COUNTING VISITORS AT NATIONAL PARKS: CONCEPTS AND ISSUES

Jay Beaman

Director, Socio-Economic Branch, Canadian Parks Service, Ottawa, Canada K1A 0H3

Dick Stanley

Chief, Socio-Economic Information Division, Canadian Parks Service, Ottawa, Canada K1A 0H3

Unless attendance is treated as a multidimensional concept, it can mislead a manager as to what is happening in his park. The Canadian Parks Service recently revised its attendance measures so that they can provide both meaningful information about individual parks and be used consistently in all parks.

1. The History of Attendance in the Canadian Parks Service

In the sixties and early seventies, many myths grew up around attendance reporting in the national parks of Canada. Some park managers firmly believed that budget allocations for their parks depended in some direct and rational way on levels of attendance. Some managers believed that their positions depended on attendance levels. There was also a widely held organizational view that it was important that attendance be increasing from year to year, even if one was not sure exactly what attendance was. On the other hand, some managers diligently sought an accurate measure for attendance in the belief that it could tell them something useful, even indispensable, about their operations, and in particular, workload.

Of course, the number of visitors has something to do with workload in a park. It is almost axiomatic, and like a lot of axioms, we do not often think much about it. If a park offers services to visitors or tries to influence their behaviour (e.g., tries to keep them from destroying the resource), then more visitors means more serving and more influencing. However, number of visitors has a very complex (and poorly understood) relationship with workload. And that relationship varies from place to place and from time to time. Where things are complex, not well understood, and surrounded by myth, you would expect a lot of confusion, and a diversity of approaches. Not surprisingly, therefore, there were a lot of different things being reported as park attendance across the system of parks.

In the mid-seventies, the "socio-economic" group at the Canadian Parks Service headquarters was charged with the task of putting some order into attendance counting and reporting. We took our mandate seriously, and tried to find and correct arbitrary and unsound definitions and ad hoc and unstable counting methods. Our aim was to transform attendance measures with little or no management utility into new reporting formulas so that attendance figures would show us what was going on in parks, or at least would signal events that required management attention.

While trying to untangle the confusion, we recognized that visitors do not all impact workload the same way in every circumstance. We came to understand that different parks were going to need different attendance measures: there was no one best way to do the job. For the same reasons, we also recognized that attendance was not necessarily a useful measure at all parks. Still, for political reasons, it was necessary to count attendance everywhere. Can you imagine a park manager not being in a position to answer the question "How many people visit your park?"

As a result, in some parks we had ways of counting attendance that measured nothing other than entering traffic. In some cases, we measured party nights of camping, and in other cases, we measured entries for the purpose of visiting the park and using the services provided. The reason for this variety is that in some parks, the primary workload is related to pass-through traffic. Attendance that does not include pass-through traffic (even though these people can only loosely be called visitors) does not reflect true park workload. In other parks, virtually everyone who comes to the park spends time overnight camping. Party-nights of camping is therefore a good measure of the park's workload. In some parks, there is a mix of day-users, campers, sightseers, and some pass-through traffic. The best solution here seemed to be to identify, through surveys, those persons who enter the park and make some use of it consistent with our primary mandate of heritage appreciation and recreation and calibrate their proportions against some easily recorded measure.

It is important to recognize that attendance has not been the only statistic reported by the Canadian Parks Service. Since 1971, a variety of use statistics, reflecting day to day operations, have been recorded by the parks. For example, if you wanted to know if camping in a park was increasing, you did not look at attendance, you looked at the camping statistics. If you wanted to know if vehicle entries to a park were increasing, you looked at vehicle entry figures by gate. If you wanted to know what was happening in interpretive programs, there were interpretive program utilization statistics.

2. The Search for Consistency

We did not, however, live happily ever after. After over ten blissful years of relatively satisfactory attendance reporting, a new force arrived on the scene: the Auditor General of the Canadian Parliament. In 1983, a comprehensive audit was conducted of the Canadian Parks Service. The Socio-Economic function, of course, was one of the groups whose activities were reviewed. Where we saw healthy and useful diversity in attendance reporting, the Auditor General's team saw inconsistency and cross-purposes. They looked at the attendance calculating formulas for different parks and saw that different things were being measured. Therefore, according to them, no overall measure of attendance was possible, so that what we had been reporting nationally as attendance was erroneous and misleading.

