PROCEEDINGS

1980 National Outdoor Recreation Trends Symposium

Volume I

U.S. Department of Agriculture, Forest Service
Northeastern Forest Experiment Station
370 Reed Road, Broomall, PA 19008

Sponsored by:

Northeast Agricultural Experiment Stations, NE-100

USDA Forest Service

USDI Heritage Conservation and Recreation Service

Recreation Working Group, Society of American Foresters

University of New Hampshire, Recreation and Parks Program

In cooperation with:

Clemson University

Journal of Leisure Research

Purdue University

USDA, Forest Service General Technical Report NE—57

FOREWORD

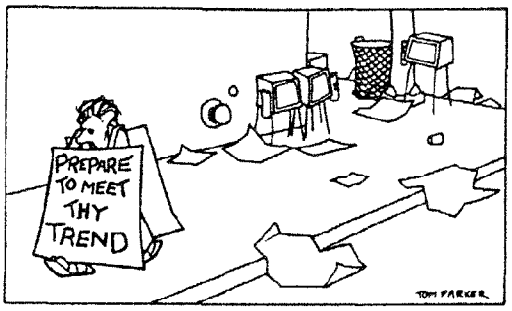
Most of us would probably endorse a one-year moratorium on meetings, conferences, conventions, workshops, and symposia. In fact, this planning committee was so reluctant to assemble another conference that it spent nearly 2 years identifying the needs and developing the program. When the rate of change is as great as it has been in outdoor recreation, conferences such as this one become essential. This is an exceptional conference because it focuses on that change, documents it, and attempts to determine what its future implications may be.

Ten years ago, a Forest Recreation Symposium was held at Syracuse, N.Y., for the purpose of "consolidating and synthesizing past research efforts in outdoor recreation." Even a hasty comparison of these proceedings with those from Syracuse suggests the enormous volume of research that has occurred over these 10 years. Equally apparent is the change in the kinds of research information that are available today; from the static descriptive and prescriptive studies of the late 1960's to examinations of trends, shifts, and changes in the outdoor recreation economy. Effective planning requires this dynamic view of outdoor recreation. Because planning, whether for corporate investment or public development, is a long-range activity, it needs information that goes beyond simple statements of "what is" into the realm of "what has been" and "what will be."

Statistical reporting is a critical function of government. Without this essential service, it would be difficult, if not impossible, to assess the state of the economy, the quality of health care, or the adequacy of public education. Price indexes, business slumps, new construction, pollution levels, production facts, and employment figures pop out of Washington bureaus onto boardroom conference tables with almost biologic regularity. Agriculture, mining, housing, manufacturing, wholesale and retail trade, doctors, dentists, educators, butchers, bakers, and even high school guidance counselors have more federally-sponsored statistics to plan with than do the providers of America's outdoor recreation opportunities. We attempt to plan the future of the Nation's recreation resources in the absence of facts about the present level and rate of growth of private investment in leisure industries. We define policy on the basis of out-of-date data and ideas about public participation in recreation activities. And, we invest scarce research dollars in "problems" which may not exist, or might at least look different if we had adequate statistics with which to view them. This symposium will not correct the situation. It can only serve to heighten your present uneasiness over the quantity and quality of available trend data. But we hope it will instill an urgency within you to demand better, more current, and more comprehensive statistics on outdoor recreation in America.

Good planning has been described as a two-step process. "First you figure out what is inevitable. Then you find a way to take advantage of it." In assembling this collection of speakers and topics, we have provided you with the best available information on, if not the inevitable, at least that which is highly probable and highly improbable. Step 2 -- how you take advantage of that information --is what recreation researchers will be monitoring in the years ahead.

WILBUR F. LaPAGE, Chairman
Program Committee



American Demographics, September 1979.
Used by permission.

THE 1980 NATIONAL

OUTDOOR RECREATION TRENDS SYMPOSIUM

Held at the New England Center for Continuing Education
University of New Hampshire
Durham, New Hampshire
April 20-23, 1980

SPONSORED BY

Northeast Agricultural Experiment Stations, Project NE-100

USDA Forest Service, Northeastern Forest Experiment Station

USDI, Heritage Conservation and Recreation Service

Society of American Foresters, Recreation Working Group

University of New Hampshire, Recreation and Parks Program

PROGRAM COMMITTEE --

Wilbur F. LaPage
Malcolm I. Bevins
Robert D. Greenleaf
Floyd L. Newby
Gerald L. Cole
A. Robert Koch
Herbert E. Echelberger
Douglas M. Knudson

CONCURRENT SESSIONS --

Floyd L. Newby
Herbert E. Echelberger
Douglas M. Knudson
Kenneth J. Hock
Marvin W. Kottke
Mahmood Seyala

ADVISORS --

R. Duane Lloyd
Earl Patric
Basil J. F. Mott
Roland Robinson
Meg Maguire
Robert McLellan
Barry Tindall
Hugo John
Fred Knight

LOCAL ARRANGEMENTS --

Robert D. Greenleaf
Herbert E. Echelberger
Gus C. Zaso
Lawrence A. Rondeau
Patricia C. Merrill

COOPERATORS

Clemson University, Department of Recreation and Park Administration

Journal of Leisure Research, NRPA

Purdue University, Department of Forestry and Natural Resources

CONTENTS

TRENDS OR METHODOLOGICAL DIFFERENCES?[1]

Daniel J. Stynes, Malcolm I. Bevins
and Tommy L. Brown[2]

Abstract.--Inconsistency in data collection has confounded attempts to identify and forecast outdoor recreation trends. Problems are highlighted through an evaluation of the methods employed in national outdoor recreation participation surveys and projections. Recommendations are advanced for improving data collection, trend measurement, and forecasting within outdoor recreation.

INTRODUCTION

Forecasting is an exciting and challenging endeavor. This is especially true in an environment as dynamic and complex as that which surrounds outdoor recreation in North America. For some, forecasting is merely fun and interesting while for others, it is an integral part of the decisionmaking and planning process.

Planners may be classified into three groups based upon their views of forecasting. Incrementalists do not try to forecast, generally reacting to events instead of trying to anticipate and guide them. For comprehensive-rationalists forecasting is essential. Their view of planning is based upon an ability to identify goals, formulate alternative courses of action, forecast, and evaluate the alternatives within a changing environment. Trend identifiers assume a compromising stance recognizing both the difficulties of forecasting as well as its importance. They typically monitor changes over time, extrapolating from these observations to produce short-range forecasts on which to base planning decisions.

We shall ignore the incrementalists based upon a belief that there is no such thing as "not forecasting" (Mendell 1969). A willingness to observe and explain outdoor recreation systems leads to an ability to predict their behavior. This may take the form of trend identification or forecasting by means of formal or informal models.

Recent attempts to identify outdoor recreation trends and to forecast the future provide some guidance for planning and decisionmaking. These efforts also raise many questions, suggesting that both historical trend data and forecasts of outdoor recreation be interpreted and applied with considerable caution.

Those who use trend information and forecasts must be aware of the underlying methods in order to evaluate the quality of the information and its applicability to a given decisionmaking situation. Our purpose here is to provide guidance to users of trend information and to make recommendations for improving the quality and applicability of outdoor recreation trend monitoring and forecasting efforts. While the focus is upon methods, the final objective is to contribute to the provision and utilization of better information to aid decisionmakers who must cope with and plan for a changing outdoor recreation system.

The title of our paper embodies several questions, some related to data collection and measurement and others related to the selection of a forecasting method or model. For trend identification, the basic question is whether the outdoor recreation participation data collected in the past is indicative of trends or merely reflects a wide range of differences in data collection methods.

---

[1]Paper presented at the National Outdoor Recreation Trends Symposium, Durham NH, April 20-23, 1980.

[2]Department of Park and Recreation Resources, Michigan State University; Department of Agricultural and Resource Economics, University of Vermont; and Department of Natural Resources, Cornell University; respectively.

Forecasting involves the application of a model to an existing data base. In evaluating forecasts one must be concerned with both the quality of the data base and the tenability of the assumptions underlying the forecasting model.

Our ability to predict where we will be in the future depends significantly on knowing where we were in the past and where we are today. Thus, we begin with a discussion of trend identification and trend measurement before delving into the more complex field of forecasting.

TREND IDENTIFICATION

Trend identification is based upon past measurements of outdoor recreation participation (and related events) and present monitoring of these same variables, as a guide to predicting future participation rates. Attempts to identify outdoor recreation trends from secondary data sources have met with numerous problems.

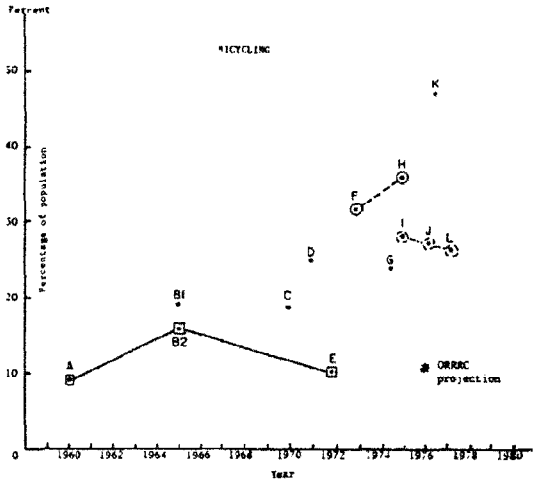
Trend identification is dependent upon reliable and valid data collected in a consistent manner over time. For outdoor recreation most basic data is either not collected, is inaccurate due to poor measurement techniques, or is collected inconsistently over time using non-comparable methods. Much historical data is not documented sufficiently to permit evaluations of its quality or applicability.

The state of existing outdoor recreation data bases suggests that trend identification and forecasting have not been the primary purposes of most data collection efforts. Most data collection is aimed at solving current problems. Potential use of such data for trend identification and forecasting is largely accidental. A good case in point is the series of nationwide outdoor recreation participation surveys.

Trend Identification from Nationwide Participation Surveys

Bevins and Wilcox (1979) have examined 22 national outdoor recreation participation surveys conducted between 1959 and 1978 in an attempt to identify trends in American's participation in a variety of outdoor recreation activities. The surveys examined include the Outdoor Recreation Resources Review Commission (ORRRC) studies, five nationwide surveys sponsored by BOR (now HCRS), three national camping market surveys (USFS), four surveys of hunting and fishing (USFWS) and seven market surveys conducted by private organizations.

Trend data for 28 different activities were examined, revealing some trends and a host of trend identification problems. The difficulties involved in discerning trends from nationwide surveys may be illustrated by an example. Figure 1 graphically depicts trend data for bicycling participation from 13 different surveys. Lines connect those surveys employing similar methods. The sponsor and methods for each survey are summarized in Table 1.



Symbol	Characteristics of survey: Contact	Recreation period	Age of respondent	Position in household
⊙·····⊙	Personal	Year	18+	Individual
⊙-----⊙	Telephone	Year	All	Individual
⊡——⊡	Personal	Summer	12+	Individual

A, B, C, . . . letters refer to surveys as identified in Table 1.

Figure 1. Bicycling—trend lines connecting surveys with similar methodologies, 1960-77.

