This paper reviews the question of the cost and benefits of data for the analysis of state-policy questions, and concludes that in the abstract the question cannot be answered; rather, it must be considered on a continuing basis whenever data is to be collected in the support of baseline studies or special projects. The paper suggests six criteria which should be considered in determining how an agency will go about acquiring the necessary information to accomplish its mission, and makes the points that addressing these considerations in an open and candid manner and involving the various parties to the data flow should significantly improve the relationship of benefits to costs in data collection.

(Author/DEA)
THE STATE COUNCIL OF HIGHER EDUCATION AS COLLECTOR AND TRANSMITTER OF DATA

Session Number 3.07

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Denis J. Curry
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Submitted by:
Denis J. Curry
Council for Postsecondary Education
State of Washington
ABSTRACT

This paper reviews the question of the cost and benefits of data and concludes that in the abstract the question cannot be answered, but is one which must be considered on a continuing basis whenever data is to be collected—in the support of baseline studies or special projects.

The paper suggests six criteria which should be considered in determining how an agency will go about acquiring the necessary information to accomplish its mission. These are:

1. Are the data necessary to provide either relevant baseline information or to meet the terms of specifically mandated responsibilities or studies?

2. Are the data reasonably available?

3. Is a suitable proxy for the information available in reports which are already being provided to other state or federal agencies?

4. Is comprehensive or census-type information necessary or can the objectives be met through sampling?

5. Can the information be best gathered through survey instruments or through site visits on the part of the agency staff?

6. Is it apparent that the collection of the data will be a recurring or a one time occurrence? If recurring, on what time frame should the collection be based: quarterly, annual, biannual, or greater time interval? Is that time interval justifiably based on evident or perceived need?

The paper makes the points that addressing these considerations in an open and candid manner and involving the various parties to the data flow should significantly improve the relationship of benefits to costs in data collection.
Is the cost of data worth it? This is a question which oddly enough cannot be answered in the abstract, but which must be considered continually by any agency of government which is charged with the functions of data collection, evaluation, analysis, and the preparation of policy recommendations. For example, the basic charge to the Washington Council for Postsecondary Education is to "engage in overall planning for postsecondary education in the state which shall include the collection and analysis of necessary data from public and, where appropriate, private institutions of postsecondary education." An important word is contained in that charge, and that word is necessary. This is an important modifier since it conditions the terms on which data are collected, transmitted and/or used in the analysis of policy questions.

The basic directive from the legislature is related to certain purposes. These include the assessment and definition of educational needs; the determination of whether defined needs are being met; and the determination of priorities among the defined needs and the specification of resources necessary to meet them. As you can see, this is a broad charge; and it becomes increasingly so when a state coordinating agency is presented with a variety of legislatively mandated special studies. In one recent year, the Washington Council, with a professional analytic staff of ten persons, was presented with thirteen separate legislative resolutions for studies in addition to its basic responsibilities.
The type of responsibilities I have mentioned requires a state agency to be both complete and discriminating in its attitude toward data collection. The following criteria, of necessity, must be employed in determining how the agency will go about acquiring the necessary information to accomplish its mission. These include:

1. Are the data necessary to provide either relevant baseline information or to meet the terms of specifically mandated responsibilities or studies?
2. Are the data reasonably available?
3. Is a suitable proxy for the information available in reports which are already being provided to other state or federal agencies?
4. Is comprehensive or census-type information necessary or can the objectives be met through sampling?
5. Can the information be best gathered through survey instruments or through site visits on the part of the agency staff?
6. Is it apparent that the collection of the data will be a recurring or a one time occurrence? If recurring, on what time frame should the collection be based: quarterly, annual, biannual, or greater time intervals? Is that time interval justifiably based on evident or perceived need?

In addition to its responsibilities within the state, a council of higher education or postsecondary education (as the current phraseology often requires) is also engaged in certain aspects of the flow of information from institutions to the national government. What role should it play and to what extent should the state agency attempt to utilize that information flow to avoid duplicate data collection requirements? The body of
this presentation will include a discussion of these questions and considerations both from the point of view of the state as initiator of data requests, and its role as a service entity in the transmission of information both to the state, other interested parties within the state, and the federal government.

Before proceeding with a discussion of those topics, it might be worthwhile to first discuss the question which is the topic of this symposium, "Is the Cost of Data Worth It?". I indicated that this was an abstract question since the answer must of necessity be subjective and will vary with the perspective of individuals involved. For example, let me recount a personal experience. I like to teach (on a part-time basis) whenever my schedule permits and when I put on my teaching hat, I am distinctly offended by the need to fill out official institutional forms even though, in the abstract, I can recognize that these probably serve some useful purpose. When I must interrupt my planned discussion of the topic, which of course includes many gems of wisdom, to engage in administrative functions, I get upset. From the standpoint of the instructor, I can see little if any benefit to me in the collection of information which escapes into the vast maw of the administration and is used for who knows what arcane purposes. In this instance, there is a cost (in terms of time) and no evident return.

