lenert. 2000. Comparison of preferences for health outcomes in schizophrenia among stakeholder groups. *Journal of Psychiatric Research*. 34(3):201-210.
- Manning, R., W. Valliere, and B. Minter. 1999. Value, ethics, and attitudes toward National Forest management: An empirical study. *Society and Natural Resources*. 12(5):421-436.
- McFarlane, B. L. and P. C. Boxall. 2000. Factors influencing forest value and attitudes of two stakeholder groups: The case of the Foothills Model Forest, Alberta, Canada. *Society and Natural Resources*. 13(7):649-661.
- Milligan, G. W. and M. C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*. 50:159-17.
- National Environmental Policy Act. 1969. 42 United States Congress, 4321-4347. Washington DC.
- Needham, M.D., R.B. Rollins, C.J.B. Wood. 2004. Stakeholders Perceptions of Bear Viewing Tours at an Alpine Ski Area in the Summer. *Human Dimensions of Wildlife*. 9(2):153-156.
- Nielsen, J.R., C. Mathiesen. 2006. Stakeholder Preferences for Danish Fisheries Management of Sand Eel and Norway Pout. *Fisheries Research*. 77(1):92-101.
- Podnar, K. Z. Jancic. 2006. Towards a Categorization of Stakeholder Groups: An Empirical Verification of a Three Level Model. *Journal of Marketing Communications*. 12(4):297-308.
- SAS Institute Inc. 1985. *SAS User's Guide: Statistics*. 5th edition. Cary, NC: SAS Institute Inc.
- SAS Institute Inc. 2000. SAS System for Windows. Release 8.01. Cary, NC: SAS Institute Inc.
- Schultz, P. W. and L. Zelezny. 1999. Values as predictors of environmental attitudes: Evidence for consistency across 14 countries. *Journal of Environmental Psychology*. 19(3):255-265.
- Steger, M. A. E. and S. L. Witt. 1990. Gender differences in environmental orientations: A comparison of publics and activists in Canada and the U.S. *Western Political Quarterly*. 42:627-650.
- Thompson, S., C. Gagnon, and M. A. Barton. 1994. Ecocentric and anthropocentric attitudes toward the environment. *Journal of Environmental Psychology*. 14:149-157.