Source: Bevins and Wilcox (1979)

No clear trend emerges for bicycling. Participation rates range from a low of 9 percent in the 1960 ORRRC survey to a high of 47 percent in the most recent BOR (HCRS) telephone survey. Some of the differences are partially explainable by differences in methods, survey populations, or the period for which data was requested. We expect participation rates from surveys of year-round activity to be higher than for surveys restricted to the summer months. Studies including younger age groups should yield higher participation rates in youth-oriented activities than surveys of adults. Other differences in measured participation rates over time may reflect legitimate trends. Sorting out which differences are due to methods and which are the result of trends is a difficult, if not impossible, task.

Table 1.--BICYCLING-participation rates according to nationwide recreation surveys, 1960-77.

Code	Year	Name	Characteristics of survey: Type of bicycling	Contact	Recreation period	Age of respondent	Position in household	Percent of population participating
(A)	1960	ORRRC	All	Personal	Summer[a]	12+	Individual	9
(B1)	1965	BOR	All	Personal	Year	12+	Individual	19
(B2)	1965	BOR	All	Personal	Summer[b]	12+	Individual	16
(C)	1970	BOR	All	Mail	Year	12+	Individual	19
(D)	1971	BOR	All	Personal	Year	10+	Individual	25
(E)	1972	BOR	All	Personal	Summer[a]	12+	Individual	10
(F)	1973	Nielsen	All	Telephone	Year	All	Individual	32
(G)	1974-75	TGI	All	Personal	Year	18+	Individual	24
(H)	1975	Nielsen	All	Telephone	Year	All	Individual	36
(I)	1975	TGI	All	Personal	Year	18+	Individual	28
(J)	1976	TGI	All	Personal	Year	18+	Individual	27
(K)	1976-77	BOR	All	Telephone	Year	12+	Individual	47
(L)	1977	AFI	All	Personal	Year	18+	Individual	26

[a]The summer period is June-August.

[b]The summer period is June-Labor Day

SOURCE: Bevins and Wilcox (1979)

Privately sponsored market surveys have shown greater consistency in methods than those sponsored by BOR. Nielsen measured an upward trend in bicycling between 1973 and 1975. Studies by Simmons Market Research Bureau (TGI) and the American Forest Institute. (AFI) indicate a slight downward trend since 1975. The BOR surveys illustrate a rather amazing variety of methods and results, providing little guidance for trend identification.

For selected other activities long term trends are more readily identified. Camping exhibits a gradual upward trend in all data series (Fig. 2) and hunting shows a fairly consistent gradual decline since 1960 (Fig. 3). Bevins and Wilcox (1979) report similar graphs for participation in 28 activities and sales of selected recreation equipment. The interested readers should consult that report for further details.

Influence of Survey Methods on Results

Examination of these nationwide studies may tell us more about the effects of alternative survey designs than about trends in participation. Unfortunately few users of the data produced in these surveys carefully examine the underlying methods in order to accurately interpret and evaluate the results. Participation rates are often reported without specifying the time period represented, the age groups included, the data collection method, the sample size, and other pertinent information. The resulting figures are thus prone to misuse.

Although we know that the methods employed will affect survey results, methodological research to date does not yield conclusive information on the magnitude of these effects. In comparing surveys conducted on different populations at different points in time, it is impossible to separate methodological effects from trend effects. A clearer picture of the potential influence of methods upon results is best obtained through controlled experimental designs. Little such research has been carried out with respect to recreation participation surveys, although some guidance is available in the general survey research literature.

Lacking controlled experiments, some insight may be gleaned by comparing surveys with similar methods conducted during the same year. This opportunity was provided in 1977 when both HCRS and NE-100[3] conducted surveys of outdoor recreation participation. Stynes (1979) addresses questions of survey comparability in the two studies.

[3]NE-100 is a Northeastern Regional Research Project examining "Recreation Marketing Adjustments in the Northeast."

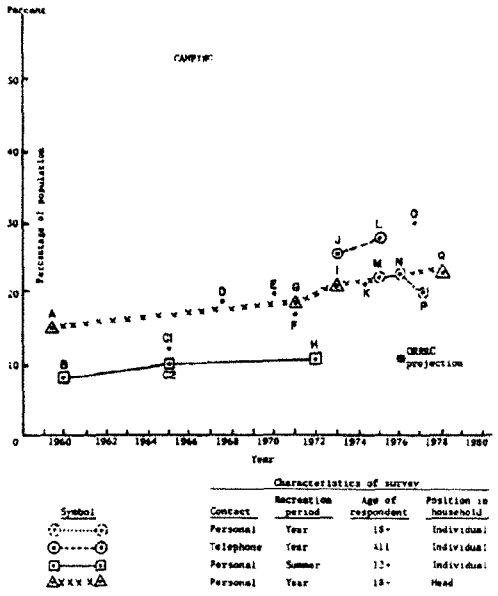


Figure 2. Camping—trend lines connecting surveys with similar methodologies, 1959-78.

SOURCE : Bevins and Wilcox (1979)

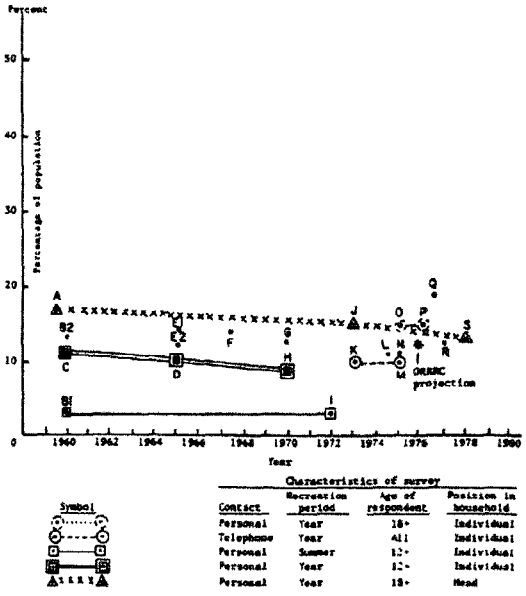


Figure 3. Hunting—trend lines connecting surveys with similar methodologies, 1959-78.

SOURCE: Bevins and Wilcox (1979)

The NE-100 survey included persons 18 years of age and older living in the Northeastern United States, while the HCRS survey was national in scope and included individuals 12 years and older. By selecting only those national survey respondents 18 or older and residing within the Northeast, a subsample comparable to the NE-100 population was obtained. Allowing for sampling errors in the two surveys we would expect (at a 95% confidence level) that the estimates of participation in outdoor recreation activities from the two surveys would be within from two to four percent of each other. Table 2 shows this to be the case for activities that were defined in a similar manner (those above the dotted line).

Of the 22 activities included in the NE-100 survey and 30 in the HCRS study, only 12 are directly comparable. Of these dozen activities, only picnicking reveals differences significantly greater than sampling error tolerances. The variation in the definition and grouping of activities below the dotted line (Table 2) illustrates a common problem in comparing the results of two surveys. When definitions or groupings of activities change even slightly from survey to survey or year to year, few clues about recreation trends for these activities can be discerned.

Differences in recall periods, response rates, question wording, question sequencing, timing of the survey, and a variety of method-related variables will also affect survey results. Examination of the variety of nationwide survey methodologies suggests that many differences in survey findings over time are due to methodological differences and are not necessarily indicative of outdoor recreation trends.

This means that those who use survey data must exercise caution in their interpretation and application. As a guide we offer a checklist of factors to consider in comparing or evaluating outdoor recreation survey results for the purpose of identifying trends (Table 3).

In using a given survey result for a planning or management decision both the quality of the survey methods and the generalizability of the results to the situation at hand must be evaluated. In trend identification it is important that each survey employ comparable methods on comparable populations. Differences in outdoor recreation participation measured in two surveys may result from different survey populations, different contact methods, or minor differences in question wording or recall periods. Differences may simply reflect sampling error, may be due to non-sampling errors, or may be indicative of a trend.

TABLE 2. -- COMPARISON OF NCRS AND NE-100 OUTDOOR RECREATION PARTICIPATION ESTIMATES

NE-100		NCRS	
Activity	Percent Participating	Percent Participating	Activity
Bicycling	42	42	Bicycling
Boating (sailing)	11	14	Sailing
Golfing	16	13	Golf
Hunting	11	13	Hunting
Picnicking	60	69	Picnicking
Skiing (downhill)	11	9	Downhill Skiing
Skiing (cross country)	4	3	Cross Country Skiing
Skiing (water)	10	12	Waterskiing
Snowmobiling	8	10	Snowmobiling
Ice Skating	22	24	Ice Skating outdoors
Tennis	30	30	Tennis outdoors
Canoeing	12	13	Canoeing, Kayaking, River Running
Fishing (freshwater)	24	60	Fishing
Fishing (saltwater)	19		
Boating (motor)	11	32	Boating (other)
Camping (RV)	11	20	Camping in a developed area
Camping (tent)	16	16	Camping in a primitive area
Camping (backpack)	8	27	Hiking or Backpacking
Hiking	25		
Jogging	27	68	Walking or jogging for pleasure
Motorcycling	11	21	Driving vehicles or motorcycles off-road
Swimming	59	63	Outdoor Pool Swimming or Sunbathing
		48	Other Outdoor Swimming or Sunbathing

SOURCE : Stynes (1979)

FORECASTING

Inconsistencies in data collection have made trend identification difficult. Since trend data is a basic input to forecasting models outdoor recreation forecasting has been similarly constrained. The lack of time series data on recreation participation and explanatory variables has limited the types of forecasting models that might be used. In evaluating a forecast one must evaluate the accuracy of the data inputs (the base period figures, for example) and the validity of the forecasting model assumptions.

Given a set of historical data, different models applied to this data may yield different results. Differences in forecasts will be due to the different assumptions of each model. An understanding of these assumptions is essential to evaluation of the forecast. Ideally one evaluates the tenability of the various assumptions for the situation under study in order to evaluate the applicability of alternative models. Practically, questions about the quality or availability of data inputs for outdoor recreation have often preempted evaluation of forecasting model assumptions themselves.

Recreation Forecasting Methods

There are a variety of both qualitative and quantitative forecasting models to choose from. Martino (1972) provides an excellent summary of long-range technological and social forecasting techniques and Wheelwright and Makridakis (1973) review shorter range techniques widely applied in business management.

Moeller and Echelberger (1974) review the forecasting techniques most often applied in outdoor recreation. These tend to be middle-range forecasts. For predicting future levels of outdoor recreation participation, two basic types of forecasting models dominate: (1) trend extension models and (2) structural models relating participation in outdoor recreation to explanatory variables that are more readily forecasted. Each method has advantages and disadvantages.

