Studying the Value of Library and Information Services. Part II. Methodology and Taxonomy

Tefko Saracevic and Paul. B. Kantor
School of Communication, Information and Library Studies, Rutgers University, 4 Huntington Street, New Brunswick, NJ 08903. E-mail: tefko@scils.rutgers.edu, kantor@scils.rutgers.edu

> These ambiguities, redundancies, and deficiencies recall those attributed by Dr. Franz Kuhn to a certain Chinese encyclopedia entitled *Celestial Emporium of Benevolent Knowledge.* On those remote pages, it is written that animals are divided into (a) those that belong to the emperor, (b) embalmed ones, (c) those that are trained, (d) suckling pigs, (e) mermaids, (f) fabulous ones, (g) stray dogs, (h) those that are included in this classification, (i) those that tremble as if they are mad, (j) innumerable ones, (k) those drawn with a very fine camel's hair brush, (l) others, (m) those that have just broken a flower vase, and (n) those that resemble flies from a distance.
> Jorge L. Borges (1966) *Other inquisitions 1937–1952.* p.108.

This report, presented in two parts, is derived from a large study sponsored by the Council on Library Resources. Two of its objectives were to develop a taxonomy of value-in-use of library and information services based on users assessments and to propose methods and instruments for similar studies of library and information services in general. In the first part of the report, we discussed the underlying concepts related to value, which must be clarified to proceed with any pragmatic study of value. This established a theoretical framework, i.e., a theory of use-oriented value of information and information services. The theory drives the models used for the pragmatic part of the study resulting in a *Derived Taxonomy of Value in Using Library and Information Services.* In this, the second part, we deal with the specifics of the study: importance of a taxonomy; the method used for gathering data on user assessments of value in five research libraries, involving 18 services and 528 interviews with users; development and presentation of the Taxonomy; and statistics and tests of the Taxonomy. A novel aspect is the division of value of information services into three general classes or facets: (i) *Reasons* for use of a service in the given instance; (ii) quality of *Interaction* (use) related to that service; and (iii) worth, benefits, or implications of subsequent *Results* from use. We conclude the paper with suggestions for applications in practice, particularly in evaluation of services, and general discussions regarding the principles of taxonomic development.

Received February 7, 1996; revised June 19, 1996; accepted June 19, 1996.

1. Introduction

1.1 Organization of the Report

We report on an empirical study of value of library and information services which resulted in a *Derived Taxonomy of Value in Using Library and Information Services* (later called Taxonomy). The term ''derived'' in the name of the Taxonomy reflects the fact that it is derived from what users had to say about the value of services received. We divided the reporting in two parts. In Part I, we addressed the following: importance of study of value in library and information services; the notion of ''value'' in philosophy and economics; value of information; value of information services; the connection between value and relevance; and related studies. The result was a framework, that is, a *theory of use-oriented value of information and information services*, establishing the rationale for proceeding with the pragmatic part of the study, reported in this part. We developed two use-oriented models: one relates to use of information, called the *Acquisition-Cognition-Application* or A-C-A model, and the other relates to use of information services, called the *Reasons-Interaction-Results* or R-I-R model. In this part, we will refer implicitly or explicitly to this theoretical framework and to the models in particular but will not repeat their description.

In this part, in the Introduction, we present the connection with Part I and the objectives of the study. In Section 2, we discuss importance of a taxonomy. In Section 3, we describe the methodology for collection of data. In section 4, we present the methods for analysis of data to derive the Taxonomy. In Section 5, we present the resulting Taxonomy and provide examples. Section 6 contains statistical results on the *Empirical Taxonomy* and

Section 7 tests of the *Derived Taxonomy*, including a novel theoretical base for such tests. Section 8 suggests practical applications elsewhere. Section 9 discusses the principles of the Taxonomy, followed by conclusions.

1.2 Objectives of the Study

This section appears in both parts of the report, in order to make them self-contained. This report is a result of a 15-month study sponsored in large part by the Council on Library Resources. The goal of the study was to address the problem of developing models and methods for studying the value and cost of library and information services in a way that can be pragmatically generalized and applied by services wishing to conduct similar studies. The practical aim of the study was to provide libraries and information systems, in general, and those oriented toward research, such as large research libraries, in particular, with methods for gathering information on value and cost of their services—information that will aid in justification and decision making. In the process, we also worked toward a theoretical framework for value of library and information services. The objectives of the study were to: (1) develop and test a taxonomy of value of library and information services based on user assessment; (2) determine costs associated with specific services; (3) develop methods for combining and correlating cost and value data; and (4) provide detailed descriptions and manuals that will allow for replication of these types of studies.

This report is one in a series of papers on the study. The report is devoted and restricted to reporting on the *first objective only, that is, the report deals with framework and description of value only.* Subsequent and related papers deal with costs and with correlation of costs and value (Abels, Kantor, & Saracevic, 1996; other articles are in preparation). A progress report on the study, which can be considered the predecessor of this paper, was presented at a conference (Kantor & Saracevic, 1995).

Achievement of these objectives required an empirical project and study involving collection and analysis of large amount of data from several research libraries and a number of different services, as described in Part II. Details of the project, (including rationale, methods, and results, plus appendices presenting instruments, procedures, and manuals for use in replication) are presented in the project's Final Report (Kantor, Saracevic, & D'Esposito-Wachtmann, 1995). A Manual for application of the Taxonomy is presented separately (Huttenlock, Dawson, Saracevic & Kantor, 1995).

2. Importance of Taxonomies

Why devote such effort to a taxonomy? What use are taxonomies? These are legitimate questions that need clarification.

Taxonomy is a form of classification; in philosophy and science, it also refers to the study of classification. The root comes from a Greek word meaning arrangement. Human minds are, by both nature and nurture, relentless classifiers. Classification is a basic process of our minds. It is essential for understanding ourselves and our surroundings, for association, and for other cognitive processes.

Classification is the oldest method in scholarship. Aristotle, for instance, classified everything in sight, in order to describe it in terms of attributes and to relate the derived classes, often, to moral and ethical values or propriety. Classification is also a basic method in science; scientific method implicitly or explicitly starts from some classification, and often results in classification. However, modern scientific and scholarly thought, starting in the 17th century, does not relate derived classes to moral or ethical values or propriety, as did the scholars in previous centuries, but it relates them to some inherent properties or characteristics—inherent in terms of what the scholar thought or observed to be of importance. As a result, some famous scientific classifications have emerged. Carolus Linnaeus (1707–1778) established the modern method of naming and classifying animals and plants; he used their observed structure as the base for the taxonomy used to this day. Gregor Mendel (1822–1884) formulated the basic laws of heredity after classifying seven pairs of traits in garden peas and following the changes in classified traits in successive generations. This served as a basis for modern study in genetics. Dmitri Mendeleev (1834–1907) introduced order in inorganic chemistry by classifying elements, using their atomic weight as a base. The classification served, among other, as a guide for discovery of unknown elements.

There are many similar examples in social sciences, although, on the whole, less dramatically successful. As was shown in Part I, Adam Smith's (1723–1790) classification of value into *value-in-exchange* and *value-in-use* served as a base for development of modern economic thought. Abraham Maslow (1908–1970) developed a classification of human needs, based on an ordered series of levels. Closer to home, Melvil Dewey (1851–1931) developed a pragmatic library classification scheme, based on an epistemic interpretation of the organization of knowledge, and so on.

All these classifications have one thing in common: they have successfully served as a foundation for further work. This is the strength of and even the reason for a number of classifications or taxonomies. In scholarship, many classifications have been used as a first step for development of a theory and/or as a guide for observation and explanation of some phenomena or processes. In pragmatic situations, classifications have been used for decision-making, as in Dewey's case, for arrangement of library materials. This highlights not only the strengths, but also the limits of classification or taxonomy in general.

It does not *explain,* it only *describes.* It is a first step toward application and/or explanation.

Taxonomies can be developed in many ways. They can be developed on the basis of preconceived notions and beliefs. They can be based on careful, selective observation and even measurement. In either case, certain principles have to be followed, as elucidated by Bailey (1994). The danger of proceeding without principles is illustrated by the delightful classification of animals quoted in the heading. It was not successful as Linnaeus's taxonomy of animals based on observation of structure.

Our Taxonomy does not *explain* the value of library and information services. But, we hope that the Taxonomy developed here has two features that go beyond the Taxonomy itself. First, we believe that it can be used in practice, that is, that it can be applied as a base for proceeding with a successful and pragmatic study of value of library and information services. Second, we also believe that it may be used in experimental and theoretical work, that is, in observations exploring effects of variables and in furthering of theory and explanations. We discuss possible applications in Section 8 and present the related application-based principles that guided us in construction of the Taxonomy in Section 9.

3. Data Collection

Approach. Our approach was to identify several large research libraries, select a few services for study at each, interview the users of these services, analyze the transcripts of the interviews, and on the basis of the analysis, develop and test taxonomies of value. In other words, the basis is observation and analysis of users′ assessments of value. The collection and analysis of data involved complex procedures, details of which are provided in the appendices of Kantor, Saracevic, and D' Esposito-Wachtman (1995). Thus, only a summary is given here.

Services Involved. Five research libraries (labeled L1–L5) participated. We asked each library to select the services to be studied. A total of 18 services were selected as shown in Table 1. Some of the services were on-site, and we intercepted users in person after the use, others were off-site (electronic network based) services, and we intercepted the users through an online invitation (placed as a screen) for a telephone interview.

Focus Groups. Prior to interviewing users we conducted two focus group sessions in order to: (i) test the applicability of concepts and questions incorporated in a preliminary version of our survey instrument (questionnaire) and (ii) obtain general views about perceptions of the value of a library for users, together with difficulties they have in assigning value. The sessions were moderated by a trained moderator, following written instructions. Focus group interviews were conducted at L1 with 11 faculty members, each representing a different discipline, and at L2 with 13 graduate students representing the three disciplines corresponding to the selected services. Discussions were taped and transcribed. We analyzed the transcripts and modified the questionnaire as appropriate. The focus groups were also useful for confirmation and modification of the study framework and model.