Attendance formulas were being used consistently from year to year at individual parks, and more and more parks had attendance figures that had value to park management. However, this was not enough to make the attendance measures acceptable to the auditors. The fact that attendance was virtually defined as pass through traffic in some cases, as campers in another, and as something else in other cases was viewed as inconsistent. There is no denying it, it was inconsistent! To be fair however, the Auditor General was looking for something quite different from what we had been aiming for. The Auditor General's Office wanted a "performance indicator" for the whole of the Canadian Parks Service, a single number used consistently at over 100 parks and historic sites which could be used to indicate to the

Canadian Parliament how the Service was performing. We, on the other hand, were trying to find measures in each park and historic site that would indicate to individual park managers something about how their individual parks were doing. So the problem was not a question of inconsistent attendance measures, but of inconsistent objectives.

Of course, the search for the single number to describe performance is an activity surrounded by at least as many myths as once surrounded attendance measurement. In fact, it is somewhat akin to the search for the holy grail: the biggest myth is that it exists at all. If attendance is to mean anything at all to anybody, it must somehow relate to workload at a park. However, through a somewhat Arrow-like paradox, what is meaningful at the individual park level is not meaningful when it is aggregated to the system level. This is, of course, because visitors affect workload differently in different parks, and there is no common denominator like the dollar to reduce all this diversity to a comparable unit so that a single revenue or profit or return on investment figure can be calculated. So what do you do to comply with the Auditor General's direction?

3. Analytical and Theoretical Approaches to Defining Dimensions

You will have noticed that in this paper, we have avoided using the term visitation, a term which has often been used to describe the number of people who visit a park. The Concise Oxford Dictionary defines a visitation as an inspection by a bishop, or a divine dispensation of punishment or reward. We certainly welcome any bishops who would like to visit our parks, and auditors as well (there appears to be a part of the definition for each of them), but these kinds of visitations are only one dimension of park attendance. Our concern with the Auditor General's comments was that there were too many important dimensions to pick just one.

Dimensions are of course used in science and in common speech to describe things clearly and consistently. (Beaman and Grimm, 1989) Imagine a series of boxes, all of different sizes. We could go a long way to describing them merely by telling you their length, width, and height. The mathematically inclined among you will immediately see that we have characterised a them using a three-dimensional vector. Limiting ourselves to any one number or dimension to describe the boxes will give an incomplete picture. (see Cooley and Lohnes, 1961, for a factor analysis example)

Now "dimension" is often used very loosely in the social sciences, as in "the dimensions of a problem". It can however be used with rigour. Take a variety of variables that measure things that go on in a park, such as answers to a questionnaire. The values of these variables can be arrayed as a table or matrix. If there are 20 variables, any one set of 20 values for those variables (one respondent's questionnaire) defines a point in a 20-dimensional space. By factor analysis, one could determine how much of the variance in that space appears to be concentrated around one, two, or three, or a limited number of underlying factors or dimensions. (Harman, 1976 and Muliak, 1972).

Operations in a park can be "explained" by visits using just such a multi-dimensional model. Some sort of "complexity of operations" or "workload" variable could be specified as a function of a variety of other variables, as in the following:

> WORKLOAD = f (number of visits of various types, number of vehicles entering, number of picnics, length of stay of various visit segments, number and complexity of programs run for these visit segments, number of campers, ...)

If all these variables could be defined operationally (especially the dependent variable "workload", which has eluded us for many years), and appropriate data collected, a factor analysis could be run to reduce the variables to their underlying factors. In fact, we suspect there would probably be around three or four that would explain 90% or more of the variance.

What was just described was an analytic approach to finding dimensions. There has however been much reaction in the social sciences to applying methods without knowing what problem was to be solved (Beaman, 1977). The alternative to ad hoc use of methods is a theoretical approach, that is, an approach which identifies quantitative measures that are understandably related to the main aspects of the business that we are in, and that do not just duplicate each other. For example, the main business of some of our parks, and the item which absorbs the majority of the budget of those parks, is servicing the major highways that pass through them. The measurement of the number of entries by persons to that park, regardless of the purpose of their entry, provides an indicator that reflects a major workload of such parks. On the other hand, if we measure the number of people who stop in a park to do something, such as participate in a recreational activity or an interpretive program, we have a measure of another kind of workload, and it is "somewhat" independent of the number of entries.