Trend Extension

Trend extension follows naturally from trend identification and monitoring, and is highly dependent upon the quality and consistency of trend data. As its name implies, the method simply involves the extension of trends as revealed in historical measurements of use or participation. The method works well as long as the underlying forces producing a given trend do not change significantly. This assumption is generally valid for making projections one to three years into the future

TABLE 3: CHECKLIST FOR EVALUATING OR COMPARING OUTDOOR RECREATION SURVEYS

1. What is the study population?
   - the study region
   - the minimum age for inclusion
2. What is the method of contact?
   - telephone
   - personal interview
   - mailed
   - on-site or household design
3. For household designs, how was the respondent selected within the household?
   - person who answers
   - household head
   - randomly selected
   - based upon quotas
4. How successful was the sampling design?
   - representativeness of the sample
   - response rate/possibility of non-response bias
     weighting procedures
5. Sampling Errors
   - are confidence intervals reported for population estimates?
   - on how many respondents is each figure based?
6. When was the survey conducted?
   - time of year
   - interviewing on weekends-weekdays; afternoons-evenings?
7. Measurement
   - definition of recreation activity categories
   - time period for which data is requested
   - length of recall period
   - wording of questions
   - open or closed-ended questions
   - question sequencing
8. Reporting of Results
   - Are methods clear enough so that you could replicate the study?
   - Are non-sampling errors discussed?
   - Are study limitations discussed?

and when the system under study is reasonably stable.

Even in trend extension, one must consciously or unconsciously make assumptions. Figure 4 illustrates several distinct projections based upon observations from three time periods. With data for only three previous years, assumptions of a linear, exponential or logistic (S-shaped) growth pattern all seem reasonable. These different assumptions yield quite different projections. Selection of the form of the equation must be based upon logical as well as statistical considerations. Do we expect the trend to exhibit a linear, exponential, or logistic growth pattern during the years for which we are projecting? Too often linear models are selected based upon statistical convenience or ignorance of alternative growth curves.

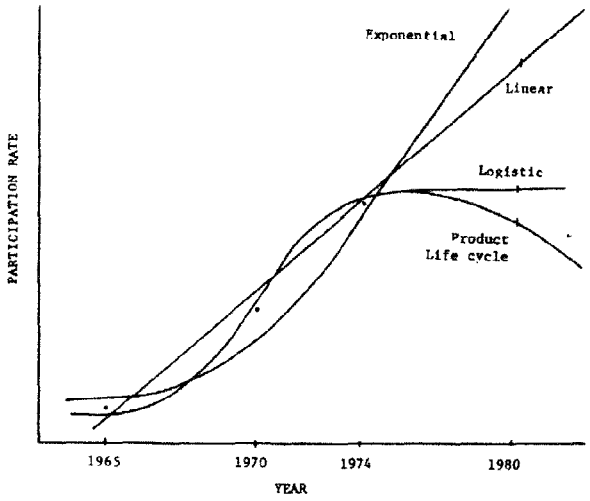


Figure 4. Trend Extension - Alternative Growth Equations

In fact, very few processes grow linearly. Logistic and exponential growth are much more characteristic of population and diffusion processes. Product life cycle curves often exhibit logistic patterns of growth, usually with a decline towards the end of the life cycle. We expect the adoption of outdoor recreation activities to exhibit similar patterns of growth.

The advantage of trend extension methods is their simplicity. However, these methods do not address the underlying forces that are producing the trend. Clearly an understanding of these forces should result in much better and much more useful forecasts, especially within a changing environment.

Structural Forecasting Models

More complex models are generally required to forecast further than five years into the future. In this case the assumption that the pattern of causal forces behind the trend remains unchanged no longer holds and an understanding of these forces becomes an important component of the forecasting model.

The most widely applied outdoor recreation forecasting model (other than seat of the pants) is the two step linear regression technique. The method was applied in the ORRRC studies (Report #26) and has subsequently been refined and further expounded by Cicchetti (1973) among others.

For a person with a given set of characteristics and recreation opportunities his or her probability of participating in a given recreation activity is estimated in the first step. Then in the second step the frequency of participation is estimated for those who participate. Both steps employ linear regression techniques to estimate coefficients in an equation of the form:

$$y = f(\text{socioeconomic characteristics}, \text{recreation opportunities}, \ldots)$$

where y is the probability of participating in the first equation and the frequency of participation in the second. Recreation survey data is used to statistically estimate the model parameters and the resulting equations are applied to forecasts of each independent variable in order to forecast future participation.

The independent variables are selected based upon their influence on participation and the ease with which they may be measured and forecasted. Age, income, and gender have proven to be the best predictors of outdoor recreation participation in such models. These variables are convenient for forecasting since predictions of future age-sex structures, and income levels are generally available. More recent applications have incorporated supply and price variables, although measurement and forecasting of these variables are more problematic.

The merits of these models, their problems, and the underlying assumptions will be discussed using the ORRRC projections as an example.

Although accuracy should not be our prime criterion for judging forecasts[4], the evaluation of how close these forecasts come to predicting 1976 participation rates will shed some light on the two-step linear forecasting model.

Evaluation of the 1976 ORRRC Projections

Brown and Hustin (1979) evaluated the ORRRC projections by comparing them with participation rates measured in the 1976 HCRS national telephone survey. A comparison of projected versus measured participation rates finds that the ORRRC study grossly underestimated participation in virtually all activities. For many activities the measured 1976 participation rate is more than double the rate projected by ORRRC (Table 4).

Table 4: Comparison of Proportion of Persons 12 Years and Older Projected to Participate in Outdoor Activities in 1976 vs. Number Participating in 1976 Survey

Activity	Projected by ORRRC[1]	Participating in 1976 Survey	z score[2]
Driving for pleasure	56%	69%	26.68
Swimming	55	70	31.38
Walking for pleasure	37	68	67.51
Sightseeing	47	62	32.76
Picnicking	57	73	33.45
Fishing	32	55	51.33
Bicycling	11	47	122.04
Attending outdoor sports events	27	61	81.26
Boating (not canoeing or sailing)	28	35	15.73
Nature walks	16	49	95.24
Hunting	14	20	18.23
Camping	11	37	88.67
Horseback riding	8	15	27.99
Water skiing	9	17	29.11
Hiking	8	26	78.15
Attending outdoor concerts, plays, etc.	12	40	90.89

[1] From ORRRC Study Report 26, pg. 27.

[2] Normal approximation of a binomial proportion, corrected for continuity (Snedecor and Cochran, 1973). The null hypothesis of equal proportions projected vs. surveyed can be rejected for $p < .99$ at $z = 2.33$.

SOURCE: Brown and Hustin (1978)

The ORRRC projection model was based upon six independent variables: (1) income, (2) education, (3) occupation, (4) age-sex, (5) urban-rural residence, and (6) leisure time. Relationships between participation in each activity and the six variables were estimated based upon a 1960-61 survey of outdoor recreation participation. A linear model was assumed.

[4]Forecasts should be judged on their usefulness in decisionmaking, not their accuracy (Martino 1972).

Projections to 1976 were then made for each of the six independent variables and the model was applied to these projections to derive the 1976 estimates of participation for each activity.

Three types of errors are involved in this forecasting process.

1. Measurement, sampling and other errors in the 1960 and 1976 surveys.
2. Errors in the forecasts of the independent variables.
3. Model specification errors.

The accuracy of a forecast always depends upon the accuracy of base period fugures (1960 in this case). It should be noted that the model parameters were determined solely from a single survey conducted in 1960-61. This survey is subject to measurement and sampling errors. In evaluating the accuracy of the forecast, we also assume that the participation rates measured in the 1976 HCRS survey are accurate.

A more subtle difference between the 1960 and 1976 surveys is easily overlooked. The ORRRC projections are based upon a survey of activity between June and August 1960. Projections are therefore participation rates for the summer months of 1976. The 1976 HCRS survey measured participation over the entire year. It is not possible to separate out the summer months and hence any comparison of the ORRRC projections and the 1976 survey must be based upon assumptions about differences between summer and year-round participation rates.

We should expect summer rates of participation (i.e. the proportion of the population participating at least once during that period) to be less than year-round participation rates. The magnitude of these differences will be greater for activities with substantial non-summer use (eg. fishing and hiking) and relatively insignificant for predominantly summer activities like swimming and boating. In the following we assume no difference between summer and year-round participation rates.

The second source of error is in the projections of each independent variable. Comparing the ORRRC projections of these variables with actual estimates for 1976 reveals that income, education, occupation, and age-sex distributions were projected fairly accurately. No comparable 1976 estimates are available for residence or leisure time, making the projections somewhat suspect.

Model specification errors are the most serious in this case. A principal assumption of the two step method is that the relationship between the explanatory variables and

participation, as estimated from 1960 data, continues to hold in 1976. Specifically, this assumes constant participation rates by demographic segments over time. Any change in participation rates predicted by the model must be due solely to changes in the explanatory variables. The increasing participation of women in many outdoor recreation activities (Bevins 1979) is one example of a direct contradiction to the model assumptions.

Model formulations that include supply and price variables are an improvement, but introduce additional measurement problems. The linearity of the relationship might be questioned, but logit and probit forms of the participation equation do not yield much improvement (Smith and Munley 1978). The use of these types of models for forecasting involves six common problems.

(1) A failure to include substitution effects. Separate and independent equations are generally developed for each outdoor recreation activity. A few of the models that include a price variable also include prices of close substitutes or complements, but these are rare (Talhelm 1973). Recent research into substitution among outdoor recreation activities does not yet provide much help for forecasters.

(2) Difficulty of incorporating supply factors. One's likelihood of participating in a given recreation activity is clearly related to the quality, quantity and price of available opportunities. Cicchetti (1969) and Beaman (1976) discuss the inclusion of supply variables and the accompanying problems.

(3) Reliance on cross-sectional data. Cross-sectional surveys are clearly not the best way to measure change or the forces producing change. Projecting 16 years into the future based upon observations during a single year is clearly suspect. Brown and Wilkins (1975) demonstrate that more accurate forecasts can be developed using structural models estimated from time series data, when this data is consistent and accurate and corresponding time series data is available for the explanatory variables. Unfortunately such data rarely exists.

(4) Other missing explanatory variables. The identification of explanatory variables continues to be a subject for research. For long-term forecasting a variety of social and economic variables seem relevant. Changing value systems, energy policies, family structures, and leisure time patterns are seldom included in recreation forecasting models. West (1977) suggests the addition of variables related to fad and fashion in leisure activity.

(5) Aggregation problems. Selecting the appropriate aggregation level continues to trouble recreation planners and forecasters. Regression models are generally estimated using data about individuals and then applied to aggregate populations. The ORRRC and most other similar projections apply the resulting equations to forecasts of the means of the explanatory variables.

The participation rate of a person with average income, average age, and an average amount of leisure time will not necessarily be a good estimate of the participation rate of the population as a whole. A more valid technique is to develop rates of participation for various population subgroups and apply these to the forecasted numbers of people within each subgroup. This requires future distributions of the population over all explanatory variables rather than just the projected population means for each.

(6) Statistical rather than process models. The structural models developed to date are for the most part statistical rather than process models. They do not capture the underlying decision processes that determine outdoor recreation participation. An individual clearly does not take a weighted sum of his income, age, occupation, and leisure time in deciding whether or not to participate in an activity. These models therefore yield little incite into the dynamics of outdoor recreation participation decisions.