Questionnaire. The questions followed the R-I-R model. In particular, we framed the interview to focus on the specific task or project in relation to which a user approached a library service. Thus, although a general questionnaire was used, the study sought to isolate the context of the present use event as a ''critical incident.'' The questionnaire contained two types of questions: (i) Likert-type scales providing a user with a choice of a number on a scale from 1 to 7 to assess a given aspect of the service, and (ii) open ended questions addressing:

1. Reasons why a user came to the library or accessed the given library service at this time.
2. What the user got out of the use; what benefits did he/she receive.
3. What would a user do otherwise, if the service were not available at that library.
4. Elaboration on why did the user give a particular score on the Likert scales.

A master questionnaire was developed and tested in pilot applications and then modified as needed for each service so that the questions incorporate the specific name and form of the service. Consequently, for the 18 services there were 17 versions; for Services 2 and 3 (Reference Service, but in different branches) we used the same version of the questionnaire. The questionnaires are reproduced in *Appendix A,* and the interviewer training manual (including instructions for interviewing) in *Appendix E* of Kantor, Saracevic, and D'Esposito-Wachtmann (1995).

Interviews. We recruited, paid, and trained interviewers to intercept the users in person or by telephone and elicit and tape-record the answers to questions. We selected interviewing rather than asking users to fill out questionnaires by themselves for two reasons. First, we could ensure a more complete response to open-ended questions and thus obtain more data for analysis. Second, and even more important, we knew from our pilot study and focus group interviews that users have a hard time in maintaining a focus on questions about value even for short periods of time. We trained the interviewers to pursue the focus on the question at hand.

It may be of interest to report on the cost of interviewing and transcription: At the rate of $10/hr for interviewers, $0.80 for keypunching numeric data per interview, average of $8 for transcription per interview, the *average cost of labor* (excluding travel expenses but including administrative and training costs) for interviewing, keypunching, and transcribing were $20.20 per

TABLE 1. Services studied.

Service no.	Library	Service name	Description
1	L1	Information Desk	At library entrance. Addresses ''simple'' reference & other questions
2	L1	Reference Desk, Branch 1	Reference service by librarians
3	L1	Reference Desk, Branch 2	Reference service by librarians
4	L1	Reference Consultation Service	In-depth reference search
5	L1	Materials Delivery Service	Campus-wide delivery from all libraries in the University
6	L2	Art and Archeology Library	Use of non-circulating collection
7	L2	Biology Library—Electronic Services	Online & CD-ROM services to a number of databases & catalogs
8	L2	Psychology Library—Reference	Reference service by librarians
9	L3	Enhanced Online Catalog Service	Online catalogs & databases
10	L3	Undergraduate Reserve Service	Library & faculty materials
11	L3	Document Delivery Service	Provided at the Biology Library
12	L3	CARL Uncover	Faculty accounts for direct order
13	L4	Electronic Reference Resources	Electronic sources in diff. subjects
14	L4	Music and Media Center	Recordings, tapes, films, video . . .
15	L4	Interlibrary Loan Service	Personal and mail delivery
16	L5	Automated Reference Service	Online & CD-ROM databases
17	L5	Traditional Reference Service	Reference service by librarians
18	L5	Patents Service	Patent depository & service

Number of libraries, 5; number of services 18.

interview in Wave 1 and $14.13 per interview in Wave 2, where there were fewer open-ended questions.

Waves. To accommodate for possible differences in assessment of value by users as the time passes by (and presumably as they had a chance to use the information for the task at hand), we intercepted users at two different time periods. In Wave 1, in order to gather immediate assessments of value, we interviewed users as soon as they have finished using the service. In Wave 2, we explored the longer term assessments of value, contacting the users at the time of use and then by phone about 2 weeks after they used the service. In terms of phases in A-C-A model, Wave 1 may be thought as more associated with *Acquisition* and *Interaction* phase, while Wave 2 adds the *Application* phase. Wave 1 was conducted in Fall 1993 and Wave 2 in Spring 1994, involving the same services but different users.

Users. The users interviewed were graduate students, faculty, or professionals—undergraduates were not the target population—with the exception of Service No. 10. (Undergraduate Reserve Service) where faculty and undergraduates were included. To obtain representative responses from each service, we established beforehand quotas of users for each service; we met about 70% of the quota. Because of our screening criteria we had to intercept in person or by phone more users for each actual interview. About one in four of intercepted users resulted in an interview. We obtained altogether 534 interviews—388 (73%) in Wave 1 and 146 (27%) in Wave 2. Of those 534 interviews, 528 resulted in transcripts that were usable for analysis. Thus, the Taxonomy reported here is based on 528 cases. Table 2 presents some characteristics of the 528 users.

4. Analysis and Development of Taxonomies

Approach. The analysis of transcribed interviews formed the basis for development of two Taxonomies of value. The analysis was conducted in two separate rounds. In the first round, we focused on discovering the terminology and vocabulary of users, as lifted from the interviews. This resulted in an *Empirical Taxonomy* (ET), as presented in *Appendix B* of Kantor, Saracevic, and D'Esposito-Wachtman (1995). The *Empirical Taxonomy* was organized according to specific questions which elicited the answer. In the second round, we focused on generalization from the *Empirical Taxonomy* and on logical arrangements and testing following the R-I-R model. The result was the *Derived Taxonomy of Value in Using Library and Information Services* (DT) presented in this part. Wherever we refer to ''Taxonomy'' we mean the *Derived Taxonomy*.

Our development of the *Empirical Taxonomy* can be said to reflect the *grounded theory approach,* which suggests that models be built from the bottom up, using raw data (Corbin & Strauss, 1990). Our specific method was, of course, *content analysis of text,* which permits classification of large quantities of text—in this case using texts of interviews with users (Weber, 1990).

Finally, in development of both *Empirical* and *Derived Taxonomy,* we have considered the *principles of taxonomies in social sciences,* as outlined by Bailey (1994).

TABLE 2. Characteristics of users.

From Libraries		Status		Gender	
L1	96	Faculty	92	Male	258
L2	115	Graduate student	334	Female	270
L3	110	Undergraduate	91		
L4	95	Other	11		
L5	112				
Area of study		**Age**		**Frequency of library use**	
Humanities	142	Under 18	3	Less than monthly	74
Professions	142	18–25	172	Once a month	49
Social sciences	96	26–29	119	2–3 times a month	97
Natural science	83	Thirties	120	4 times a month	39
Engineering	14	Forties	73	More than 4 times	127
Other	40	Fifties	26	Missing	142
Missing	11	60+	12		
		Missing	3		

Number of users: 528.

Empirical Taxonomy (ET). Starting with the specific questions and the actual vocabulary of users' answers, categories in ET were defined, refined, and redefined in a series of steps. These steps were reiterative, involving considerable amounts of feedback. First, one analyst went through all interviews and identified representative keywords and phrases. Second, the analyst grouped these into categories associated with specific questions in the interviews and gave each category a label. Third, these were grouped into a working classification scheme to be used for testing. Fourth, the analyst went back to the interviews and classified each response according to that working scheme. Fifth, the analyst wrote then a set of instructions for encoding to test the reliability of the scheme when the interviews were encoded by other analysts. Sixth, two analysts, not previously engaged, tested the scheme for intercoder agreement; the consistency was acceptable (kappa $\sim$ 65%). (Further statistics on the *Empirical Taxonomy* are given in Section 6.) As mentioned, in this report, we present only the *Derived Taxonomy* as the final product. We regard the *Empirical Taxonomy* as an intermediate product.

While developing the *Empirical Taxonomy,* we discovered a number of interesting aspects about the behavior of users' in describing their perceptions of value. We found that despite the intention of specific questions, the users in effect tend to speak about the entire experience in most responses. They have difficulty in assigning value or even relating to value. As a result, for instance, one of the categories in the *Empirical Taxonomy* with the working title ''What did you get from using the service?'' seems to span answers to many differing questions facing users in assigning value.

Derived Taxonomy (DT). Development and testing of DT was a second activity, also involving a number of reiterative steps. First, another analyst took the *Empirical Taxonomy* and on the basis of the theoretical framework and R-I-R model (described in Part I), constructed the first version of the *Derived Taxonomy.* Second, that version was then tested by the same analyst going back to the data, i.e., the interviews. Third, results were used for corrections. Fourth, two new analysts were engaged to test for intercoder agreement (consistency) in application of the Taxonomy using the original interviews, to make suggestions for refinements, and to prepare a manual for encoding (Huttenlock, Dawson, Saracevic, & Kantor, 1995). The result is the *Derived Taxonomy* presented next.

5. Derived Taxonomy of Value in Using Library and Information Services

5.1 Structure and Attributes

Structure. The *Derived Taxonomy* can be thought of as a faceted classification, incorporating levels as subdivisions. Following the R-I-R model, there are three major classes or facets: *Reasons, Interactions, and Results,* each with subclasses and specific categories or nodes. The structure of levels is as follows (where X. represent letters and n numbers in the codes):

X. General classes (facets)
  X.n Subclasses
    X.n.n Subclasses
      X.n.n.n Specific categories

As we have developed it, the Taxonomy has three major classes, 12 subclasses, three of which have further subclasses—there are 10 such sub-subclasses—and 90 specific categories.

We have structured the Taxonomy in an outline form. Graphically, the outline can be represented as a tree: from the trunk of value, branch three different major branches

of categories or facets, from each facet branch smaller branches or subcategories, and then leaves as specific categories or nodes. We have sought to arrange the distinct components of value on this tree in such a way that proximity in the tree (as measured by the number of steps required to get from one aspect of value to another) corresponds in a reasonable way to the relation between the ideas represented by the corresponding points in the tree. For example, if users speak of a sense of accomplishment or of a sense of confidence or of a sense of frustration, these three are all clearly different and yet they are somehow closely related to each other in that they represent the affective state of the user. Thus, they are placed together under a single subclass corresponding to affective outcomes, (in this case, subclass C.2 of Figure 2) and are rather more distant in the tree from concepts such as saving time through using this service (C.5 in Figure 2), which are places in different branches. It is to be noted that the principle of mutual exclusivity is applied to *concepts* and not to the *labels* attached to individual nodes or leafs of the tree. That is, we might find the phrase ''term project'' appearing both in the *Reasons* class or branch of the tree and in the *Results* branch of the tree. Although the words are the same, the full concept is represented by the entire path from the node to the branch, and these are different.