Because there are as many ways to measure visitor behaviour as there are analysts to do it, the Auditor General's quest for a measure of attendance which is both valid and generalizable to all parks led us to make somewhat arbitrary selections from among the various alternatives available. Taking a general view, we reasoned, three things can happen at a park. A visitor passes through the park on his way to somewhere else, the park being a convenient travel corridor. A visitor stops for a brief time to participate in an activity (e.g., a picnic). A visitor stops for a significantly longer time to participate in a series of activities, often camping overnight. This kind of split provides a good general purpose way of presenting what happens. We are talking about very generalized behaviour when talking about reporting attendance for the Auditor General's purposes).

The time dimension can be tricky to capture here. It may be useful to split attendance into only two segments: those who enter just in order to pass through, and those who enter for park related purposes (heritage appreciation or recreation) for any period of time, long or short. Once the latter group is identified, it can then be measured two ways, by absolute number, and by some measure of duration, since clearly, the longer you stay, the more workload you represent, although here again the relationship is complex, regrettably ill understood, and of course, surrounded by myths.

The result of our considerations was a set of four measures: the person-entry, the person- visit, the person-visit-day and the person-visit-hour. A person-entry occurs whenever a person enters a park for any purpose. A person-visit occurs when a person enters a park for the first time on a given day or for the first time on the first day of his stay in the park, and his stay in the park is for the purpose of heritage appreciation or recreation, our mandated reasons for having the park. A person-visit-day occurs when a visitor stays in a park for a day or part day. Each day he or she stays counts as an additional person-visit-day. The person-visit-hour measure is used for visits to places like historic sites, where the visitor cannot stay overnight (there are no facilities provided for accommodation), but where different lengths of stay, ranging from minutes to hours, can be significant for management.

4. Testing Our Measures

Were our measures good ones? We tested the attendance data using factor analysis to see if we could reduce the number of variables from three. Then we submitted the data to cluster analysis to see if the measures provided any meaningful perspective on what was happening in parks.

To conduct these tests, we used person-entry, person-visit, and person-visit-day data for all national parks for the months of July and August for 1988, 1989, and 1990. For the purposes of this paper, we restricted ourselves to national parks. Since, as pointed out above, historic sites do not have overnight accommodation, have a length of stay of hours or minutes, and rarely have pass-through traffic, historic sites should be considered as different spaces, measured against different dimensions.

We have also limited the data in our example to the peak summer season (July and August). Some national parks do not operate in the winter. Others have radically different "operational" parameters in winter since accommodation, trail use, etc.,play a different role in park use at that season. The difference reflects both what can be offered and the different markets being served. We conducted a factor analysis on the three attendance variables person-entry, person-visit, and person-visit-day. Tables 1 and 2 show the results of two different factor analyses. The analysis in Table 1 shows the initial result.

Table 1. Results of factor analysis on untransformed variables.

VARIABLE	FACTOR LOADINGS		
	FACTOR 1	FACTOR 2	FACTOR 3
person-entry	0.92		
person-visit	0.97		
person-visit-day	0.98		
eigenvalues	2.769	0.213	0.019

All the variables load on Factor 1. One factor explains over 90% of the variance. However, it turns out that this is because the sheer scale of visits to parks overwhelms all other variation. Some parks receive as many as 650,000 person-visits in a month (with a commensurate number of entries and visit-days) while some parks get as few as 15 visits. Any subtle (or not so subtle) relationship between person-entries, person-visits and person-visit-days is wiped out by the fact that high attendance parks also have high person-entries and high person-visit-days as well.

We performed the factor analysis again on data transformed to make two of the variables "independent" of the third. We divided person-entries and person-visit-days by person-visits in order to turn them into the ratios "entry per visit" and "visit-day per visit". The results of the factor analysis are shown in Table 2.

We now had three variables. "Visits" showed the number of visits as before, with the influence of scale left in (high attendance parks still had big numbers; low attendance parks had low numbers). "Entries per visit" showed the proportion of entries to visits: where there was a great deal of pass-through traffic, the proportion of entries to visits should be high, regardless of whether the park had high or low attendance. "Visit-days per visit" indicated the proportion of long stays to one day visits: Where visitors stayed a long time, the ratio would be high; where visitors tended to stay only for a day, the ratio would approach one, again regardless of the scale of person-visits to the park.