Systems and simulation techniques are better suited to the identification of feedback effects, timelags and other dynamic characteristics of recreation systems. An examination of activity adoption, participation, and dropout decisions in the light of changing family structures and recreation opportunities might yield a better understanding of the dynamics of recreation participation. Models relating participation to equipment sales also seem promising.

CONCLUSIONS

The stimulus for this paper was the question of whether outdoor recreation participation data collected in the past is indicative of trends or merely reflects differences in data collection methods. We must conclude that some data series indicate changing patterns of participation and others are the result of measurement errors or methodological differences. The problem is that the task of telling which is which is not an easy one, and in many cases is impossible.

Those attempting to divine trend information from published research must exercise considerable caution. An understanding of data collection methods and an ability and willingness to dig into documentation of survey results in order to evaluate the quality and applicability of the findings is advised. Better data collection techniques and documentation of results are also needed.

Multiple sources of trend data should be sought and compared before drawing conclusions. Consistency and comparability of findings from different sources increases our confidence in the conclusions. This is especially true when distinct methods independently yield similar results.

Many of our recommendations for improving data collection, trend measurement and forecasting have been made before. The need for more consistent and comparable data collection efforts over time is clear. Before this can happen the importance and utility of trend information and forecasts must be recognized. There have been few systematic efforts to collect outdoor recreation data for the purpose of identifying trends. Available data has been collected primarily for other purposes. The data requirements for trend identification and forecasting are somewhat unique. Unless data is collected for these purposes the chances are that it will be inadequate or useless for forecasting.

Given the costs of data collection and the vast amount of trend data that might be collected, a systematic examination of trend information needs is required. This must be based upon an identification of the kinds of planning decisions that trend information might contribute to, and an assessment of the likely improvement in planning decisions as compared with the costs of data collection. Institutional frameworks must be established to ensure periodic and systematic collection of the data and to develop appropriate information systems to facilitate its wise use.

Research is also needed: (1) to develop clear standard definitions of recreation terms, (2) to improve measurement techniques and develop acceptable measures for variables that are difficult to quantify, (3) to identify variables that may help explain changes within outdoor recreation, and (4) to experiment with forecasting models that more fully capture the processes that are taking place within outdoor recreation.

LITERATURE CITED

Beaman, J. 1976. Statistical Projections That Go Beyond Projections of Past Trends. Canadian Outdoor Recreation Demand Study. Technical Note 13. Ontario Research Council on Leisure. Toronto.

Bevins, Malcolm I., Tommy L. Brown, Gerald L. Cole, Kenneth J. Hock, Marvin W. Kottke, Wilbur F. LaPage, Richard W. Stammer, and Daniel J. Stynes. 1979. Changing Patterns of Outdoor Recreation Participation in Northeastern United States. Northeast Regional Research Publication. Delaware Agricultural Experiment Station, Newark.

Bevins, Malcolm I. and Daniel P. Wilcox. 1979. Evaluation of Nationwide Outdoor Recreation Participation Surveys, 1959-1978. Vermont Agricultural Experiment Station. Burlington.

Brown, Tommy L. and K.L. Hustin. 1978. Evaluation of the 1976 ORRRC Projections. Unpublished Report. Department of Natural Resources, Cornell University. Ithaca.

Brown, Tommy L. and Bruce T. Wilkins. 1975. Methods of Improving Recreation Projections. Journal of Leisure Research 7(3):225-234.

Cicchetti, C.J., J.H. Seneca, and P. Davidson. 1969. The Demand and Supply of Outdoor Recreation. Bureau of Economic Research. New Brunswick, NJ.

Cicchetti, C.J. 1973. Forecasting Recreation in the United States. D.C. Heath. Lexington, Mass.

Martino, Joseph P. 1972. Technological Forecasting for Decisionmaking. American Elsevier. New York.

Mendell, Jay. 1969. There is no such thing as NOT forecasting. The Futurist (Dec.): 161.

Moeller, George H. and Herbert E. Echelberger. 1974. Approaches to Forecasting Recreation Consumption. Outdoor Recreation Research; Applying the Results. U.S. Department of Agriculture Forest Service General Technical Report NC-9. St. Paul, Minn.

Outdoor Recreation Resources Review Commission. 1962. Prospective Demand for Outdoor Recreation. Study Report 26. U.S. Government Printing Office. Washington, D.C.

_____. 1962. National Recreation Survey. Study Report 19. U.S. Government Printing Office. Washington, D.C.

Smith, V. Kerry and Vincent G. Munley. 1978. The Relative Performance of Various Estimators of Recreation Participation Equations. Journal of Leisure Research 10(3):165-176.

Stynes, Daniel J. 1979. A Comparison of the 1977 National Telephone Survey with a Northeastern Regional Survey of Outdoor Recreation Participation. Unpublished Report. Department of Park and Recreation Resources. Michigan State University. East Lansing.

Talhelm, Daniel. 1973. Economic Evaluations of Sport Fishing in Michigan: A Comparison of Methods and their Application. Paper presented at American Agricultural Economics Association meeting. Edmonton, Alberta.

U.S. Department of the Interior, Heritage Conservation and Recreation Service. 1978. 1977 National Telephone Survey. Unpublished Report. Washington, D.C.

West, Patrick C. 1977. A Status Group Dynamics Approach to Predicting Participation Rates in Regional Recreation Demand Studies. Land Economics 53(2):196-211.

Wheelwright, Steven C. and Spyros Makridakis. 1973. Forecasting Methods for Management. John Wiley and Sons. New York.

THE TREND OF MEASURING PUBLIC USE OF THE NATIONAL PARKS[1]

Kenneth E. Hornback, Ph.D[2]

INTRODUCTION

Since the outburst of interest in outdoor recreation travel in the period following World War II, measurements of outdoor recreation have been made in many ways. The measurement of outdoor recreation is decentralized activity being carried out by a large number of private and federal agencies with a variety of goals and purposes. Future improvements in federal statistics will partially come from the emergence of a combined strategy for statistical data gathering. Partially due to decentralization, such a strategy has not emerged even though the need for trend data is wide-spread. To envision the characteristics of a more coordinated and cost effective program of outdoor recreation measurement, it is helpful to consider the history of outdoor recreation studies. Changes in studies conducted for the National Park Service (NPS) illustrate a trend with several periods.

Travel and Tourism Studies

From the early 50's to the early 60's, the Federal Highway Administration (FHA) engaged in cooperative programs with state highway departments to collect specific measurements of travel and tourism activity. The primary concerns of these studies were to gather basic data for economic and highway planning. Major national parks were sites for many of these studies.

Because of the relatively narrow applications intended for these studies, the scope of inquiry was limited to a few basic issues. The information collected usually included vehicle type, home town, trip purpose, trip mileage, overnight stays, duration of stay, and trip expenditures. As a result of the narrowness of focus, the comparability between travel studies is surprisingly high.

These travel and tourist studies used entrance station interviews or mailback questionnaires. The brief sets of questions could be completed by the visitor in a few minutes. Appendix A contains information about how field operations were carried out (Cape Cod study of 1963). Appendix B shows several questionnaires used in the Grand Canyon Tourism Study (1953). Several of the studies used identical interview schedules (Grand Canyon and Yosemite studies in 1953).

The findings of the travel and tourism studies were generally reported as descriptive statistics in various tables and figures. The focus was on the market area served by a particular park, the forecast of future use which could be expected from that market area, and the economic benefits of visitation to surrounding areas. The reports vary in quality and tend to assume that the usefulness of the results is self-evident. Descriptive findings are typically presented with expressions like "it is of interest that". . . "interesting to note", and "as reasonably could be expected. . ." (Shenandoah Travel Study, 1952). There is no record of any raw data ever being centrally stored for later use although in one case findings were compared to eight other travel studies (the 1963 Cape Cod survey).

---

[1]Prepared for publication in the proceedings of the Outdoor Recreation Trends Symposium, April, 1980, Durham, N.H. Forest Service, Department of Agriculture.

The views expressed here are those of the author alone. These views are not official policy of the National Park Service and do not involve material approved by any office of the Department of Interior. I would like to acknowledge the helpful reviews provided by Dr. Don Field, Pacific Northwest Region Chief Scientist, NPS, and Dr. William Key, Consultant to the Chief Scientist, NPS, Washington.

2 Kenneth E. Hornback, Ph.D Chief, Statistical Office, U.S. Dept. of Interior, National Park Service, Denver.

The travel and tourism studies reflect several advantages and disadvantages of measuring outdoor recreation. Advantages included the narrow focus, general comparability between studies, cost effectiveness of acquiring data, and the use of client participation in the collection effort (FHA funds enabled State highway departments to collect data or hire local people to do the work). Disadvantages included the failure to provide for the common storage of raw data and the failure to document applications of data to planning and management problems. These studies were one-shot studies and were not concerned with the problem of learning about trends.

Travel and tourism studies received lower priority beginning in the early 60's. As planning for the inter-state highway system neared completion, the work of the FHA concentrated more on the problems of urban area transportation planning. But as this particular type of study declined in frequency, a new type of study began to appear.

Visitor Use Studies

The study of travel and tourism had many beneficial effects for the NPS. Among the benefits was the interest stimulated in the visitor as a factor to be treated in both planning and management. The benefit, however, did not come directly from the original travel and tourism studies but from the great variety of "visitor use studies" which were sponsored by diverse interest from many different parts of the agency. These studies tended to be exploratory and responsive to the new curiosity of park staff about the visitor, visitor attitudes and behavior. The visitor use study appeared in several distinct forms.

The user satisfaction study was conducted to gather information needed to evaluate interpretive programs and to gather basic performance data on how well the park was doing its job. These studies often concentrated on the things which visitors liked, found useful, or judged to be of value to them. To the extent that they focused on manageable conditions, these studies discovered that people were more satisfied when the litter was controlled and the trails and facilities were maintained. Often, however, the focus of these studies was vague and they often identified circumstances which were beyond the realm of manageable conditions, e.g. campers were satisfied when other campers were pleasant to be around. Such studies often attempted to measure the perception of environmental quality or determine the meaning of the visiting experience. To this extent, indicators were developed to capture differences in the expressions of feelings, experiences, and perceptions. These types of studies did not lend themselves to comparison because they used a great variety of definitions, concepts, and methods of measurement.

Closely related to the "satisfaction" type of study was the attempt to measure sociological "carrying capacity." Emerging in the early seventies, these studies attempted to parallel natural science studies of resource wear and tear. The sociological measurement problem was to determine when visitor density itself prompts changes in the quality of the visiting experience. These studies were plagued by a variety of confounding conceptual, theoretical, and operational problems. Some visitors, for example, valued having others around while others wanted to avoid people. The "sociological carrying capacity" type of study, however, did serve to sensitize managers to the varying needs of user groups.

The "visitor profile" was yet another form of visitor use study. Characterized by their taxonomic style, these studies grouped people according to various types or styles of behavior, e.g. backcountry users, straight and counter-culture users, family users, etc. These studies served to refine statistical generalizations such as emerged from the tourism studies by demonstrating that while the "average visitor" did not exist, there are patterns of behavior that distinguish people from one another in ways that bear on management actions. The focus on differences between visitors is important because it suggests that there can be corresponding differences in management actions. Such studies, for example, contributed to thinking about ways to control potential conflict among visitors by zoning activities. Knowledge of seasonal changes in types of visitors enabled changes to be made in the content of interpretive programs.