Relations. The relations within and between different levels of the taxonomy (classes, subclasses, and categories) are based on set-member relationship, as in faceted classifications. A class X has subclasses which are its members; subclasses are sets with other subclasses or specific categories as members. A specific category ''*Research*'' belongs, with a number of other categories, to a subclass ''*Task*''; subclass ''*Task*'' belongs, with a number of other subclasses, to a class ''*Reasons.*'' Because there are three top-level facets or classes (*Reasons, Interactions, Results*) and within each further sub-facets or sub-classes, this is a polyhierarchical structure. Specific categories denote or label a user's response and connote or imply the class to which it belongs. That is, selected user responses are grouped into specific categories (lowest level) that denote (label) those responses. In turn, the categories are grouped in subclasses that connote (imply) a common attribute shared by the set of specific categories. Finally, subclasses are grouped in classes that connote an attribute shared by the set of subclasses.

Nonuniqueness. A response may include several conceptually distinct parts. Thus, more than one specific category may be assigned to a given user response. These categories can come from any of the three general facets.

Open Endedness. The Taxonomy was derived from a sample of users. While the sample was large and representative of a number of library services, there could be other users and services which call for additional specific categories of value. By necessity, any user-derived taxonomy of value is not complete; it is also not of uniform depth. Additions can be anticipated and included, particularly on the level of specific categories. Furthermore, if and when this Taxonomy is applied to a context different than ours, some specific categories may be modified or, if not relevant to that context, may be omitted.

Stability. We expect, however, that in different applications and contexts: (i) the three general classes will remain stable; (ii) the subclasses will be relatively stable; and (iii) the greatest need for change will be at the level of the specific categories, or the ''leaves'' of the tree. In other words, as this Taxonomy is adapted the changes will almost certainly appear in specific categories to fit specific contexts and applications in different libraries and information services. For instance, modifications in specific categories can be expected if this Taxonomy is applied to special libraries and information centers in given organizations, to account for the given context, users, and use.

5.2 Contents of the Derived Taxonomy

We present The *Derived Taxonomy of Value in Using Library and Information Services* in two exhibits. Figure 1 shows the overall structure at glance. We list the general classes and subclasses, without the specific categories under each. At this level we believe that the Taxonomy will remain stable for most, if not all, library and information services.

Figure 2 shows the whole Taxonomy in all its detail, i.e., we incorporate general classes, subclasses, and specific categories. As mentioned, we expect that changes/modifications may occur on the level of specific categories when the Taxonomy is applied to a specific library or information service within a given context (organization, institution).

5.3 Description of Major Classes and Subclasses

In this section, we concentrate on discussion and illustration of the three major classes or facets (Class A, *Reasons;* Class B, *Interaction;* and Class C, *Results)* and their associated subclasses as listed in Figure 1.

Class A: Reasons. This class covers the causes, motives, bases, purposes, and/or rationales underlying the use of library services. Why do users use the library and information service? What do users want to get from the service? We have subdivided *Reasons* into three subclasses:

A.1 *Task or project:* Activities, work, or problems in which the user is engaged that are the direct cause for using the library service. What are the users doing that prompted the use of the service? What were the users working on or wanting to work on? What problems brought them to the service? Included are tasks such as: research, compiling a bibliography, class assignment, project proposal, and a host of others. Most of the times

A. **REASONS** for using a library or information service
- A.1 For a **TASK** or project
- A.2 For **PERSONAL** reasons
  - A.2.1 Cognitive reasons
  - A.2.2 Affective reasons
  - A.2.3 Reasons for substitute choice
- A.3 To get an **OBJECT, INFORMATION** or perform an **ACTIVITY**
  - A.3.1 Physical (tangible) objects
  - A.3.2 Intangible objects
  - A.3.3 Perform an activity or work

B. **INTERACTION** with a library service
- B.1 **RESOURCES** - availability, accessibility
- B.2 **USE** of resources, services
- B.3 **OPERATIONS and ENVIRONMENT**
  - B.3.1 Policies, procedures
  - B.3.2 Facilities, organization
  - B.3.3 Staff performance
  - B.3.4 Equipment performance

C. **RESULTS** of using a library services
- C.1 **COGNITIVE** results
- C.2 **AFFECTIVE** results
- C.3 **ACCOMPLISHMENTS** in relation to tasks
- C.4 **EXPECTATIONS** met
- C.5 **TIME** aspects
- C.6 **MONEY** estimates

FIG. 1. General classes or facets.

there is some tangible element in *Tasks,* an element that is directly reflected in the responses.

A.2 *Personal reasons:* Private, individual reasons for using the library services. Most of these are intangible. What is the motivation of the users for using a library service, as far as knowledge or emotion is concerned? The *Personal reasons* are further subdivided into subclasses:

A.2.1 *Cognitive reasons,* related to various aspects of learning and knowledge, staying current, or orienting oneself;

A.2.2 *Affective reasons,* related to a person's own emotional feelings or desires for using the services, like relaxation, stress reduction, feeling of satisfaction;

A.2.3 *Reasons for substitute choice,* related to a personal decision to use the given library or information service instead of some other choice.

A.3 *To get an object, information or perform an activity:* Covered are reasons or objectives associated with what the users intend to request from a service or what activities they intend to engage in at the library. What do the users actually want to get from the service? What will they do in the library? This subclass is further subdivided into:

A.3.1 *Physical (tangible) objects:* Reasons or objectives related to getting a book, article, or other objects available from a library service;

A.3.2 *Intangible objects:* Reasons or objectives related to getting information, facts, data; being directed to other sources of information;

A.3.3 *Perform an activity or work:* Studying, searching, browsing and other activities a user intends to perform. Includes nonlibrary activities, such as the use of a computer, if available for general use.

Class B. Interaction. This class covers the assessment by users of the qualities of various aspects of library services. How does a user assess or evaluate the encounters with the library while seeking and using a service? We have divided *Interaction* into four subclasses:

B.1 *Availability and accessibility of resources, services:* Covers availability, accessibility, and quality of given materials, items or services. Availability asks: Does the library have a given resource, item or service? Does it have a given book or journal and is it available for use on this occasion? Accessibility asks: How readily can the access to the service be obtained? Quality asks: How current, timely, complete is a given resource or service?

B.2 *Use of resources, services:* Covers a number of aspects connected with user experiences while using a resource or service. How convenient is it to use the service? What effort is required or how easy is it to use? What is the degree of frustration in using it? How much

DERIVED TAXONOMY OF VALUE IN USING LIBRARY AND INFORMATION SERVICES

A. **REASONS** for using a library or information service

A.1 For a **TASK**

- A.1.1 Research
- 1.2 Dissertation/thesis
- 1.3 Project work
- 1.4 Professional and other occupational work
- 1.5 Paper, report, article - writing, starting
- 1.6 Book - writing, starting
- 1.7 Bibliography, references, citations, sources - compiling, checking
- 1.8 Class assignment, requirement for grade or degree
- 1.9 Exam, test, comprehensive
- 1.10 Teaching, instruction - preparation, gathering materials
- 1.11 Presentation, oral report
- 1.12 Proposal for grant, funding
- 1.13 Job search, job application, interview for job, employment
- 1.14 Review, assessment, appraisal, evaluation of a book, proposal, application and other objects, materials
- 1.15 Planning for some activity, work
- 1.16 Delegated work - doing it for or helping somebody else

A.2 For **PERSONAL** reasons

A.2.1 Cognitive reasons

- A.2.1.1 Learning something, confirming something
- 2.1.2 Staying current, catching up with an area, topic
- 2.1.3 Orienting oneself to the library, resources, services, equipment; learning how to use them

A.2.2 Affective reasons

- A.2.2.1 Relaxing, pleasure, recreation, leisure, curiosity
- 2.2.2 Reducing stress, worry

A.2.3 Reasons for substitute choice

- A.2.3.1 Using this library service instead of other choices - other information resources, services people

A. **REASONS** for using a library or information service **(cont.)**

A.3 To get an **OBJECT, INFORMATION** or perform an **ACTIVITY**

A.3.1 Physical (tangible) objects

- A.3.1.1 Getting a book, periodical, article, recording, image, film, video
- 3.1.2 Using interlibrary loan or materials delivery service to get an item
- 3.1.3 Using recall to get an item

A.3.2 Intangible objects

- A.3.2.1 Obtaining information, facts, data; clarify something
- 3.2.2 Pointing to another source(s) of information within or outside the library

A.3.3 Perform an activity or work

- A.3.3.1 Studying, reading in the library
- 3.3.2 Viewing a film listening to a recording, using special equipment
- 3.3.3 Searching electronic resources - catalogs, databases
- 3.3.4 Searching print resources - catalogs, indexes, tables of contents
- 3.3.5 Browsing
- 3.3.6 Copying
- 3.3.7 Use computers for non-library task (where available as service)
- 3.3.8 Performing other non-library or non-information activities or work

FIG. 2. General classes and specific categories.