On the other hand, from the perspective of an individual in charge of collection and analysis of information involving multiple institutions and major issues of public policy, I must wonder why everyone does not recognize the necessity for every element of information which might have a bearing on the problem at hand. In this case, I see the needs and the benefits and have less of an appreciation of costs.

Another example is the budget. In the public sector at least the largest single outpouring of data is not to the higher education general
information survey but in support of or to support special studies of the institutions' annual or biennial budget request. To dispense with the provisions of this information would, of course, have extremely negative fiscal implications for the institution. A subjective decision is made that this information is worth it. The institution can see a direct cause and effect relationship between providing the information and receiving funding. The questions become far more hazy when this cause and effect relationship is not as clear. This is most often the case in terms of base line analyses dealing with either the central planning responsibilities within a state, the establishment of broad federal priorities or the response to particular legislative resolutions unless they strike at the roots of the very existence of the institution.

I submit that in the final analysis questions posed in this symposium cannot be answered without modifying the question with "Costs to Whom?" and "Benefits to Whom?".

From the standpoint of a state planning and/or coordinating agency, data are an absolute necessity to effectively accomplish its responsibilities. It must, however, address the questions I posed earlier in deciding the size, shape, and extent of data collection.

Let's review these questions in more detail. First, are the data necessary? This seems like an easy question to answer, however, in practice one of the most difficult. As hindsight is more accurate than foresight, so, too, is it easier to evaluate whether what you did was appropriate than what you are planning to do will be appropriate. One needs to evaluate the level of detail and the definition of the elements to determine whether the direct evident purposes and reasonably possible ancillary purposes incorporated in the study or project can be met. It is my opinion that this process
can best be accomplished through the give and take involved in discussing the subject with the individuals who are required to collect and provide the information and open one's preliminary judgment to the critique of other knowledgeable individuals.

Second, are the data reasonably available? This question must be considered within the context of the first. If the data are determined to be absolutely necessary to accomplish the study or project, the availability question is secondary as long as the project itself is well justified. If the answer does affect the decisions on timing or the possibility of sampling, however.

Third, are suitable proxies available in existing reports? This is a very important aspect which not only applies to states but the federal government, central boards, and institutions themselves. Oftentimes reasonably similar elements of information are provided in ongoing, routine reports. It is important before beginning a new data collection effort to thoroughly review existing reports to determine their applicability. It is not enough, however, to conclude that merely because the information is related to what is needed, it automatically meets the needs of the study or project. For example, financial data are provided in a variety of forms and formats. A careful analysis of the necessity question will provide the evaluator with the insight as to whether those reports are sufficient depending upon the purpose of the study. An ancillary benefit of this analysis is the potential for modifying ongoing reports so they do have greater applicability.

Fourth, can sampling be used? This is one of the most important questions since I believe there is a suspicion on many of our parts that statistical techniques are suspect, and that we would run the risk of overlooking
some significant item if we did anything less than to survey the entire universe. Sampling can and should be used to a greater extent, again, depending upon the study or project itself and the necessity for the information.

Fifth, should one survey or collect information on-site? I have a distinct bias in this area since I feel that there are many additional benefits to working directly with individuals in institutions through the site visitation process. A better understanding of the basic recordkeeping system results, better communication between the provider and the requester is often established and as a matter of fact, more is learned through this process than merely sending out a survey form and manipulating the data which are returned.

Sixth, what about the timing and frequency of data collection? Oftentimes this question cannot be answered since it is someone else's decision to determine whether information will be again requested. However, every effort should be made to estimate whether the needs will be recurring and at what time intervals. One of the most important yet often overlooked aspects of the data burden question is a lack of appreciation of the frequency of the data collection issue. If a basic recordkeeping system can be adjusted to produce information on a known time frame basis, the second collection is not nearly as burdensome as the first. It does, however, place a burden on the requester to thoroughly analyze needs to minimize modifications. In addition, when institutional recordkeeping systems are changed, adequate lead time needs to be made available. As a general rule, we try to allow institutions at least ten months lead time in making basic modifications to record systems to respond to changes in data elements or the shape and scope of ongoing data requests.

It is, of course, much easier to discuss these questions on a general basis than to fully adhere to all of the points I have just noted. I know
from personal and sometimes sad experience that I have deviated from my own
good advice. It is likely that we will always have a degree of conflict
between the provider and recipient of data in terms of its worth and its
benefits. This applies to relations between the federal government and the
states, the states with institutions and institutional administrations with
their component parts. Attempting to address each of the questions which I
have outlined above, in an open and candid manner, through involving the
various parties to the data flow should serve to significantly tilt the cost-
benefit ratio toward the benefit side.