One of the most distinctive types of visitor study was the regional, interagency transportation study. These studies attempted to bring the accumulation of many earlier styles of inquiry to bear on outdoor recreation behavior at one time. While few of these were completed (Great Smoky Mountain National Park study in 1975 and Yellowstone-Teton National Park in 1978), they played a key role in the spread of interest in the visitor and impacts of visitation on parks and the surrounding communities.

However, because of the broad focus, the large variety of participants, and wide areas of geographic interest, the volume of survey data collected by these studies grew to enormous proportions. In contrast to the brief tourism studies, the regional visitor use studies pursued answers to scores of questions ranging from general recreational activities to social attitudes. A major result of this type of study was the sensitivity it created about cost effectiveness (time and monetary cost to the agency as well as reporting burden on the participating public). It seemed evident that if information about the public as a consumer group was to be applied routinely to national park management, it would have to be timely and through a more limited, cost effective method.

Throughout the sixties and into the mid-seventies, various forms of visitor use studies have been conducted on behalf of the NPS. Throughout this period, unfortunately, an integrated body of organized knowledge failed to emerge even though great effort was invested in the task. If anything, the great diversity of effort suggested the need for establishing some centralizing control to insure economy of effort.

Research Studies

The proliferation of ad hoc studies during the 60's served to bring the need for a more systematic program of studying the visitor and visitor behavior to the attention of NPS management. In the early 70's, efforts were completed to establish several regional centers of social-scientific research in the NPS.

The Cooperative Park Studies Units were created at various universities for the purpose of blending the intellectual power of academia with the apparently intractable difficulties of resource management.

Centralizing the expertise for social science work served a variety of purposes. At last there were places where consultation could be acquired, where research coordination could be made, where quality control over study design could be exercised, and where findings and raw data could be assembled for later use. The institutional formalization of social science activity prompted greater research coordination and indirectly contributed to the assembly of interdisciplinary perspectives. These centers served as a focus of communication between the academic and governmental communities and achieved some degree of reorientation of effort for their mutual benefit. In addition, the monumental task of inter-agency coordination was carried out by these centers, eliminating some redundancy and improving the transfer of social science technology to NPS management at minimal cost, e.g. the computerized backcountry permit system developed by the Social Science Program in the Pacific Northwest Region, NPS.

Throughout the period of growing experience and mobilization of effort, however, a persistent problem served to confuse the establishment of more viable social science goals: who was to be served by the application of social science to management and planning, and how? Was the level of application to be at the site, within the region, or at the national level? Nowhere is the conflict of applications more evident than in various public opinion polls which were conducted.

Public Opinion Polls

This type of work differs from other visitor studies in that the population studied is the general rather than the visitor population. The methods, therefore, tended to be off site telephone and personal interviews. Such studies were needed because studies done for research purposes or on behalf of unit managers did not yield information needed about broader problems.

From 1968 to 1972 a variety of national and regional public opinion polls were designed and carried out for the purpose of expanding knowledge of participation in and opinions about outdoor recreation, a purpose which was not being served by other studies. From these studies valuable perspectives about the relationship between visitors and nonvisitors was gathered and made available to upper level management. In addition, a series of short "People in the Parks" reports were prepared to explain the general usefulness of the data to planners and managers. Theoretically, information would "filter down" until it found an application. The reports were propagated because they "may also be useful to other divisions for any number of purposes. . . How it may be useful to each division will, of necessity, be decided within the division."

Unfortunately the expected integration of general survey findings into planning and management documents did not spontaneously happen, possibly because there were no experienced social scientists available to help other professionals find the meaning of general findings for specific actions. On the other hand, social scientists themselves may not have had sufficient experience in the work of planning and resource management. Although much effort was put into the task of delivering findings to other professional groups, social scientists often stopped short of saying exactly how information was to be used.

Discussion

The preceding account of studies done for the National Park Service represents one viewpoint on how social science work was carried out and what was accomplished over the years. While many subjects have been explored and a variety of study tactics have been tried, the cumulative impact of the findings on planning and management is less than might have been expected judging by the effort (time and cost) invested. The failure to communicate better could be due to a variety of problems such as the delays necessary to accomplish the studies themselves, lack of general experience of scientists and planners in working together on managerial problems, the artificial "freezing" of data in the text of a written report, failure to focus on comparative and trend data, and mismatched expectations of what studies can and cannot do. While studies can make crucial impacts on thinking, study findings too often were applied cosmetically if at all.

Up to the present both the scientific and managerial communities have been learning about the role of social and economic data in natural resources management. While the period of learning is far from over, a period of applications needs to be started. The tactic of gathering data by studies may itself be a problem. While scientific studies will continue to be essential sources of new knowledge, studies may not represent the tools needed to effectively bring information to planners and managers. Other ways of gathering and communicating the use of information need to be explored.

The coming decade promises to bring revolutionary changes in the economy which will alter previous circumstances in which out-door recreation has occurred. The context of planning will soon be unlike what it has been in the past. The rate and degrees of change may be so fast and extreme that conventional studies will be inadequate to assemble required data for planning and management. In the future, natural resources planning and management will rely more heavily on limited programs to monitor public use and an improved system of federal statistical reporting about out-door recreation.

Monitoring Public Use

From the limitation of studies as a method of collecting information for park management, a complementary method was developed in the late seventies: continual monitoring of public use. While not an evolutionary step beyond the ad hoc study, continual monitoring serves purposes which cannot be as effectively served by conventional studies. While monitoring must be an activity which is restricted to collecting a few variables, the benefit of monitoring is that it yields comparable data on a continuing basis. The focus of monitoring is on change and, more importantly, comparative change. This is an important development in the effort to bring social science to bear on park planning and management because with this type of information it is possible to focus on the trends of public use and the question of what differences exist between types of areas, geographic locations, and forms of management.

The content of the current public use monitoring program under development by the NPS is limited to a few items which can be quickly gathered using a free-form interview schedule. Questions include (1) where people entered the park, (2) when they entered, (3) how often they come, (4) where they live, and (5) what other places they have visited on their trip. In addition, the visitor is asked "on your first day in the park, (6) where did you go, and (7) what did you do?" This last question is a general probe designed to recover the details of the visit using a notational coding method. For the purpose of reducing costs, this type of interview is designed to be conducted by members of the park staff who are specially trained by the NPS Statistical Office. Seasonal or permanent staff members gather the data in the context of their day-to-day contacts with visitors. The interviews serve to give visitors an opportunity to ask questions and give comments (both positive and negative) about their visit and to give park staff a systematic way to learn about the visitor and the pattern of park use. All data are carefully edited and errors are returned to the interviewer for training purposes.

The major source of cost effectiveness, however, is the entry of data into a data base management system (DBMS) which is accessed by conventional English commands (INQUIRE). The timeliness of data collection and editing creates a "live" foundation of statistical data for planning and management. Rather than prepare reports on findings, park service statisticians prepare a library of graphic and tabular outputs which is made available to clients of the system as needed (Computer Assisted Management Program - CAMP). Parks draw data interactively using conventional computer terminals. Use of the system and its application to various kinds of work are explored in "Applications Clinics" conducted at the parks by members of the agency statistical office.

One of the major limitations of the monitoring type of activity is the question about the quality of the data. If such programs are to be successful in terms of financial cost, they must take advantage of every opportunity to economize. Asking the unit manager to administrate the task with existing staff is a shortcut which eliminates one of the major expenses of this work. Although NPS field staff are not primarily assigned to conduct interviews, most have shown a genuine interest in gathering this kind of information. One of the major reasons for strong field support is the dedication of the monitoring effort to gather basic information commonly recognized as essential to routine park management. The availability of data within ten days of collection also serves to stimulate interest in the quality and utility of the findings. To the extent that active interest in the findings can be maintained by timely production of findings and to the extent that training and careful editing of data can be completed, the monitoring program can result in data of respectable quality for the purpose to which it is applied.

As the concept of monitoring as a tool of park administration spreads, more areas will be added to the system and comparative, aggregate, and time series studies will be conducted. As interest in data based management grows, new applications will be found. Program and policy evaluation will become easier and more timely as actions are reviewed for effects which are detected by the monitoring of public use.

By monitoring public use on a limited basis, the Park Service engages the active participation of visitors in the park management process. However, to insure the application of data to the largest number of managerial problems faced by the federal government, the data collected by the National Park Service needs to be a part of an integrated statistical effort.

Federal Statistical Policy and Outdoor Recreation Statistics

The current federal activity in out-door recreation measurement suffers from a variety of problems. Standardization of procedures and documentation of methods have not been fully completed. Training and quality control over field practices needs to be improved, particularly in areas where staff resources are minimal and conventional sampling is difficult to carry out. The frequency of "estimates" as a basis for determining certain figures is much greater than is desireable. At the same time, the breadth of the current statistical accounts (generally visitor hours and visits) is of slim utility for any sort of realistic planning and of little use for site management.

A major element of the problem is more or less common to all federal offices working with minor statistical programs. A situation of perpetual negative feedback tends to exist: (1) lack of resources (staff and money) requires shortcuts in statistical procedures; (2) shortcuts in statistical procedures lower validity and reliability of data; (3) data low in validity and reliability cannot be used to solve managerial problems; (4) data which do not solve problems do not receive priority for resources; (5) lack of resources. . . Although all agree that good public use data are essential to planning and management, the negative feedback cycle creates a situation which tends to maintain statistical systems at subsistence levels.

The problem, however, is widespread. In July, 1978, the U.S. Department of Commerce issued A Framework for Planning U.S. Federal Statistics, acknowledging the various problems of decentralization generally, and specifically the numerous difficulties of minor statistical programs such as the general efforts to measure trends of outdoor recreation.

It seems safe to predict that increases in resources for out-door recreation statistical programs will not change without improving the programs themselves. Fortunately, this can be done. To the extent that a voluntary division of labor is accepted by various agencies, the combined efforts to cover segments of outdoor recreation statistics can be made cost effective. In the past an agency might look into a problem and report certain findings which would have limited utility for or be at odds with the immediate objectives of another agency. The typical response is to launch an additional study/counter-study tendency is enormously costly. Redundancy of effort can be reduced if agency social science professionals guide their agencies toward better coordination. This requires, however, that the same professionals take greater interest in what is being done around them and act to influence the work of other agencies to adequately meet the needs of their parent agency as well.

The National Park Service's effort in this direction include both sharing data it produces and depending more on data produced by other participants in the measurement of out-door recreation. The U.S. Travel Data Center, for example, conducts monthly surveys of travel from a random sample of people in the nation which promises to yield valuable data which is not available elsewhere. Combined with the

surveys of national recreation and travel conducted by the Heritage, Conservation, Recreation Service and the Bureau of Census, a formidable body of data for new area and existing site planning is available. National Park Service data are already shared with the Department of Commerce, Federal Highway Administration, and Department of Energy. Current planning work is especially sensitive to the need for information exchanges such as near areas where boundary land exist. Here again, the NPS Cooperative Park Studies Units at various universities have played a key role in improving interagency exchanges by easing the problems of information access. Ideally, creating better access to statistics at the federal and state level, including distribution to places of higher learning, will have a positive influence and will enable basic improvements to be made as higher demand justifies better support for programs to measure outdoor recreation.