B. INTERACTION with a library service

B.1 RESOURCES, SERVICES

- B.1.1 Availability of desired materials, item(s) - degree of
- 1.2 Completeness of given resource, service - degree of
- 1.3 Currency, timeliness - degree of
- 1.4 Accessibility, ability to use a given resource, service - degree of

B.2 USE of resources, services

- B.2.1 Convenience in using the resource or service - degree of
- 2.2 Effort required in using it; ease of use - degree of
- 2.3 Frustration in using it - degree of
- 2.4 User performance - degree of perceived ability
- 2.5 Effort in getting from one resource or service to a complementary or subsequent one - degree of

B.3 OPERATIONS AND ENVIRONMENT

B.3.1 Policies, procedures

- B.3.1.1 Clear - degree of
- 3.1.2 Conducive for ease, convenience, effectiveness of access, use - degree of
- 3.1.3 Requirements upon users, fairness, reasonableness - degree of

B.3.2 Facilities, organization

- B.3.2.1 Space - degree of adequacy
- 3.2.2 Physical layout, design, and organization - degree of quality
- 3.2.3 Adequacy, quality of intellectual organization of resources, materials, services - degree of
- 3.2.4 Comfort, ambience of facilities - degree of quality

B.3.3 Staff performance

- B.3.3.1 Knowledgeability, expertise - degree of
- 3.3.2 Helpfulness, empathy, sensitivity - degree of
- 3.3.3 Efficiency - degree of

B.3.4 Equipment performance

- B.3.4.1 Technical functioning - degree of
- 3.4.2 Availability and clearness of instructions, guides, documentation
- 3.4.3 User friendliness, ease of use - degree of
- 3.4.4 Difficulty in operating equipment - degree of

FIG. 2. continued.

effort is required to get from one service to another (e.g., from searching for and finding references to getting articles)?

B.3 *Operations and environment:* Covers the experience of a user in relation to the working performance and environment of the service. Four subclasses are included:

B.3.1. *Policies, procedures:* How clear, convenient, reasonable are they?

B.3.2 *Facilities, organization:* How adequate or of what quality is the space, physical layout, intellectual organization? How comfortable are the facilities?

B.3.3 *Staff performance:* How knowledgeable, helpful, efficient is the staff?

B.3.4 *Equipment performance:* Technical functioning; instructions; difficulty; user friendliness?

Class C. Results. This class covers users' assessment of outcomes. What did a user get out of the service? What did a user accomplish? Were the expectations met? How is the service related to time and money? Six subclasses are included:

C.1. *Cognitive results:* What was learned? Getting ideas? Reinforcement? Serendipity?

C.2 *Affective results:* Degree of satisfaction? Failure? Confidence? Comfort? Frustration?

C.3. *Accomplishment in relation to task(s):* Degree of contribution? Fulfilling assignment? Next step?

C.4. *Meeting expectations:* Getting, obtaining what needed? Too much, too little? Uncertainty?

C.5. *Time aspects:* Saving time? Wasting? Speed of service? Sufficiency of allocated time?

C.6. *Money estimates:* Estimate of dollar value obtained. Saved? Spent? Lost if service not available?

5.4 Specific Categories: Encoding Decisions, Examples

In this section, we provide examples of encoding decisions to demonstrate the approach to and problems of

C. **RESULTS** of using a library services

C.1 **COGNITIVE** results
- C.1.1 Learning something, stretching knowledge
- 1.2 Reinforcing knowledge
- 1.3 Changing viewpoint, outlook, perspective
- 1.4 Getting ideas, perspective, conceptualization how to proceed
- 1.5 Serendipity - getting ideas about different, tangential things
- 1.6 Getting no new ideas; did not learn anything

C.2 **AFFECTIVE** results
- C.2.1 Sense of accomplishment, satisfaction, success - degree of
- 2.2 Sense of failure - degree of
- 2.3 Sense of confidence, reliability, trust - degree of
- 2.4 Sense of comfort, good feeling, happiness - degree of
- 2.5 Sense of frustration, stress - degree of

C.3 **ACCOMPLISHMENTS** in relation to task(s)
- C.3.1 Contribution to accomplishing or proceeding with task at hand; facilitation of or help with work; - degree of
- 3.2 Fulfilling assignment - degree of
- 3.3 Providing access to people or other sources of information
- 3.4 Providing for a next step, task, information seeking activity - degree of

C.4 Meeting **EXPECTATIONS**
- C.4.1 Getting, obtaining what needed, sought, expected, requested - degree of
- 4.2 Getting too much
- 4.3 Getting nothing
- 4.4 Confidence in sources or certain about what gotten - degree of
- 4.5 Exceeding expectations, getting additions to what expected
- 4.6 If not gotten what expected, degree of hurt; seeking of substitute action

C.5 **TIME** aspects
- C.5.1 Saving time as a result of using the service - amount, comparison
- 5.2 Wasting, losing time in using the service - amount, comparison
- 5.3 Waiting time to access the service - amount
- 5.4 Speed, quickness of service - comparison
- 5.5 Time it takes to figure out or use the service - amount
- 5.6 Time available or allocated for use of the service - degree of sufficiency

C.6 **MONEY** estimates
- C.6.1 Estimate of a dollar value of results obtained from a given service, or of information obtained
- 6.2 Estimate of the amount of money saved because of use of the service
- 6.3 Estimate of the cost (dollar value spent) in using the service, or the actual amount spent
- 6.4 Estimate of what may be spent on a substitute service or activity for similar results
- 6.5 Estimate of dollar value lost in cases where service was not available or the use was not successful

FIG. 2. continued.

assigning specific categories to the texts of user responses. We also provide a sample of user responses as illustration.

Within the class *Reasons,* the specific labels in subclass *Tasks* are not exclusive or easily separated from those in subclass *Personal reasons.* There is, however, a differentiation between the concepts represented. Most of the time a distinction can be made by thinking of *Tasks* as being tangible and contextual while *Personal reasons* are intangible and represent the inner world of the user. On the other hand, labels in subclasses of the class *Interactions* (*Resources, Use and Operations*) may be simultaneously applied to a given user response, as may those under class *Results.* More than one specific label under any subclass, or combination of subclasses, may be needed to classify user answers. We found that even a brief user's response may reveal several different but overlapping aspects, each

of which naturally falls in its own place in the Taxonomy. In our test of the Taxonomy, we made an arbitrary decision to assign from one to three specific labels (codes) for a response. An alternative procedure, which is to split responses and assign one code for each, was not explored.

In probing for why the users use the library, we found that they often responded in terms of what they want to get (''to get a book''), or what they would do in the library (''do a search''), rather than in terms of the underlying reasons which drove them to get an object or do an activity, such as ''class assignment'' or ''paper writing.'' One has to probe to get to these underlying reasons— the users think of both the activity in the library and the underlying reasons as one, even though they are conceptually distinct. Similar examples can be given for a number of other questions. Users tend to bundle responses associated with *Reasons* as they do with those with *Interactions* and *Results.* This is the main basis for a lack of exclusiveness, as discussed above.

In particular, the transitive verb ''to get'', and its various tenses and synonyms, appears often in users' responses, but the ''get'' responses span two classes of the Taxonomy: Specific categories related to ''get'' appear under *Reasons* and under *Results,* but with significant distinction. Under *Reasons,* and its subclass *To get an object, information or perform an activity,* there is an intention: ''Why did you use the library today?'' ''To get a book.'' Under *Results,* and its subclass *Accomplishments in relation to task(s),* there is an outcome: ''What did you get out of using the library today?'' ''I got a book that helped me with [tasks at hand].'' Although in this case the answers to both questions are similar, they are distinct; the first refers to a reason for using the service and an intended action, while the second one refers to a result and an accomplished action. Often, they follow each other, but at times ''To get a book'' is not followed by ''I got a book.''

Figure 3 provides a sample of users' responses categorized in the three main classes. Further examples for every specific class in the *Derived Taxonomy* are provided in the Manual for encoding (Huttenlock et al., 1995).

5.5 Examples of User Responses on Possible Substitutions

After the users actually used a service, one of the questions asked was hypothetical in nature:
''*If the* [*given*] *service were not available at this library, is there something else you could have done to get the same benefit?*'' If the answer was YES, then the follow up open-ended question was: ''*What would you have done?*''

We did an analysis and classification of answers to this open-ended question. We did not include the classification of the possible substitutions in the Taxonomy because it does not reflect the *use of an information service.* It reflects only a *hypothetical substitution* for a service given that it were *not* available for use. Thus, although we asked the question and analyzed the results, we left this class out of the Taxonomy. Still, we believe that the results may be of interest to illustrate users' perceptions and provide interesting examples. Figure 4 contains a classification of users' answers about hypothetical substitutions for given service were it not available, together with examples for each specific category. Nine specific categories are provided, ranging from ''Use another service in the library'' to ''Do nothing.''

6. *Empirical Taxonomy:* Distribution of Responses

As mentioned, the *Empirical Taxonomy* was developed first, based on extraction and classification of users' responses to given questions. It is a question-driven classification following the users' vocabulary. Categories were developed and a code was given to each category. Then, an analyst went back to the interviews to encode each text that was a response to a question, following the codes in the Empirical Taxonomy. The frequency distribution and other statistics of responses were analyzed using the multiple response capability of the SPSS statistical package.

There were 528 interviews with users (cases) containing 2,132 user responses to questions in the interviews (an average of 4.04 responses per interview). Up to three codes were allowed for each response. The single coder assigned 3,497 codes to user responses (an average of 6.47 codes per interview and 1.64 codes per response). Out of these, we took the responses which were coded according to *why did a user use* the service and *what did a user get out* of the service for a presentation here, because they are the most interesting and illuminating. (Distribution of responses to other questions coded in the *Empirical Taxonomy* can be found in Appendix C of Kantor et al., 1995). Table 3 presents the predominant ''why'' and ''get out'' responses (those that appeared in more than 10% of the 528 interviews), analyzed by their content and distribution. The first column contains the text for a category from the *Empirical Taxonomy* (e.g., ''get physical object'') ; the corresponding numerical code from the *Derived Taxonomy* (Figure 2) is shown in brackets. The second and third columns provide the corresponding number of responses and percentage of all (3,497) responses, e.g., ''get physical object'' had 459 responses which was 13.1% of the 3,497 responses. The last column is significant in that it shows the percentage of all the interviews in which the particular response is given. For example, ''get physical object'' was cited in 86.9% of the 528 interviews, ''obtaining information or knowledge'' in 77.1% , and ''for a paper'' in 72.5% of interviews. Since more than one code may be assigned to each response, the last column sums to more than 100%.