Table 2. Results of factor analysis on variables transformed to eliminate effects of scale.

VARIABLE	FACTOR LOADINGS		
	FACTOR 1	FACTOR 2	FACTOR 3
entry/visit	-0.82	-0.06	
visit-day/ visit	0.41	0.78	
visit	0.55	-0.67	
eigenvalues	1.140	1.058	0.802

The factors with eigenvalues ranging from 1.1 to 1.06 to 0.80 show that if a reasonably large proportion of the variance is to be explained, three dimensions will be needed.

The conclusion is that there is no single relationship, particularly a linear one, that lets person-entries, person-visits, and person-visit-days be explained by each other. You cannot predict visits from entries by a general formula that applies to all parks. You can develop such a formula for a particular park for a peak summer season, but the formula will, at most, apply to a small group of parks that by chance have the same multipliers.

To see if we could use the three dimensions of attendance data to draw conclusions about how parks were behaving we used cluster analysis (Everitt, 1980). We ranked all the data using the "SAS PROC RANK" procedure and created a transformed normal distribution value for each ranked value using Tukey's algorithm. This is the same transformation as mentioned above, and had the result of eliminating the effects of scale. We then used the "SAS FASTCLUS" procedure to derive clusters from these transformed data. The results are shown in Table 3.

Table 3. Results of the cluster analysis.

CLUST #	N	CENTROID COORDINATES			CENTROID DISTANCE
		ENTRY/ VISIT	DAY/ VISIT	VISIT	
1	2	2.69	0.14	-2.69	2.21
2	13	0.80	1.92	-1.96	1.90
3	28	0.35	0.90	2.69	1.55
4	11	2.28	-1.60	-0.34	1.63
5	31	-1.68	-0.17	0.45	1.43
6	12	0.52	-1.60	1.12	1.03
7	26	-1.68	-2.69	-0.80	1.63
8	19	-1.68	1.17	-1.32	1.59
9	7	0.63	-0.75	-2.19	1.47
10	39	1.42	0.06	0.19	1.03

As can be seen from the centroid distances, a number of distinct clusters arose from the analysis. An easier way to understand this information is to list the parks which make up each cluster. This is done in Table 4. Recall that what were clustered were data from individual months, so it is possible for a park to appear in more than one cluster. It is a measure of the distinctness of the clusters that this "splitting" of a single park into several clusters occurred, with one exception, only to parks that are remote or extremely hard to get to, and so receive very few entries, visits or visit-days. The numbers in brackets are the number of months of attendance of a park that occurs in a cluster.

Table 4. Parks arranged in clusters.

CLUSTER 1: Northern Yukon (2)
CLUSTER 2: Auyuittuq (4), Northern Yukon (3), Ellesmere Island
CLUSTER 3: Banff, Jasper, Kootenay, Riding Mountain (4), Cape Breton Highlands
CLUSTER 4: Glacier (5), Mount Revelstoke
CLUSTER 5: Kejimkujik, Pacific Rim, La Mauricie, Kluane (1), Elk Island, Point Pelee
CLUSTER 6: Forillion, Prince Edward Island
CLUSTER 7: Pukaskwa (1), Kluane (5), Nahanni (1), Grasslands (2), Mingan Archipelago, St. Lawrence Islands, Wood Buffalo (4), Georgian Bay Islands (1)
CLUSTER 8: Pukaskwa (1), Nahanni (5), Auyuittuq (2), Prince Albert, Georgian Bay Islands (5)
CLUSTER 9: Pukaskwa (4), Northern Yukon (1), Wood Buffalo (2)
CLUSTER 10: Waterton Lakes, Yoho, Glacier (1), Riding Mountain (2), Fundy, Terra Nova, Kouchibouguac, Gros Morne

Anyone familiar with the Canadian Parks Service parks will have a good feeling about the groups as they are displayed above. Quite frankly, it was a surprise that we obtained this good a grouping from the attendance data. For example, Cluster two consists of very remote northern parks, with no roads, let alone any pass-through traffic. When visitors get there, they tend to stay for a relatively long time, and probably camp in the park. Cluster three consists of parks that have mixed use, with a lot of pass-through traffic, a lot of person-visits lasting only one day (e.g., skiing at Banff) but a lot of long stays for the world class attractions, particularly in Banff, Jasper, and Cape Breton Highlands. Cluster five parks can be characterised as end use parks, appealing mostly to regional residents who would use the parks over the weekends, but have very little pass through traffic. Finally, cluster ten consists of parks which have significantly more pass-through than five, but similarly attract a lot of regional short term visitors.