Conclusion

Viewed as a series of changing tactics of measuring public use, outdoor recreation studies since WWII reflect a variety of accomplishments and shortcomings. Travel and tourism studies served to stimulate and broaden the interest of planners and managers for information about people. Visitor studies served to satisfy curiosity about a large variety of interesting questions ranging from visitor judgements to attitudes and opinions. Research studies brought methodical thinking to conceptual problems of measuring out-door recreation and sharpened the tools of study. In spite of obvious progress in the ability to contribute to planning and management, the actual adoption of public use data as a factor in planning and management has been limited. Until a better strategy for gathering data about out-door recreation can be developed, the desired impact of public use information on planning and management will not take place.

Studies may not be sufficient to communicate the applications of data to the problems of management and planning. An improved strategy will also involve the idea of continual monitoring of public use. The resulting statistics can be merged into a system which is both useful to the site manager and useful to the managerial concerns of those working at regional and national levels. Many improvements, however, need to be made in the quality of existing statistical programs. These improvements may be made without substantial additional investments if a productive voluntary division of labor can occur among agencies participating in the measurement of outdoor recreation.

These views are offered in the belief that important work has been accomplished but is undervalued. Suggestions are offered in spite of the liklihood that, here too, every solution has a problem.

LITERATURE

Although this paper has not involved direct references to other literature, the reader may have interest in related work. The following references are offered for those with continuing interest:

Buhyoff, Gregory J. 1979. Evaluation of Current Statistical Reporting Procedures and Recommended Corrective Action. Statistical Office, National Park Service, Denver, CO.

Carnie, Chere Amic. 1979. "Public Use Reporting in the National Wildlife Refuge System and Recommendations for Improvements". Unpublished mimeo. U.S. Fish and Wildlife Service, Division of Refuge Management Washington, D.C.

Duncan, Joseph. 1975. "Planning the Federal Statistical System--A Historical Perspective." A Numerator and Denominator for Change. Technical Paper 37, Department of Commerce, Bureau of Census.

Thoennes, Nancy A. and William H. Key. 1979. "Review of National Park Service Science Research". Unpublished mimeo. Department of Sociology, University of Denver, Denver, CO.

U.S. Department of Commerce, Office of Federal Statistical Policy and Standards. 1978. A Framework for Planning U.S. Federal Statistics for the 1980's. U.S. Government Printing Washington, D.C.

APPENDIX A

CAPE COD TOURIST STUDY

		COLUMN	CODE
1.	Station	1	
2.	Date	2-3	
3.	Hour	4-5	
4.	Vehicle - Sedan( ) Station Wagon( ) Trailer - Camper( ) U-Haul( ) Boat( )	6	
5.	State of Registration Mass.( ) R.I.( ) Conn.( ) N.Y.( ) Other_____ Specify	7-8	
6.	Car Occupancy Adults Male_____ Female_____ Children under 16 Male_____ Female_____	9-14	
7.	Where is your permanent residence?_____ Town State	15-20	

If on the Cape / If not on the Cape

(A) Do you have a summer residence on Cape Cod? Yes( ) No( ) — 21

Where_____ (Town) — 22-23

(B) What is the purpose of your trip?

Business and Pleasure	( )	
Personal Business	( )	
Shopping	( )	
Serve Passenger	( )	
Recreation	( )	
Vacation	( )	
Medical, Dental, Education, Church	( )	
Social	( )	
Work	( )	
Other_____ (Specify)	( )	24-25

(C) When did you arrive on the Cape?_____ (Mo. Day Time) — 26-30

If today

If prior to today

(1) Where did you stay?_____ (Town) — 31-32

(2) At a motel( ) hotel( ) cottage( ) tent( ) private home( ) other_____ (Specify) — 33-34

(D) What was the purpose of this trip?
( )vacation, ( )recreation, ( )work, other_____ (Specify) — 35-36

What was your major recreational activity?
( )beach, ( )boat, ( )fish
( )scenery ( )other_____ — 37-38

(E) Is this a rented car? ( )Yes ( )No. — 39

(F) Now for my last question, approximately how much did you spend on the Cape? $_____ — 40-43

The questionnaires were filled in by the interviewers and coded later in the office. 17,404 vehicles were interviewed out of 62,312, an overall sample of 28.4 percent. In addition to serving the requirements of the study, the stations were selected to insure the safety of the motorists and the interviewers. Wherever possible, sites were so chosen that interviews could be conducted off the road. One station was located at a rest stop. When the interview was completed, each driver received a formal note of thanks explaining the reason for the interview.

APPENDIX B

GRAND CANYON NATIONAL PARK TRAVEL SURVEY

1. Auto
2. Bus
3. Truck
4. Motorcycle

A. Residence:
   1. City or Town_____ State _____
   2. State of Vehicle Registration_____
B. Is this trip for: 1.Pleasure___ 2.Business___ 3.Both_____
C. Was your trip to the Grand Canyon the principal purpose for your trip from home? Yes___ No___
D. How many persons in your party?
   (Don't forget the babies and other children)_____
E. Where were your last two overnight stopping places before entering the Park? If more than one night in either place, then, please write the number of nights:.
   a. Last Night_____
   b. Night Before_____
   c. What type of accommodations did you use during these stops?
      1.Hotel___ 2.Auto Court(Motel)___ 3.Trailer Park_____
      4.Camping___ 5.Friends or Relatives___ 6.Home_____
      7.Other (Specify)_____
F. Overnight stops while in the Park:
   a. How many nights did you stay in the following accommodations?
      1.Hotel___ 2.Cabin with Bath___ 3.Cabin without Bath_____
      4.Camping___ 5.House Trailer___ 6.Other (Specify)_____
   b. CIRCLE your first preference above, if type desired was not available.
   c. If you are not stopping overnight, are you leaving the Park because accommodations were not available? Yes___ No_____
G. For the ENTIRE TRIP, please estimate:
   a. How many days will you be gone from home? _____
   b. How many miles will you travel on the entire trip? _____
   c. How much will you spend on your entire trip? _____
H. For that portion of your trip in Arizona, please estimate:
   a. How many days will you stay in Arizona? _____
   b. How much will you spend in Arizona? _____
I. During your stay in the Park and while enroute to and from the Park, please estimate how much you and the members of your party will spend in this GENERAL VICINITY for the items listed below: (Outer limits of this "general vicinity" includes such places as Prescott, Ashfork, Williams, Flagstaff, Cameron, St. George, Cedar City and Panguitch)

(NOTE: Include Credit Card Purchases)	TO NEAREST DOLLAR
Food	$_____
Lodging	$_____
Gas and Oil or Transportation	$_____
Other(Park Entrance Fee,souvenirs,etc,)	$_____
TOTAL	$_____

J. Where do you plan to make your next overnight stopping place after leaving the Park? No. of days
   a. Town and State_____this stop_____
   b. Please check the type of accommodations you expect to use:
      1.Hotel___ 2.Auto Court(Motel)___ 3.Camping___
      4.Trailer Park___ 5.Friends or Relatives___
      6.Home___ 7.Other (Specify)_____

K. Show order of preference with a 1, 2, 3, for the THREE features which appealed to you most in the Park:

a.	Enjoyment of Scenery_____	f.	Horseback Riding_____
b.	Mule trip into Canyon_____	g.	Ranger Talks_____
c.	Climate_____	h.	Indian Dances_____
d.	Hiking _____	i.	Wild Life_____
e.	Camping_____	j.	Evening Entertainment_____

k. Other Attractions (Specify)_____

L. Have you visited or do you plan to visit the other Rim of the Grand Canyon on this trip? Yes___ No___

COMMENTS OR SUGGESTIONS REGARDING YOUR VISIT TO THE PARK WILL BE APPRECIATED:_____

DATA BANKS FOR RECREATION SUPPLY AND PARTICIPATION[1]

E. M. Avedon & S. L. J. Smith[2]

Data archives and data banks have become increasingly important as more researchers begin to examine trends. Characteristics of data banks, sources of bias in secondary data sources and important trends in data banks are described. The paper concludes with advice about using data banks.

INTRODUCTION

Studying outdoor recreation trends presumes a source of historical information upon which description and forecast of trends can be based. In the past, forecasters relied on private sources of data, on access to private and public agency records, and on published tables and statistical documents archived in research libraries for these purposes. A relatively new source of information (primarily but not exclusively quantitative) has become more accessible to recreation researchers in the last decade, namely, machine readable data and data archives.

Since World War II, the social sciences have undergone an "information explosion". This explosion is usually evidenced by the dramatic growth in the publication of books and journals. As great as this growth has been, it does not tell the full story. The social sciences, like the physical sciences have become more quantitative, and behind every scientific publication in the social sciences there are quantitative findings upon which the report is based. However, in the social sciences more data are often collected than ever find their way into published reports. This is common practice in many contemporary surveys and historical studies. In fact, many elaborate and costly primary data gathering projects are undertaken without a clear plan for analysis of all data collected in the project.

In the past, such "excess data" would have been destroyed, but today such "excess data" represent a valuable storehouse for future research and planning and policy analyses. Information is stored without analysis or interpretation on punched cards, magnetic cards, paper tape, magnetic disk, magnetic tape, microfilm, and other mechanical or electronic media to facilitate retrieval. Unlike information stored in published form, such as the familiar printed tables of a national census, emphasis in a data archive is on rapid retrieval, custom rearrangement, processing, and summarization. Such flexibility is a boon to the myriad of potential users of any data set. Academics planners, management consultants, entrepreneurs and others can examine and use the same data set for widely differing purposes with equal ease. Although there may be different missions and policies in different machine readable data archives, and although information may vary from bank to bank, usually a data bank does not direct or control the types of information a user seeks. The secret of success and usefulness of data banks, if there is a secret, is to emphasize technology over teleology.

The concept of a data bank took hold in many facets of industrial society over the past thirty years. Airline reservations systems, department store accounting records, warehouse inventories, banking statements -- are just a few well known examples of this phenomenon. However, industrial and commercial data banks are designed for single-

[1]Presented by Dr. Smith at: OUTDOOR RECREATION TRENDS - A NATIONAL SYMPOSIUM, Elliott Alumni Centre, University of New Hampshire, April 20-23, 1980.

[2]Authors are Director and Research Associate, respectively, of the Leisure Studies Data Bank/Banque de Donnees sur les Loisirs, University of Waterloo, Waterloo, Ontario, Canada.

purpose use, by highly trained users, based upon rigidly efficient and economic systems. Within the social sciences, it was recognized that data have multiple research uses, and although some potential users may be skilled methodologists, many are relatively unsophisticated users - particularly in academic settings. Thus social science data banks could not be designed as a single purpose entity, but rather had to be designed as "banks" with multiple functions. Furthermore, data elements destined for industrial and commercial data banks were collected and prepared within the objectives of their single purpose systems. Social science data on the other hand continues to be generated through a variety of sources, and each source has its own research "perspective".