In interpreting the results from Table 3, it is very gratifying to see that the leading categories in user responses correspond to our conceptual framework and facets later

Notes: Letters and numbers in brackets refer to codes from Derived Taxonomy shown in Exhibit 2. The quotes are independent; i.e. responses from one class are not connected with responses in any other class.

Class A. Reasons	Class B. Interaction	Class C. Results
I wanted to do several searches for papers that I need to do for classes this semester. [A.1.5]	Well, the fact that I could access it at home ... you can do it at your leisure [B.2.1]	It made me realize that I shouldn't go into this project at that stage. [C.1.3]
It is my second time here and I have no clue where anything is. [A.2.1.3]	I couldn't figure out how to use the machine. [B.3.4.4]	It worked! I mean I found the book. It was on the system. [C.2.1]
I am doing a research project and it beats looking up in paper form. [A.1.1; A.3.3.3]	It was easy for me to just say "I need these two newspapers..." and I got them. [B.2.2; C.4.1]	...Invest a lot of time down there to find the articles and use the machine. [C.5.5]
I come here to work. [A.1.4]	I got what I needed with minimal hassle. [B.1.5]	It didn't have any kind of impact. [C.1.6]
I was having trouble in locating a book to start my research on a certain paper that I am starting. [A.3.1.1; A.1.1; A.1.5]	I waited two hours and was unable to have access to viewing the tape ... they did not have enough machines. [B.1.4]	By doing the search ourselves, we found that the name was already being used and that saved us from making a costly error. (from using a patent service) [C.6.2]
Various questions have come up in our patient care here, and people did enough reading or research to find out that there were articles on these topics and I sent the information to Ms [...] and she sent us back the articles, so it was to further our education on patient care. [A.1.1; A.2.1.1; A.3.1.1]	All I got is one book really. The person behind the desk told me exactly where it was. The fact that the library doesn't have enough resources on my subject is really not her fault. [B.1.2; B.3.3.2; C.4.1]	There are so many sources to use inside the [electronic services] ... it made me look into other things that I wasn't even thinking of using. [C.1.5]

FIG. 3. Examples of users' responses (direct quotes) in relation to specific categories in the Derived Taxonomy.

used in the *Derived Taxonomy.* Getting a physical object was cited in 86.9% of the cases. Obtaining information or knowledge was cited in 77.1% of the cases, and ''for a paper'' was cited in 72.5% of the cases. Looking at the top three reasons reported we see here clearly the extreme importance of the *Acquisition, Cognition,* and *Application* (as described in the A-C-A model in Part I) in users' thinking and classification. That is ''get physical object'' may be associated with *Acquisition,* ''get information or knowledge'' with *Cognition,* and ''for a paper'' with *Application.* Interestingly, however, thinking about *Application* from the outset may play a much higher role in people's thinking because the next several categories, ''task completion'', ''for research'', and ''for a class,'' all clearly relate to the application aspect.

Exchange values appear with ''saving time'' in 31.8% of cases, and just below it ''convenience and efficiency'', which refer implicitly to time as a medium of exchange. The next two categories were used to code negative experiences, neither occurring very frequently. However, 20% of the responses included some statement that information or resources were not available and 22.5% indicated that in some way the service did not fulfill the immediate goals. Last in the group of responses occurring more than 10% of the time is the simple statement that the experience was positive because the people at the library were nice. As an aside, we note that simply behaving nicely will not distort users' perception of the quality of service. Only 10% of the cases cite this quality of interaction, after an overwhelmingly more frequent appearance of other aspects and qualities.

Our objective in this study was *not* to evaluate the services involved; it was solely to develop a taxonomy and provide suggestions for applying it and tools for evaluating it. However, the preceding discussion can also serve as an example of what type of statistical and evaluative results one may obtain, when using the specific categories from the *Derived Taxonomy* as questions in a survey. Clearly a variety of correlations may be tested for significance.

Category number	Description of substitute action	Examples of users quotes
S.1	Use another service in the library	Ask somebody else, maybe at the circulation desk Wander through the stacks. Use another electronic service Go to books
S.2	Go to another library	I would go to the Library of Congress Go to some other library
S.3	Use other facility/resource outside of libraries	Go to Brussels Go to the main office in Virginia. Would have written a letter. Call the Academy of Medicine. Connect my computer with a network. Use other facilities at the University.
S.4	Purchase the book, material, etc.	Go to some kind of [subject] bookstore. Go back to a store for newspapers. Would have bought a TV and VCR
S.5	Subscribe to a service, journal, . . .	Pay for Compuserve. Subscribe to the database service on my own. Go to the nearest video store and rented a video.
S.6	Consult with a person outside of the library	Talk to my professor after Spring break. Would have gotten professional help, like going to a lawyer. I would talk to people working in this area. Make phone calls and directly ask people "Do you know of some..."
S.7	Do more work on one's own	I would have to dig deep. I might have gone back and looked at my research and found other questions to ask. Spent more time in the computer trying to do it for myself. Had to piece it together myself
S.8	No substitute available	I couldn't have done it ... without having access to the databases to prove that my strategies are valid. I don't know what I would do.
S.9	Do nothing	Go to the park and watch people.

FIG. 4. Examples of users' responses (direct quotes) in relation to question on ''possible substitutions for given library services that were used'' (Class S).

7. *Derived Taxonomy:* Statistics, Tests

To ensure that there is a clear operationalization of the classification represented by the Taxonomy, we have used multiple coders in the development and testing of the Taxonomy, and have applied a variety of statistical techniques to ascertain the significance of the agreement among the coders. As mentioned, we permitted up to three distinct codes to be assigned to a single user response. This has required the development of some techniques for analysis which we believe to be new. (A detailed manual for coding is available in Huttenlock et al., 1995).

In the test of degree of agreement (intercoder consistency), we used two coders to apply independently codes (letter and number combination representing specific categories) from the *Derived Taxonomy* to user responses to questions in 528 interviews. Standard statistical techniques exist for tests of intercoder consistency when two coders assign a *single code* for a response (e.g., kappa). However, there are no standard techniques for testing intercoder consistency when *more than one code* is assigned to a response. As we had allowed up to three codes to be assigned per response, we lacked a technique to test consistency, so we had to develop one. In itself, this poses some interesting methodological problems, because the coders may agree on some of the codes for the same response but not on others. To deal with this we have developed a statistical theory of consistency on a code-by-code basis.

We also propose that a polyhierarchical or tree-like structure of the specific codes, as in our Taxonomy, lets us define a path between two different coder's assignments for the same response and, thus, derive a path-based distance between two codes. This enables us to assess the degree of agreement between two coders with regard to a response as a whole. The theory and methods that follow from the theory are new. We believe that they will be useful to anyone seeking to modify or apply our Taxonomy and then test intercoding consistency in assignment of codes, or to anyone testing any taxonomy with a tree-like structure and multiple coding.

TABLE 3. Empirical Taxonomy (ET): Predominant encoded responses to questions on why did a user use the service and what did a user get out of the service.

ET category label & code from Derived Taxonomy	No. of codes in ET	Percent of codes	Percent of cases
Get physical object [A.3.1]	459	13.1	86.9
Get information, knowledge [A.3.2]	407	11.6	77.1
For a paper [A.1.5]	383	11.0	72.5
Task completed (degree of) [C.3.1]	212	6.1	40.2
For research [A.1.1]	192	5.5	36.4
For a class [A.1.8]	174	5.0	33.0
Save time [C.4.3]	168	4.8	31.8
Convenience, efficiency [B.2.1]	146	4.2	27.7
Didn't fulfill goal [C.4.3]	119	3.4	22.5
Not available [B.1.1]	110	3.1	20.8
Positive (people were nice) [B.3.2]	53	1.5	10.0

Number of users (cases), 528; number of user responses, 2,132; number of codes to user responses, 3,497.

In this section, we present application of those two theories to our results, that is to the codes independently assigned by two coders to 2,132 responses in 528 cases. Note that results described above for the *Empirical Taxonomy* were based on assignments by the single coder who developed the taxonomy, while here we used two coders to test consistency. Again, we allowed up to three codes per response. The fist coder later called Coder A assigned 4,617 codes and the second coder, Coder B, assigned 3,770 codes to the 2,132 responses.

7.1 Code-Based Consistency: Theory.

When there are two coders each of whom selects codes from some available set or alphabet of codes and assigns them to responses, we can ask whether their agreement is significant or not. This is somewhat different from simply asking whether the consistency is high enough for us to be comfortable about using the codes, assigned in this way, as defining properties of the responses themselves. In the latter context, intercoder consistencies of 65% or higher are generally accepted. The Cohen kappa measure, for example, considers the consistency of the two coders across the entire set of codes. (Bateman & Gottman, 1986). Because of the tree-like structure of our Taxonomy, and the availability of multiple codes for each response, this approach can not be adapted directly.

However, we are accustomed, in dealing with other phenomena, to assess their significance by considering an alternative null hypothesis and asking whether the observations could have come about under the null hypothesis. We believe that it is useful to apply this line of reasoning to the present situation. To do so, we must formulate a null hypothesis. We begin by considering what factors might prevent coders from agreeing on the code that they assign to a particular response. There are three ways this could happen:

1. The underlying taxonomy might not make sense, as in the Celestial Emporium.
2. The taxonomy might not be relevant, appropriate or adequate to the responses.
3. The coders might have been inadequately trained.

If any one of these three conditions is met, we are in the ''null hypothesis'' condition. Below we will characterize that condition mathematically. But first we note that logically, rejection of the null hypothesis involves rejection of the ''OR''-ed combination of all three possible causes of disagreement. And the contradiction of an OR is the logical AND of the contradiction of all three conditions. That is, rejection of the null hypothesis entails confirmation (although not, of course, proof) of the assertion: the taxonomy makes sense AND it is relevant to the responses AND the coders are adequately trained. Assuming well-trained coders, we call the first two parts of this the *Taxonomic Hypothesis.* In other words,:

Taxonomic Hypothesis: The taxonomy makes sense AND the taxonomy is relevant to these responses.

The mathematical statement of the null hypothesis for the purpose of statistical tests is:

Null hypothesis: Each of two coders assigns each of the possible codes in a way that has no correlation with the other coder.