5. Uses and Limitations of Attendance Measures

Table 4 shows the clustering of the parks based on the three attendance dimensions. It should be noted however that what happens in some parks is a reflection of circumstances and not what was planned. CPS may have planned that the parks which happen to be in a cluster should in fact be quite different. Examination of target markets for each park in a cluster, and the "opportunity structure" for each park could well imply that different development options should be followed. This may be particularly true if sustainable tourism has become a driving objective for new development. The attendance measures may not provide enough information to allow an analyst or planner to define the clusters that parks should be in. They do , however, give a perspective on what is happening in the park. For general managerial tracking, the three attendance measures provide enough information to give some insight, but do not provide so much detail as to be unmanageable.

To answer the kinds of questions that might be raised, however, data beyond attendance must be examined. In the Canadian Parks Service, data that are to be regularly captured and monitored are being defined in a data plan for each park. The data plan must be based on the nature of the park and what is likely to happen to it. Furthermore, because money is scarce, the plan must be developed with due consideration of what is being collected anyway, what can be collected cheaply by automation, how accurate data must be and how often they need to be collected if they are to be useful.

What is wrong with combining the three measures to get one? In terms of the earlier discussion of the way to describe boxes, there are several ways to combine length, width and height, depending on what you want to do with the boxes. If you are going to keep liquid in them, you could combine length, width, and height into volume, and have a good measure. If you want to store solid objects in them, volume is a much less useful measure. If you want to minimize space wasted, then you have to know all three dimensions in order to match your objects with the correctly sized box. Therefore, what is the "best" measure depends directly on the problem you are trying to solve. In one park, person-entries can drop as person-visit-days increase. In another park, person-visits can increase as person-visit-days decrease. Person-entries to parks can rise and fall as exogenous factors influence the volume of pass-through traffic, yet a park's workload can remain unchanged. In another park, workload can primarily be determined by entries. All three measures are needed to provide a measure of attendance which will be useful for the variety of issues parks managers face. It is important to realize that the three measures are not selected on the basis of mathematical analysis. They were selected on the practical basis of being general yet useful.

6. Conclusion

The multi-dimensionality of attendance is a concept that can easily be applied to pool attendance, community centre attendance, festival and special event attendance and so on. This fact should be kept in mind when one sees a statistical report or an entire statistical system being designed around one statistic for each facility. Worse yet is a situation where the same statistic is reported for a variety of facilities or services regardless of its appropriateness to represent the work load of the facility or service. If management wants attendance measures that are consistently defined for different locations, and yet measure something remotely meaningful and useful, they must accept that they will have multiple measures which cannot be reconciled to a single "index" that gives an "overall" perspective on what is happening.

Attendance reporting is a political necessity. It can be little more than supplying numbers to meet political requirements. The emphasis of this paper has been that attendance can be consistent, relevant, logically justified and logically related to general data collection plans. It is to be hoped that more and more attendance reporting will be of the type just described.

Literature Cited

Beaman, J. 1977. Education for tourism research for the 1980's. In: The 80's - It's impact on travel and tourism marketing. Salt Lake City, Utah: The Travel Research Association and Bureau of Economic and Business Research, University of Utah: 33-35.

Beaman, J.; Grimm, S. 1989. Park use related data recording: A new direction for the Canadian Parks Service. In: Proceedings of the 1989 Northeastern recreation research symposium. April 3-5, 1989, General Technical Report NE-132, United States

Department of Agriculture, Forest Service, Northeastern Forest Experimental Station: 69-76.

Cooley, W.; Lohnes, P. 1961. Multivariate Procedures for the Social Sciences, New York: John Wiley and Sons.

Everitt, B. 1980. Cluster Analysis. London: Heinemann Educational Books.

Harman, H. 1976. Modern Factor Analysis. Chicago: University of Chicago Press.

Muliak, S. 1972. The Foundations of Factor Analysis. New York: McGraw Hill.