Data banks developed as a practical response to the need to handle a flood of information. One major contributor to this was the rapid development and expansion of computer technology. Social science data in the 1950's and 1960's paradoxically rendered the information generally irretrievable to the unaided researcher. To overcome this, scholars and decision-makers began to cooperate to pool information resources. They found, however, that such a project easily became time-consuming, costly, and involved great practical difficulties. If an individual did not have personal knowledge of a particular information source, the research literature often had to be scanned to track down possible sources. Once the required source of the data had been identified, it was necessary to determine if these data were still in existence, and how accessible they were to the researcher's computing resources. Often it was found that the source had not been sensitive to the possibilities of further analysis on the data collected, and as a consequence, took no pains to store the information in a manner as to allow universal retrieval and re-use.

When such data were available, they were often stored in an idiosyncratic fashion, accessible only to the source, and after a while, not comprehensible even to the source. Coding may have been ambiguous, or worse, undocumented, rendering the data useless. Formats, definitions, and terminology often reflected local, arbitrary conventions. Data documentation might contain mis-codes, missing records, and labels in a foreign language. To clarify, clean, and edit these records required more time, money, and technical expertise and sophistication than most individual researchers have. It became evident, not just to individuals, but also to universities, governments, and other agencies that retaining archivists and related personnel would be necessary to acquire, prepare, store, and retrieve machine readable social science data on a continuing basis. This recognition led to the establishment of the specialized data banks for the social sciences that exist throughout the industrialized world.

Characteristics of Data Banks

The parallel between data banks and research libraries is obvious, but there is a divergence more important than any similarities. Data banks do not attempt to archive all available information related to their specific mission. Each bank has some identified theme, and collects data in keeping with that theme. Sources of information are diverse, to the point that an "outsider" may not immediately recognize the relevance of a particular data set to the archive's mission. For example, the Leisure Studies Data Bank at the University of Waterloo primarily acquires machine readable data concerning leisure-related phenomena. This perspective normally includes survey and administrative data regarding facility or resource use, expenditures, tourist origin and destination information, performing arts audience information, park attendence information, and the like. Some information is donated by expected sources, such as the Ontario Ministry of Culture and Recreation, but data are also sought from less obvious sources such as the Federal Ministry of State for Science and Technology. In this example, we obtained a survey of the impacts of science on Canadians that included useful information about attitudes toward television viewing. Another social science data bank may acquire data only regarding political attitudes, such as the Roper or Gallup poll data, and the extent of leisure-related information in that data bank would be minimal. Some data banks acquire data for specific geo-political areas only, rather than for a specific theme. When the holdings of a bank are examined, new perspectives on both themes and data sources are possible. Collecting information on different aspects of leisure not only creates new perspectives and research possibilities, it creates a collection of skilled personnel. Data banks become the loci for contacts among people trained in data collection, data management, and data analysis. Exchanges among technicians, scholars, and policy-makers provides a rich environment for all. A data bank is more than just a warehouse of numeric information, it is a source of assistance, insight, and inspiration for the researcher.

Data are initially obtained in a variety of forms and structures, from simple tabular reports through decks of standard punch cards, complex multi-punched formats to edited and labeled variable spanned matrix system files. Most banks in archiving a data file, store information in a single medium (commonly magnetic tape), that can be read easily and

Banks may also specialize in the scope of their data. Some are depositories of national and international studies. Others might serve a national clientele, but limit holdings to provincial or state data. Often, this is not the result of a single, conscious decision by administrators, but rather the result of responses to opportunities and requests. Whatever the cause, banks generally become more specialized over time.

In addition to bias arising from objectives and specializations of a data bank, there is bias imposed by the archivists' decisions about the quality of data and documentation. A decision not to include a particular data set is based on a number of reasons. Data may be of questionable value. Records may be missing. There may be coding errors, biased sampling, ambiguous questions, or problems with instrument design. On the other hand, qualities of a file might be adequate, at least for the original purposes of research, but documentation may be inadequate or missing, and this prevents further use of data by other researchers. A decision to archive or not archive is a technical one, made on objective grounds. However, the evaluation of whether a data set or documentation meets the objective criteria is often a subjective decision made by data bank personnel; and the quality of that decision depends upon the knowledge, skill, and ability of the specific staff.

If a data set is archived and made available for use, it is not guaranteed to be free from error or distortion. The purpose of a primary data gathering project, the wording of questions, the sampling frame and design, definitions of words and terms and how substitutions were made for non-respondents, the basis upon which a test has been standardized, interpretations or shifts in meaning by the original coder, all can cause bias in reliability and generalizability. This is why documentation is so important as a part of a data bank's holdings. As a user, you should be able to assume that the data bank staff has acquired, cleaned, stored, and retrieved data properly. You cannot make any assumption, however, about the inherent quality or characteristics of the data set without examining the associated documentation.

Access rights are another source of possible bias, not so much as a distortion in the interpretation of a particular data set as it is a distortion in the information available from a data bank. The Leisure Studies Data Bank does not accept any file that is restricted to only certain users. For example, we were given a copy of a survey of violence in a specific sport. Shortly after receiving the file we were advised that this information could not be released to all classes of potential users, and that only "approved" researchers could have access to these data. It was therefore decided to de-archive this file. The decision to include only publicly available information limits the data that can be archived, but it does ensure all potential users of availability and access. This is not always the case in other social science data banks.

The final source of possible bias is that of the donor. Here we refer not only to the types of information or questions that a collector gathers, but to the original treatment of these data. Some data are distorted before being released. In other cases data are "laundered". The process of "laundering" alters the validity and accuracy of a file. The level of aggregation of observations is another reflection of each nation's political concerns and its perceptions of privacy, social responsibility and individual rights. In North America, as in many parts of the world, it is not possible for the ordinary data user to identify specific individuals by name or address. Normally disaggregation is possible only to a subgrouping short of the individual case level. Privacy extends to protecting the identity and responses of corporations as well as individuals. Protection of identities means more than just eliminating names, addresses, case numbers, and some geocodes, it can also mean aggregating responses from small or lightly populated areas to thwart attempts to deduce the probably identity of a person or corporation. For example, if you were studying private campgrounds, and had the responses of an owner in a specific local region, state, or province that reported gross income and expenditure, you may be able to narrow identity to one or two campgrounds. To prevent this, disaggregation to a local region may be limited during the archiving process.

There is substantial variation among banks in different countries with respect to data availability. In one European nation researchers can only obtain the most generalized, averaged figures for most social statistics of their population; whereas in a neighbouring country, data files are so specific that it is possible to link individual income tax returns with responses on other social surveys to check on the validity and reliability of responses concerning income and expenditure on the social survey. Many governments retain registries of disabled persons which are available for research purposes from public agencies; this is generally not the case in North America. Some data banks regularly receive official government data for permanent archiving with the intent of providing wider access of data to researchers. Some banks have no liaison

worked efficiently. Generally, data banks do not place restrictions of the size of a data set that can be archived, and although most data sets are of a manageable size for research or re-analysis, storage can become a problem when a collection grows large. Large data sets are more often a processing problem for the researcher, than an archiving problem for a data bank. Tape storage is usually in a central computer tape library under climate control and access is protected by a stringent security system. The former is to prevent physical decomposition of the tapes, the latter to prevent unauthorized use of a data file, or the inadvertent destruction of a file by an unsophisticated user. Because data are machine readable, physical distance from the tapes, or for that matter from the computer, is not a problem. Access is accomplished through electronic communication, and all a user needs is a small computer terminal, a communication link, and a telephone number. For example, although we are hundreds of miles from the Leisure Studies Data Bank -- with a terminal, a coupler, and a telephone we could access and process any of the Bank's holdings from this room! This is not unusual. At the Central Social Science Archive at the University of Cologne, users are in a building on the campus in Cologne, but the computers and the tapes are in the city of Bonn, kilometers away. At the University of Odense at the Danish Data Archive, on the island of Fyn, one of their computers and its adjacent tape library is on the island of Zeeland, in a town north of Copenhagen. At this time many researchers have access to a network of information from a number of data banks, and such practice is becoming easier as technology is developed.

Although stored numeric data form the heart of a data bank, these would be useless without supplementary documentation which explains, for example, that the fourth through the ninth columns in each record gives the total population of a region; or that specific concentrations of magnetic oxides at a certain location on a tape indicates that information concerns swimming. Documentation may be in print or may also be machine readable. In fact, the latter is becoming more common place to permit greater access to data by users from a distant terminal.

At a minimum, documentation provides the following types of information:

1. description of data structure;
2. description of data format;
3. examples of both structure and format;
4. size of data set;
5. definitions of data elements;
6. explanation of abbreviations and codes used;
7. description of sampling design and technique, substitution procedures, etc.;
8. non-response rates and weighting procedures
9. source statements that generated data, instrumentation, tests;
10. bibliographic citations for publications based on use of the data set;
11. list of related data sets;
12. names and addresses of personnel or the agency responsible for collecting the data;
13. special information regarding access or processing.

To provide data and documentation to users, data bank staff members are called upon to perform a variety of tasks. They must provide enough information about the availability and contents of holdings and how to retrieve the required data set. Behind these obvious tasks are many hours of effort that are invisible to the user. The process of archiving a data set so that it can be used has become a highly technical and exacting skill. As a bank gains visibility and its staff matures, they are expected to serve as consultants to potential users not only with respect to retrieval of data sets they have archived, but with respect to computer software, statistical procedures, and eventually even research design. In time, they are called upon to make recommendations about the process of primary data collection and storage for eventual deposit in a data archive.

Sources of Potential Bias

Because of the wealth of data available, and the pressure on archival staff to provide ever greater detailed and technical advice, data banks specialize and refine their official mission. Specialization produces an inherent bias through the type of data available, who donates data to a bank, and who would likely use a bank. Although "leisure" as a subject is a specialization within social science, some archives have close, continued relationships with certain donors, or have continued experience with certain types of users, and consequently specialize within the field of leisure studies. One bank may become stronger in the area of the sociology of leisure, another might become more skilled in the economics of leisure; some collect data only on user patterns and visitation, still others might focus only on subsets dealing with sport.

Banks may also specialize in the scope of their data. Some are depositories of national and international studies. Others might serve a national clientele, but limit holdings to provincial or state data. Often, this is not the result of a single, conscious decision by administrators, but rather the result of responses to opportunities and requests. Whatever the cause, banks generally become more specialized over time.

In addition to bias arising from objectives and specializations of a data bank, there is bias imposed by the archivists' decisions about the quality of data and documentation. A decision not to include a particular data set is based on a number of reasons. Data may be of questionable value. Records may be missing. There may be coding errors, biased sampling, ambiguous questions, or problems with instrument design. On the other hand, qualities of a file might be adequate, at least for the original purposes of research, but documentation may be inadequate or missing, and this prevents further use of data by other researchers. A decision to archive or not archive is a technical one, made on objective grounds. However, the evaluation of whether a data set or documentation meets the objective criteria is often a subjective decision made by data bank personnel; and the quality of that decision depends upon the knowledge, skill, and ability of the specific staff.