This represents the notion that the assignment has nothing to do with the text or response to which it is assigned. For, if each coder pays some attention to the text, there would be some induced correlation in their assignments of the codes. The model can be realized mathematically by being a little more specific. We focus our attention on any specific code. Let R be the total number of responses coded, let A be the number of times that the first coder assigned that code, and let B be the number of times that the second coder assigned the same code. The chance $\Pr(R, A, B, k)$ that they will agree exactly k times, under the null hypothesis, is given by a ratio of certain binomial coefficients:

$$\Pr(R, A, B, k) = \frac{\binom{A}{k}\binom{R-A}{B-k}}{\binom{R}{B}}.$$

In this expression, the denominator represents all of the ways that the second coder can assign this particular code to exactly B of the R responses. The numerator represents the number of ways that the second coder can assign this code to exactly k of the A cases to which the first coder assigned it and to exactly B-k of the other R-

A cases. The binomial coefficient $\binom{n}{k}$ represents the number of ways to choose exactly k out of n items. It is computed from the factorial expression $m! = 1 \times 2 \times 3 \times \cdots \times n$ according to the rule:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{1\cdot 2\cdots k}.$$

This looks a little complicated but is easy to work out with the aid of modern spreadsheets such as Excel, for which the corresponding probability functions are built in, under the name Hyper Geometric Distribution. When a particular level (k) of agreement is observed, the confidence is calculated by finding the probability of the observed level of agreement, or higher. That is, we need to compute:

$$Q(R, A, B, c) = \sum_{k=c}^{k=\min(A,B)} \Pr(R, A, B, k).$$

This is the total probability, under the null hypothesis, of seeing the observed number of agreements (c), or any larger number of agreements, up to the maximum possible, which is of course the smaller number of assignments of that code. We find, in fact, that this number (Q) is usually quite small. For example, with 600 responses to be coded, if both coders use a particular code 20 times each, the expected number of agreements under the null hypothesis is only $20 \times 20/600 =$ two-thirds of an agreement. Thus we find that even quite small numbers of agreements, such as 5 out of the 20 possible, have a high statistical significance. Technically, the cumulated sum needed here is not a built-in function in a spreadsheet, but is implemented in terms of a look-up table which is recomputed for each of the cases.

To give an example which can be verified by the reader. Limitations of space prevent us from using a real example, in which 2,132 responses were coded. Suppose instead that there were 50 responses and that a particular code was assigned to 10 of them by one coder (A), to 4 of them by another coder (B), and that they agreed in $k = 3$ cases. The specific probabilities of this many agreements or more is the sum of only two terms: three agreements and four agreements. (Since the second coder only assigned this code 4 times, there cannot be more than 4 agreements.) Thus we must compute the sum of two fractions:

$$\frac{\binom{10}{3}\binom{40}{1}}{\binom{50}{4}} + \frac{\binom{10}{4}\binom{40}{0}}{\binom{50}{4}} = \frac{\frac{10 \times 9 \times 8}{1 \times 2 \times 3} \times \frac{40}{1}}{\frac{50 \times 49 \times 48 \times 47}{1 \times 2 \times 3 \times 4}} + \frac{\frac{10 \times 9 \times 8 \times 7}{1 \times 2 \times 3 \times 4}}{\frac{50 \times 49 \times 48 \times 47}{1 \times 2 \times 3 \times 4}} = 0.02175.$$

The corresponding calculation, executed in our spreadsheet, yields:

N	A	B	Agree	Chance better	Adjustd agree
50	10	4	3	2.175%	0.6875

It is easy to see that this is statistically significant, but the measure of reliability, adjusted for this code, is $(3 - x)/(4 - x)$ where x is the expected number of agreements $x = (10/50) \times 4 = 0.8$. So the adjusted agreement reliability measure is $(3 - 0.8)/(4 - 0.8) = 2.2/3.2 = 0.6875$. This means that, regarded as a measuring instrument, this code assignment has a reliability of less than 70%. But regarded as a test of the validity of the Taxonomic Hypothesis it has a significance of 100%-2.175%, which is better than 95%, the usual standard for confirmatory hypothesis testing. Thus the agreement on this code validates the corresponding Taxonomic Hypothesis, even though the code in this hypothetical case is not a very reliable description. We return to this issue below.

7.2 Code-Based Consistency: Empirical Results

Using the principles described above, we have calculated the chance of the observed agreement or better, under the null hypothesis, for every one of the codes used. Results for those codes used at least one hundred times are shown in Table 4. The chance of better agreement under the null hypothesis is astonishingly low, giving support to the Taxonomic Hypothesis at a correspondingly high level. This confirms both components of the Taxonomic Hypothesis: the *Derived Taxonomy* makes sense AND it is relevant to user responses. Recall that $P = 0.05$ is the usual level for confirmatory testing, and any lower value would be judged significant. On the other hand, the adjusted agreement (column 7) is often not even 50%. This means that when one of the coders has assigned the code to a case, even a low chance that the other coder will assign the same code is an occurrence much better than random chance under the null hypothesis. (Note that Coder A assigned a total of 4,617 codes to the responses; Coder B assigned a total of 3,770 codes to the set of responses.)

For example, in Table 4, all codes which were used at least 100 by one of the coders are shown. We show the number of times the code was used by the first coder and by the second coder, and the number of times that they agreed. This ranges to values as low as 0.22 (for Code C.3.1) for which the corresponding number in column 6

TABLE 4. All codes which were used at least 100 times by one of the coders.

Code	Maximum Uses	Coder A	Coder B	Agree	*P*	Adjusted Agreem.
C.4.1	323	323	318	132	~ 0	0.31066
A.3.1.1	290	286	290	164	~ 0	0.50627
C.3.1	266	266	186	59	~ 0	0.21987
A.3.2.1	256	256	252	131	~ 0	0.45432
C.3.2	153	153	58	23	9.9×10^{-13}	0.3499
A.1.5	141	131	141	108	2×10^{-127}	0.81199
C.4.4	138	115	138	57	1.4×10^{-41}	0.46075
A.1.1	138	138	125	96	1×10^{-105}	0.75194
C.5.1	130	120	130	85	3.5×10^{-89}	0.68939
C.1.4	116	88	116	48	1.9×10^{-41}	0.5193
C.3.3	114	114	44	21	2.3×10^{-16}	0.44774
A.1.8	109	109	95	79	1×10^{-101}	0.8225
B.1.2	104	70	104	23	1.7×10^{-14}	0.29414
B.4.2	101	101	94	77	5×10^{-102}	0.81016

We show the number of times the code was used by the coder A (N = 4,617) and by the coder B (N = 3,770), and the number of times that they agree. Column 6 shows the probability of the observed level of agreement or better, under the null hypothesis. This is the usual confidence level. Column 7 shows the adjusted agreement, as defined in the text.

is too small to be computed. This is, by strict statistical measures, a high level of confirmation of the Taxonomic Hypothesis. This suggests that a rule of thumb such as ''70% adjusted agreement'', based on reliability considerations, is not an appropriate scientific representation of the significance of the observed agreement. For example, code C.3.3 was used 114 times by Coder A and 44 times by Coder B. They both assigned it to the same item in 21 cases. The chance that this would occur under the null hypothesis is only 2.3×10^{-16}. The adjusted agreement is 0.4474.

Note that we are treating each of the separate codes as if its confidence level could be assessed independent of the others. Strictly, in looking at scores of codes, we expect to find some that show good agreement merely by chance. However, even such ''data dredging'' would not produce spurious confidence levels as high as shown here.

This is a new approach to the assessment of validity of coding, whose implications remain to be explored. It is clear that we are far from the usual realm of consistency levels. On the other hand, the chance that we see the observed levels of agreement is, under the null hypothesis, astronomically small (if one may use such an expression). So the Taxonomic Hypothesis is supported. In fact, the very high significance levels observed protect us against the usual fears about testing many hypotheses at once. Even with the most conservative multiple tests models, levels of $P = 10^{-12}$ remain significant. We may conjecture that such high levels of statistical significance occur together with such low levels of conventional concordance because there are many ideas inherent in each of the responses we have collected. In a sense, even allowing three distinct codes may simply be inadequate to capture the full set of ideas expressed. This will force lower levels of concordance as different coders make different choices among the entire set of codes that might reasonably be assigned to a given response. Hence, significance is a better test than reliability for determining the validity of the taxonomic structure.

In the following section we discuss how discordant code assignments might be merged to arrive at consensus judgments and some measure of the certainty of that consensus judgment.

7.3 Intercode Distances: Theory

When we have two different coders assigning codes to the same responses, and each coder assigns three codes from a tree-like structure (as shown in Figure 5), we can define the distance between their assignments. This is a somewhat subtle two-step process (Kantor, 1995).

First, we can certainly define the distance between any two responses from our tree structure. To do this we proceed in three steps. First we find the lowest common code of the tree, which contains both codes. For example, A.2.3 and A.2.4 are both contained in the common class represented by the code A.2. A.2.3.3 and A.3.2.3 have as their lowest common code A itself, and so on. We introduce a formal ''root'' of the tree represented by '' *, '' which is above all three parts of the taxonomy. Thus A.2.3 and B.4.3 have '' * '' as their lowest common code.

Second, we count up the number of steps from each code to the common code. For example, A.2.3 is one step below A.2, and A.2.4 is one step below A.2. Finally we add the number of steps to determine the difference between the two codes. So A.2.3 is 2 steps (one up the outline, and one down) from A.2.4. On the other hand, A.2.4 is only one step from A.2 itself. This formalizes the definition of the distance between codes. Note that this definition makes sense because no order is implied among several classes appearing under the same higher class in the Taxonomy.

Now we must define the distance between triplets of codes, since each coder is allowed to assign up to three

codes. This is done in the ''most lenient possible manner''. The issue is this: if one coder assigned some codes, e.g., A.2, B.1.3, and C.4.1 and the other assigned the same codes in a different order, we would want to discover that they made the same assignments. (Recall that we did not assign any special significance to the order of codes assigned.)

To do this, we compute the pairwise distance between codes assigned by the first coder and codes assigned by the second coder for all possible pairings of the codes. So if one coder assigned codes x, y, z and the other assigned codes w, x, y, we consider six different pairings, such as (x with w) + (y with x) + (z with y). That is not a particularly intelligent way to pair them, as it maximizes disagreement. But one of the pairings is (x with x) + (y with y) + (w with z). This is the best possible pairing in this case and yields the lowest possible summed distance between codes. In general then, *we define the distance between the set of codes assigned by one coder and the set of codes assigned by the other as the minimum value that can be found by considering all possible ways of pairing them up.* If one coder assigned more codes than the other, the excess codes are ignored in computing the distance. So, for example, the assignments (x, y, z) and (x, y) are considered to have 0 distance between them.

This kind of measure is a natural one given the tree-like structure of our Taxonomy. One may think of it as respecting the relation between a heading and a subheading but not giving any particular significance to the order in which subheadings appear. That is, A.2.4 is not closer to A.2.5 than it is to A.2.7. This reflects our logic in building the structure to begin with. We do not believe that dimensions of the value can be arrayed in a simple linear order any more than we believe that, for example, the books in a library can be arrayed in a simple linear order which captures, for all readers, the relations among subjects.

In the present research, this concept served as a supporting tool in the refinements of the Taxonomy. We wanted to be able to identify those cases in which the coders have a high level of agreement (distance is zero) and those in which they had a high level of disagreement (high distance). Particularly during the development of coding rules and the structure of the Taxonomy this was a tool to help zero in on the responses that were most problematic, as they were the ones for which the distance between the two assigned sets of codes was the largest. Thus, in the manual for encoding (Huttenlock et al., 1995), we provided appropriate and more detailed explanations based on the findings about distances. This can serve as a general example on how to test our Taxonomy (or any other tree-like taxonomy) in practice, in order to identify codes that are less or more problematic, and subject them to additional analyses.

7.4. Intercoder Distances: Empirical Results

The distribution of distances between the scores, in the best possible alignment is shown in Table 5. For instance, in 284 or 13.3% of 2,132 encoded responses the distance between two coders was zero, i.e., there was maximum possible agreement, and in four cases, the distance was 19—the maximum achieved disagreement (or minimum achieved agreement) between those two coders.

TABLE 5. The distribution of intercoder distances.

Intercoder distance	No. of responses
0	284
1	7
2	56
3	15
4	376
5	185
6	163
7	112
8	247
9	147
10	130
11	104
12	108
13	86
14	57
15	24
16	11
17	9
18	7
19	4
Total	2132

Number of responses encoded, 2,132; number of coders, 2. Each distance is computed after the codes assigned to user responses, by the two coders, have been paired so as to give the minimum possible distance between them A distance of 0 means there was no disagreement. If one coder assigned the same codes as another, plus an additional code, the distance is still counted as 0.

The data in Table 5 show that there was perfect agreement in about 13% of the cases, in a tri-modal distribution with the most prominent intercoder distances being zero ($N = 284$ responses), four ($N = 376$), and eight ($N = 247$).

This measure of distance, or consistency can also be used to assign codes, with degrees of certainty, in the case of general polyhierarchical coding. The idea is to define the ''lowest common code'' of two different codes. This is simply the highest point in the tree of codes, through which we must pass to move from one code to the other. Examples of the coding results for two responses is shown below. First example is of a question and answer that resulted in a maximum agreement between two coders, and the second that resulted in a poor agreement.

Case 103. *Question 6:* Why did you use the Information Desk today?

Answer: Because I had to find a journal on [ ] and had no idea where it would be and I needed help.

Subquestion: OK. Can you tell me a little more about it?

Answer: Um. Well, I am using the [ ] system to find journals, so I found the journals but I had no idea if we

even have 'em anymore or what the story was, so I went to desk to ask for help.

Subquestion: OK. And can you tell me the context of what you are going to use the journal for?

Answer: I had to write a paper for Computers and Education.

Codes assigned by both coders: A.3.11; A.2.1.3; A.1.5. *Lowest common code:* Those three codes. *Distance:* 0

Case 312. *Question 3a:* What did you get out of using the [Online Service] on that occasion?

Answer: I found that the available information, my source information did not match with the database. For example, it had a volume number and a date and when I put it in . . . of course the title . . . there were 20–30 of the same date, but no volume numbers. So now I have to go and browse all the shelves. So it was fruitless. Well, you know it was not all that satisfactory. And I can tell you why. Because I feel that the data entry that is going on to log these books is quite haphazard at [ ]. Are you familiar with what goes on with that? . . . for example, between 10%–50% of the books are not even cataloged.

Codes assigned: Coder 1: A.3.3.5; B.3.3.3; B.3.2.3. *Coder 2:* C.1.1; B.2.3 *Lowest common code:* *, B. (* = root). *Distance:* 17

In this way, the cases with lower or higher distances in encoding can be identified and subjected to further analysis. In the first case, all three codes can be considered to have the highest confidence, as there was complete agreement between the codes. In the second case, the existence of a lowest common code (B) is virtually meaningless because the overall distance between the codes assigned.

8. Possible Applications and Replication

One of the objectives of the project, as stated in Section 1.2, was to propose and develop methods for possible applications of the Taxonomy and replication of this study. Consequently, an important part of the project was aimed at providing elaborate tools for reproduction and conduct of similar studies, as presented in great detail in Kantor, Saracevic and D'Esposito-Wachtmann (1995) and in the manual for encoding (Huttenlock et al., 1995). Here, we will not elaborate on these tools, but we wish to make some suggestions on how to apply the *Derived Taxonomy* and the associated framework in a practical situation or to replicate the work (or part thereof) in a similar study.

In a study that is directed toward getting users' assessment of value for a given library or information service, the Taxonomy can be used in several ways.

Questionnaires. The Taxonomy can serve as the basis for the structure and content of a survey instrument to be applied in a study of value; in particular, it can guide development of questions and Likert-type scales in the instrument. Or, modifications to the Taxonomy can be made beforehand (particularly at the level of specific categories) to accommodate the context of the service. A given library or information center can study a number (if not even all) of its services using the same structure for the survey instrument, but modify the wording of questions to accommodate specific contexts, as did we. In this way uniform data can be obtained across a number of services.

Encoding. In cases where open-ended responses are obtained from users, the Taxonomy can be used for encoding of responses. As mentioned, we have developed a manual for encoding just for that purpose (Huttenlock et al., 1995). That manual contains rules and examples for encoding in each category; it can be used or modified for application in other studies. The Taxonomy together with the manual can save considerable time that would have to be spent in development of encoding structures and rules.

Evaluation. We have *not* set out to evaluate any of the services studied. However, it is clear that the Taxonomy (or selected parts of interest) can be used in user-centered evaluation of services. *Use in evaluation may be the most important application of the Taxonomy.* To give an example of evaluative results that could be obtained: in Table 3, some of the encoded responses pertain directly to evaluation. For instance, in 168 responses (used in 31.8% of cases) users indicated that they ''saved time;'' in 119 responses (used in 20.8% of cases) users indicated that they ''didn't fulfill goal.'' Clearly, one can get an evaluative picture by follow-up, focused on cases and reasons for these (or any other evaluative) response. Likert-type scales can be constructed for aspects of evaluative interest, and open-ended interviews conducted after a user indicates a numerical score for reasons of that score. In this way, a numerical score for a service and the rationale behind the score can be obtained as evaluation. On the basis of results, decisions can be made for modifications, if any, in the service. Results of such evaluations can also be used for benchmarking and continuous improvement, the hallmarks of the total quality movement (TQM).

Replication, Extension. In scholarly studies of value, the Taxonomy, together with the theoretical framework, models, and testing tools, can serve as the basis for similar studies elsewhere, to compare, contrast, and illuminate results. In our tests, we barely touched the distribution of responses as a variable. As this was not in our objectives, we have not dealt at all with possible effects and behavior of variables that can be readily identified from the Taxonomy. However, the Taxonomy can be used for such objectives. Replications and extensions of this study are needed if we are to learn more about value related to library and information services. As in other areas, development of

a taxonomy is a prelude for more in-depth studies. But it can focus such studies and provide for accumulation and organization of findings and knowledge.

A Caveat. We wish to offer a note of caution. Any study of value is a complex undertaking. It requires surprisingly high levels of resources, time, and commitment. It also requires the usual concomitants of research: considerable preparation and close attention to often boring details. In addition, the analysis is not likely, in this early stage of the development of the field, to follow simple and predictable paths. Therefore, institutions and investigators contemplating a study of value should be fully aware of these factors. In addition, we found that, as a rule, users cannot or do not sustain a focus on value when interviewed about their assessments of the value of a service. The focus is not maintained even for a short period of time. In practice this means that an interviewer should be trained and prepared to refocus the discussion. During the analysis, a number of variables enter simultaneously into responses by users, and the theoretical challenge is to disentangle these variables into a focusable list.

9. Principles Used in Development of the Taxonomy

While development of a ''grounded'' taxonomy is, in principle, free to discover ''whatever it wishes,'' we have been guided by a number of principles which have been mentioned at places above. We bring them together here and illustrate how they have helped us.

The first of these principles recognizes that a taxonomy ought to speak effectively to four overlapping audiences: library professionals; institutional administrators, library users; and other researchers. This goes beyond the usual requirements of taxonomies as elucidated by Bailey (1994) as well as beyond the general requirement that research simply be consistent with, or an extension of, previous research. Because we are engaged in the study of self-aware organizations, which may wish to adopt or adapt our methods, we must be aware of their constraints and goals. To put it more simply, a taxonomy of library value cannot be persuasive if librarians or users themselves find it at odds with their own experience.

To ensure that a taxonomy be meaningful to *library and information service professionals,* we must include those aspects of value which can be influenced by these professionals. A professional, *qua* professional is interested in research only in so far as it leads to improved practice, and one cannot improve that over which one has no control. In general, these aspects correspond to the class of *Interaction* in our Taxonomy.