If a data set is archived and made available for use, it is not guaranteed to be free from error or distortion. The purpose of a primary data gathering project, the wording of questions, the sampling frame and design, definitions of words and terms and how substitutions were made for non-respondents, the basis upon which a test has been standardized, interpretations or shifts in meaning by the original coder, all can cause bias in reliability and generalizability. This is why documentation is so important as a part of a data bank's holdings. As a user, you should be able to assume that the data bank staff has acquired, cleaned, stored, and retrieved data properly. You cannot make any assumption, however, about the inherent quality or characteristics of the data set without examining the associated documentation.

Access rights are another source of possible bias, not so much as a distortion in the interpretation of a particular data set as it is a distortion in the information available from a data bank. The Leisure Studies Data Bank does not accept any file that is restricted to only certain users. For example, we were given a copy of a survey of violence in a specific sport. Shortly after receiving the file we were advised that this information could not be released to all classes of potential users, and that only "approved" researchers could have access to these data. It was therefore decided to de-archive this file. The decision to include only publicly available information limits the data that can be archived, but it does ensure all potential users of availability and access. This is not always the case in other social science data banks.

The final source of possible bias is that of the donor. Here we refer not only to the types of information or questions that a collector gathers, but to the original treatment of these data. Some data are distorted before being released. In other cases data are "laundered". The process of "laundering" alters the validity and accuracy of a file. The level of aggregation of observations is another reflection of each nation's political concerns and its perceptions of privacy, social responsibility and individual rights. In North America, as in many parts of the world, it is not possible for the ordinary data user to identify specific individuals by name or address. Normally disaggregation is possible only to a subgrouping short of the individual case level. Privacy extends to protecting the identity and responses of corporations as well as individuals. Protection of identities means more than just eliminating names, addresses, case numbers, and some geocodes, it can also mean aggregating responses from small or lightly populated areas to thwart attempts to deduce the probably identity of a person or corporation. For example, if you were studying private campgrounds, and had the responses of an owner in a specific local region, state, or province that reported gross income and expenditure, you may be able to narrow identity to one or two campgrounds. To prevent this, disaggregation to a local region may be limited during the archiving process.

There is substantial variation among banks in different countries with respect to data availability. In one European nation researchers can only obtain the most generalized, averaged figures for most social statistics of their population; whereas in a neighbouring country, data files are so specific that it is possible to link individual income tax returns with responses on other social surveys to check on the validity and reliability of responses concerning income and expenditure on the social survey. Many governments retain registries of disabled persons which are available for research purposes from public agencies; this is generally not the case in North America. Some data banks regularly receive official government data for permanent archiving with the intent of providing wider access of data to researchers. Some banks have no liaison

with government and have access only to academically generated data. Still others may only have access to commercially produced data. Still others may only have access to commercially produced data. Availability of social science data for research thus differs from nation to nation, and this bias influences the scope of data a researcher has for analysis.

Trends in Data Banks

One of the most important trends is the apparently contradictory tendency to become both more specific and more general. The growing specificity of a bank is the result of continued relationships with certain donors and users. A bank that develops a good working association with particular agencies, tends over time, to specialize in the information these agencies provide and need. A subtle, but important force thus slowly influences the mission of each archive. Growth of other data banks into related subjects also encourages specialization.

At the same time, there is a broadening in the perception of data files useful to researchers in a specific subject. In past years, only data files obviously related (as indicated by title of a file) to the mission of a bank would be archived. There is now a recognition that data files from unlikely sources can be of great value. For example, we have recently obtained a copy of projections used by the local public school board for educational planning. This file contains information on the number, gender, and ages of children expected in the regional population in years to come. Although this information was collected for educational planning, it is also valuable for doing feasibility studies and needs assessments for public recreation facility planning.

Another trend in data banking is toward use of more efficient and generalized technology. The newest generation of computers, and the availability of new memory technology will speed the time necessary for data processing, reduce computing costs, and increase available storage. Software packages, such as SPSS, SAS, BMD, OSIRIS, and others have been improved to match improvements in hardware and operating systems. It is now possible to work with files that would have been considered monstrous only a few years ago. The Leisure Studies Data Bank, for example, regularly assists users in working with files that contain over 1,000 variables or that have as many as 50,000 cases.

In addition to greater power and efficiency, there is a trend to greater flexibility and compatibility. Work is proceeding to develop a universal interchange file that will facilitate linking data sets or the output of one system with any of several software packages.

Just as data banks were developed to help researchers cope with data, archivists are beginning to see a need for providing assistance in helping potential users cope with the growing number of data banks. The first step is the development of a system to allow a potential user to query a bank's holdings for information about a specific topic, geographic region, or other characteristics of a data set. Because of the kind of specificity inherent in different thematic research approaches, efforts are underway to develop hierarchical modes of inquiry that may be shared among several cooperating data banks. A user at the Leisure Studies Data Bank for example, will one day be able to browse not only LSDB holdings, but leisure-related data that is part of the holdings of other universities in other countries. The significance of this system of hierarchical study descriptions or file precis is not only to allow an efficient search to be made, but to provide for common terminology and descriptions among several archives.

This growing cooperation is international in scope, and thus there is a trend in the polylingualization of archives. The international language of computers may be English, but file precis, variable labels, catalogues, and the like will need to become available in all the major languages of the world. Several data archive organizations have been established to encourage system and file exchange, cooperation and consistency among member archives. In 1965, as one example, the Council of Social Science Data Archives was established to further these goals among two dozen United States archives. Unfortunately, the differences of opinions among members was so great that the council collapsed. This is a problem and a challenge to data banks in the United States. In Canada, data archives, government, academic, and private are members of the Canadian Data Organization Committee of The Social Science Federation of Canada. A similar organization exists for Western Europe. One of the most important organizations promoting inter-archival cooperation today is the International Federation of Data Organizations, an associate member of the International Social Science Council - a UNESCO organization. Member archives are from both east and west Europe, the United States, and Canada.

Perhaps the last major trend to cite is the growing importance of data bank personnel as research consultants. Because of familiarity with different problems in research design and analysis, they acquire an overall perspective on the production and use of new data collections. Our own staff have provided consultation to a number of government and private organizations on the design of data gather-

ing projects, on coding and weighting of data after collection, and on other related matters. Data banks also have the potential to organize groups of individual users to pursue new lines of inquiry. Because of the potential to serve as "spokespersons" for both data and computer users, these personnel can help to provide the impetus for developing computer software systems, and can become effective spokespersons for social scientists to the computer industry.

Using a Data Bank

Users of data banks fall into a number of relatively well identified groups: the unsophisticated new researcher who has not been a data user, and knows almost nothing about computers; the researcher who has had some primary research experience and some familiarity with data, but little computing experience; the experienced researcher who has considerable methodological skill, and computer literacy; the sophisticated researcher and computer user. Each of these types of users require different types of assistance from a data bank staff. The more unsophisticated a potential user is, the more likely the first few visits to a data bank will be a "fishing expedition". The more sophisticated a user is, the more specific and technical is the use of a data bank. Preliminary visits to a data bank by any user concern documentation rather than data, regardless of the level of sophistication.

In order to ensure that users have access to the archived data that will be of most use to them, documentation is usually organized on five levels.

1. File Identification -- a user may discover that a data file exists from an entry in a library catalogue, an inventory of data sets, a data bank catalogue of holdings. These citations are often cryptic, and will often include only the name of a file, and some general identifying information.

2. File Precis -- a user, upon discovering a data file that may meet research needs, then proceeds to examine more detailed information about a file. File precis are available in some library reference rooms, in data bank offices, and many are available in a hard copy form that can be mailed from a data bank to a potential user. Some banks offer machine-readable precis that can be examined on-line through an interactive computer system at a CRT terminal. A file precis describes the data, presents sampling and weighting information, indicates types of variables in the file, lists published reports based on these data, and provides a summary of the research which generated the data.

3. Source Documents -- after examining the file precis, a user may wish to consult some of the sources cited in the precis. Depending upon the practice of the bank, some data archives in collaboration with libraries have catalogued data files in the same manner as related source documents. Thus a user can examine related sources documents within a library or through inter-library loan by using the same catalogue classification codes. In many instances, a data bank will also have copies of these source documents, but these are usually for on-site use.

4. Variable Lists -- generally these are available for in-depth examination of a data file for use when designing a research plan. Hardcopy lists are usually available for each file and can be mailed to a potential user. Sometimes these lists include the source statements (from a questionaire or psycological test, etc.) which generated the variables in the data set.

5. Codebooks -- offer the user who has developed a research plan for use of a specific data set detailed information regarding the characteristics of the file structure, a complete listing of code elements specifying the variables and all values for each variable. Univariate tables in a codebook specify the frequency for each value within a variable, precise information about codes needed for processing, matrix information, CPU requirements and the like. Codebooks may be in hardcopy and sent through the mail, or may be machine readable and available on-line locally or can be used at a distance from the bank with the necessary hardware and software, and software systems documentation.

A word should be said about the nature of the files that are usually available within a bank. There are three basic types of files: raw data files, edited system files, and process produced files.

1. Raw data files are files that have been cleaned for errors, wild codes, etc. These files are stored in their cleaned "raw" form. Generally they are available to a potential user that wishes to write a special analysis programme for these data, and does not wish to use a package programme.

2. Edited system files are files that have been put through the archive process and been prepared to be used with a package analysis programme such as SPSS, SAS, OSIRIS, BMD, etc. There are often different versions of these files, such as simplified editions for new student researchers that may have some of the ordinal data grouped and labeled for simplified analysis, or the file might be reduced to a sample of cases to enhance useability when there are large numbers of records. Or a special file might have been created with an inverted matrix for use in certain factor analytic programmes by experienced researchers.

3. Process produced files are computer-generated files, based upon a researcher designed programme. Input for these files will be variables from a number of different files. The final file will be a "raw" data file, distinguished from researcher-collected data because data collection is actually internal to the computer used.

The process of using a data bank is generally the same in most locales. First a user needs to have a computer account number in order to access the required data. Access in many places today can be either through batch mode or through an interactive system. A user can either do all of his own computing or can seek assistance from data bank staff. In effect, the process is analogous to designing any research project, the only difference is that after designing the project (in this instance with the use of the data documentation) and instead of going out to collect data from households or on-site, or from some administrative reports or documents, the researcher writes a computer programme and collects the required data from the computer--data that have been stored by the data bank. Although this seems somewhat oversimplified, any working researcher knows just what all that simple formula can involve.

In many instances, there really is no need to expend the time and money to collect new data, particularly concerning leisure-related topics because there is a wealth of data available today that has never been subjected to analysis. Perhaps the biggest problem leisure researchers face today is identifying where useable data reside. Because finding these data is so difficult, the tendency for researchers is to develop research designs that include new data collection. This is not only a common practice within the field of leisure studies, but throughout the social sciences. However, as data banks become more common in universities, and are used more frequently as laboratories in the teaching of social science research, there will be a greater tendency for new researchers to think about analysis of secondary data before embarking on new data collection.