To ensure that a taxonomy be meaningful to *institutional administrators,* we must reflect those aspects of value which bear on institutional responsibilities and goals. From the point of view of administrators, the library information service as a whole has only a contributory value. In general, those institutional concerns are reflected in the facets of our Taxonomy dealing with the *Reasons* for using the library information service, and *Results* obtained, as reflected by users.

To ensure that the Taxonomy be meaningful to *users,* we have based its development, substantially, on what users themselves say when talking about the value of library information services. This has led to the *Empirical Taxonomy,* which served as a foundation for the *Derived Taxonomy,* and the Taxonomical Hypotheses presented and tested in the present paper. We note also that it is the users alone who are capable of speaking to the *Reasons* for use and to the *Cognition* aspects of the A-C-A model of the value of use of information itself.

Finally, normal research practice requires that we relate our Taxonomy to other research and *researchers* or, in other words, that we relate the conceptual structure of the Taxonomy to other models for understanding information systems. Our primary concern has been the relation to philosophical and economic models, as detailed in Part I of this paper, which served as the base of development of our models, and thus of the Taxonomy itself.

Turning now to the classificatory principles, we have structured the Taxonomy in an outline form. Graphically, the outline can be represented as a tree (Figure 5). We have sought to arrange the distinct components of value on this tree in such a way that proximity in the tree (as measured by the number of steps required to get from one aspect of value to another) corresponds in a reasonable way to the relation between the ideas represented by the corresponding points in the tree. For example, if users speak of a sense of accomplishment or of a sense of confidence or of a sense of frustration, these three are all clearly different and yet they are somehow closely related to each other in that they represent the affective state of the user.

Although our point of view has been to approach library and information services from the perspective of *value-in-use,* we find reference to the *value in exchange,* such as money or time. In more quantitative analyses, to be reported elsewhere, we have found time, rather than money, was the most effective dimension for assessing value. As have others, we found that users have great difficulty in assigning monetary value to library and information services. But they refer easily and often to time (''saved me time'', ''loss of time'' . . .). Various aspects of time may be the strongest indicator of exchange value for library and information services.

10. Conclusions

We have undertaken a study of value-in-use of library and information services. We present the results in two parts, the first one dealing with framework and the second with results. In the Conclusions of Part I, we summarized major concepts from the framework. In Sections 8 and 9 in this part, we provided specific conclusions dealing with

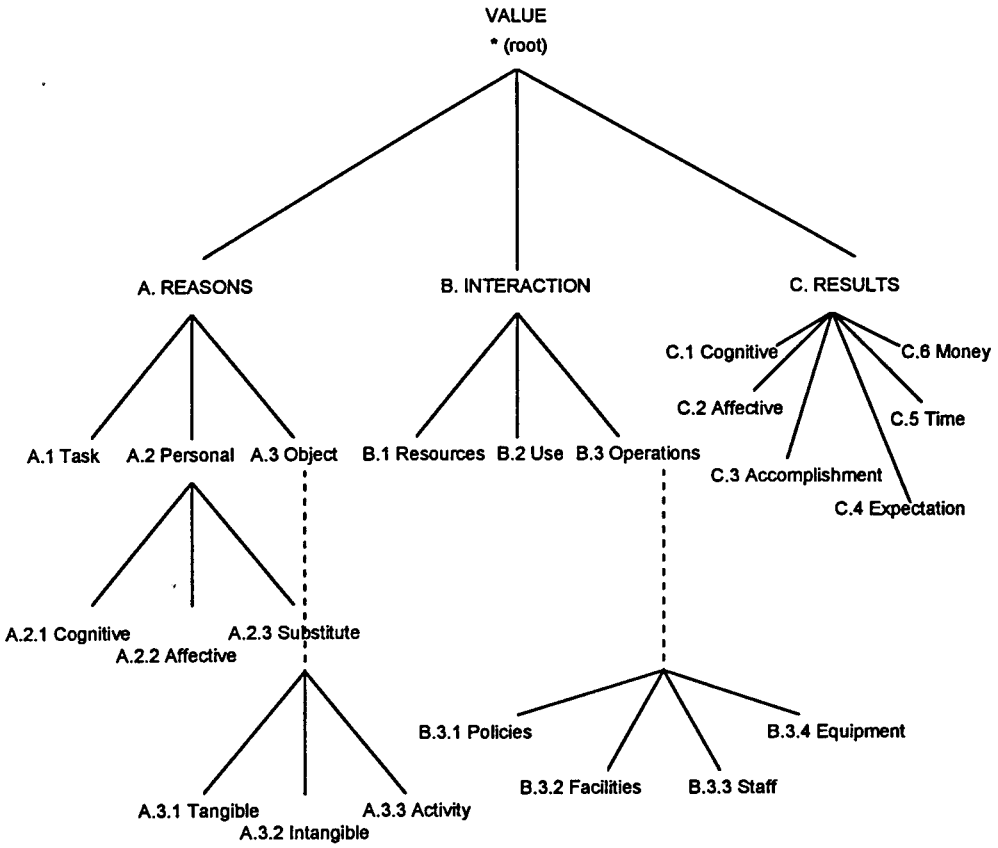

FIG. 5. Tree-like structure of the *Derived Taxonomy* (specific categories NOT included).

applications and general principles. Here we provide a few general remarks.

Our framework and model for study of use of information services calls for studying individuals using a service as related to a given task or problem. Using this framework, we interviewed over 500 users of 18 services in five large research libraries.

On the basis of these interviews we developed first an *Empirical* and then a *Derived Taxonomy of Value in Using Library and Information Services.* The *Derived Taxonomy* has a faceted structure, incorporating three major classes, each with subclasses, and then specific categories.

We believe that the Taxonomy developed and tested here, and the associated methodology are generalizable. That is, they can be used and modified for the study of value and for evaluation in a great many other library and information services. Over the years, there has been much debate about value of such services. At the same time, there has been relatively little progress in our understanding from a theoretical and methodological point of view. Too few data-based studies were conducted. Future progress in this area will require more pragmatic and theoretical studies. We believe that the present model of development and validation of a complex taxonomy will further the conduct of such studies.

While studies and determination of value are a difficult and involved proposition, they are only the first step in meeting a larger challenge. The challenge is to connect studies of value with some appropriate economic indicators. Given the present transitory circumstances and pressures, the fundamental question for library and information services is not whether they should undertake the difficult task in assessing their value, but whether they can afford not to.

Acknowledgments

First and foremost our thanks go to the users who were willing to participate in the study. Without them there would be no study. But, as is the custom, they remain anonymous.

This study became a reality because of the hard work, ideas, and cooperation of many colleagues. Joann D'Esposito-Wachtmann served as Project Manager. She also conducted the focus group interviews and played a large role in the definition, testing, and application of the successive instruments. Michael Wilk developed the *Empirical Taxonomy*. Dong Li was responsible for analyses, and Alan Zhao developed programs to support the analysis of transcribed texts. Terry Huttenlock and Pat Dawson coded the open-ended responses according to the *Derived Taxonomy* and developed the coding manual, and Kwong Bor Ng developed the analytical programs to asses the coding.

The project would not have been possible without the

generous cooperation of the five participating libraries. Our thanks to Elaine Sloan and Carol Mandel at Columbia University; Frank Polach, Marianne Gaunt, Margie Epple and Emily Fabiano at Rutgers University; Carlton Rochell, Nancy Kranich, Lucinda Covert-Vail, Nancy Friedland and Marlayna Gates at New York University; Donald Koepp, Nancy Klath, Mary Chaikin, Patty Gaspari-Bridges, David Goodman, and Janice Powell at Princeton University; Joanne Harrar and Danuta Nitecki at University of Maryland, and Eileen Abels at the College of Library and Information Services, University of Maryland.

Financial support was provided by the Council on Library Resources (CLR) through a grant to the Alexandria Project Laboratory (APLab) at the School of Communication, Information and Library Studies (SCILS), Rutgers University. Additional support was provided by Dean Richard Budd of SCILS.

References

Abels, E. G., Kantor, P. B., & Saracevic, T. (1996). Studying the value and cost of library and information services: Applying functional cost analysis to the library in transition. *Journal of the American Society for Information Science, 47*(3), 217–227.

Bailey, K. D. (1994). *Typologies and taxonomies. An introduction to classification techniques.* Thousand Oaks, CA: Sage Publications.

Bateman, R., & Gottman, J. M. (1986). *Observing interaction: An introduction to sequential analysis*. Cambridge: Cambridge University Press.

Corbin, J., & Strauss, S. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology 13*(1), 3–21.

Dervin, B., & Nilan, M. S. (1986). Information needs and uses. *Annual Review of Information Science and Technology, 21,* 3–33.

Huttenlock, L. T., Dawson, P., Saracevic, T., & Kantor, P. B. (1995). *Derived taxonomy of value in using library and information services: A manual for encoding of responses.* New Brunswick, NJ: School of Communication, Information and Library Studies, Rutgers University. Technical report APLab/95–5. Available in electronic form from: ftp://scils.rutgers.edu/pub/APLab/Cost.Value.Study/taxonman

Kantor, P. B. (1995) Of time and the library: New approaches to assessing the impact of the library on individuals and institutions. *Proceedings of the Conference Economie d l'information.* May 18–20 1995 (pp. 307–330). Lyon-Villeurbaine, France: Centre des Etudes et de recherches en science de l'information.

Kantor, P. B., & Saracevic, T. (1995). Studying the value of library and information services: A taxonomy of users assessments. *Proceedings of the American Society for Information Science, 32,* 35–44.

Kantor, P. B., Saracevic, T., & D'Esposito-Wachtmann, J. (1995). *Studying the cost and value of library services: Final report.* New Brunswick, NJ: School of Communication, Information and Library Studies, Rutgers University. Technical report APLab/94–3. Available in electronic form from: ftp://scils.rutgers.edu/pub/APLAB/Cost.Value.Study

Weber, R. P. (1990). *Basic content analysis.* (2nd ed.) Newbury Park, CA: Sage Publications.