Libraries nationwide are in yet another phase of belt tightening. Without an understanding of the economic factors that influence library operations, however, controlling costs and performing cost-benefit analyses on services is difficult. This paper describes a project to develop a cost model for collaborative virtual reference services. This cost model is a systematic description of all expenses incurred by a library in providing virtual reference service as part of a collaborative.

Introduction

Recent years have seen an upswing in interest in return-on-investment (ROI) and cost-benefit studies for libraries—and for good reason; federal and state governmental bodies show an increasing reluctance to provide funding for library operations. Perhaps the best-known recent cases of threats to libraries from the governmental bodies that fund them were the proposed closure in 2005 of the three public library branches in the city of Salinas, California, and the proposed closure in 2003 of the State Library of Florida. These two proposed closures were averted, largely due to public outcry. What is worse, however, is that these proposed closures are merely the tip of the iceberg. The ALA's Federal Libraries Wiki, wikis.al.org/fedlib/, tracks "federal library threats, reorganizations, and closings" over the past two years; and this wiki has tracked many closures and downsizings of the libraries within federal American Library Association agencies and military operations. Similarly, a slate of public libraries in states across the United States have closed their doors or significantly reduced their hours and personnel in the past two years due to budgetary reasons. The "Library Funding Report to the ALA Executive Board" notes that, in 2007, at least $15.5 million in library funding and budget cuts have occurred in the United States and that an additional $1.6 million in cuts are pending. Only some impoverished school libraries have seen increases in funding in recent years, thanks to the Laura Bush 21st Century Librarian Program.

It appears that libraries nationwide are in yet another phase of belt tightening. Libraries of all types can benefit, therefore, from a greater awareness of how to control their own expenses so as to make the most of decreasing or flat funding allotments. To do so requires focus on the economic aspects of library operations; but, unfortunately, very little work has been done on the economics of individual service elements within a library. Without an understanding of the economic factors that influence the individual components of library operations, controlling overall costs is difficult; and making comparisons between service areas on the basis of how much they cost in relation to how much benefit they provide is impossible.
This paper describes a project to develop a cost model for collaborative virtual reference services. Virtual reference is defined here as synchronous online reference service, using a commercial application, such as Questionpoint or Tutor.com. (Reference services that use instant messaging [IM] applications such as AOL Instant Messenger are also often referred to as virtual reference. Due to the functional limitations of IM applications, however, such services are rarely, if ever, collaborative.) A collaborative service is one in which multiple libraries participate and share the load of questions submitted by users to the service. A cost model is a systematic description of all expenses incurred by an organization while engaging in a specific activity—in this case, by a library in providing virtual reference service as part of a collaborative.

This project is the first stage in a two-stage effort to, first, develop cost models both for individual libraries participating in a collaborative virtual reference service and, second, for the collaborative as a whole. At the individual library level, it is important to identify where libraries participating in a statewide collaborative may find it useful to "tap into" shared resources and expertise and where participating libraries may find it useful to "go it alone." To make this assessment, however, the library needs a clear understanding of exactly what costs it is incurring. Thus, the scope of this stage in the project is to develop a cost model at the individual library level in order to help the library clarify its costs since many of the costs of participating in a collaborative virtual reference service are borne by the libraries themselves. In fact, as Eileen Abels points out, it is difficult to differentiate individual reference services provided by a single library because most resources are allocated in complicated ways across numerous services.

Literature Review

Recent library ROI studies have tended to focus upon the overall benefits provided by libraries to their users. For example, several studies conducted in North America have investigated these benefits at a statewide and nationwide level. Charles McClure and colleagues found that public libraries in Pennsylvania and Florida benefit the local community and the state in a variety of ways, though they did not place a dollar value on those benefits. Leslie Fitch and Jody Warner found similar benefits for public libraries in Canada. Four other studies employed special methods of economic valuation to actually place a dollar value on the benefits provided by public libraries and have found impressive returns on investment: José-Marie Griffiths et al. found that the ROI for public libraries in the state of Florida is $6.54 returned for every $1.00 invested; Sarah Aerni and Donald King found a $5.50 return on $1.00 invested in the state of Pennsylvania; Daniel Barron et al. found a $4.48 return on $1.00 invested in South Carolina; and the firm Levin, Driscoll & Fleeter found a $3.81 return on $1.00 invested in southwestern Ohio. It is important to note that the differences in the ROIs for these states are due to differences in valuation methods employed in these studies and not to higher or lower quality of service provided in some states than in others. These studies are all large scale, addressing the issue of the economic value of the existence of libraries in a state as a whole.

By contrast, few studies exist that address the smaller scale issue of the value of specific library services. One of the few studies of this type was conducted by Anne Morris, Margaret Hawkins, and John Sumsion. This team investigated the value to the
library user of borrowing books from libraries (as opposed to buying the same book) and found that most users would be willing to pay between 30 pence to £1 per book (approximately 55 cents to $1.80 US, at the exchange rate as of this writing.) A few studies have been undertaken to assess the costs of acquisitions of periodicals and other full-text resources and of technical services more generally.\(^\text{13}\)

There also have been few ROI studies of specific services offered by a single library to its user community (as opposed to services offered by libraries generally, as with the studies cited in the previous paragraph). Indeed, there are far more studies with the ambitious goal of determining the ROI for the existence of libraries at a statewide level than there are studies with the more modest goal of determining the ROI for a single library. The few that have this more modest goal include a study by Glen Holt, Donald Elliott, and Amonia Moore, who found that the ROI for the St. Louis Public Library is greater than $4.00 returned for every $1.00 invested.\(^\text{14}\) David Harless and Frank Allen found that student users of an academic library were willing to pay $5.59 per semester and that faculty were willing to pay $45.76 per year for the continuation of the existing library reference service.\(^\text{15}\)

To find ROI studies of library reference services one must look farther into the past. A number of studies in the late 1960s and 1970s arrived at surprisingly high, as well as a surprisingly wide range, of cost-per-transaction figures—from $2.20 to $10.80 per reference transaction.\(^\text{16}\) Inflation-adjusted, these figures correspond to approximately $12.00 to approximately $60.00 in 2008 dollars.\(^\text{17}\) However, these studies focused on desk, as opposed to virtual, reference services, and Abels argues that many of these studies suffer from methodological problems.\(^\text{18}\) Virtual reference services must add to any other expenses the cost of licensing software, the cost of maintaining servers to host the software, the cost of training librarians to use these applications, and other costs related to the service being computer-mediated. In fact, Brenda Bailey-Hainer suggests that virtual reference is likely to be cost-effective only at the statewide scale and predicts that an individual institution will be unable to sustain this service over the long-term.\(^\text{19}\) If true, this could account for the many cases of virtual reference services that have been discontinued in recent years.\(^\text{20}\) What this also makes clear, however, is that given the difficulties and cost of sustaining a virtual reference service, it is all the more important that libraries identify the costs they will incur by providing this service, whether individually or as part of a collaborative.

Although there have been almost no published cost studies of reference services in decades, some authors have articulated methods for conducting such cost studies and have detailed the categories of costs that must be taken into consideration. Eileen Abels, Paul Kantor, and Tefko Saracevic, for example, describe a functional cost model they developed to assess the unit costs of nine reference services. Their explanatory article provides key insights into the structure of cost models, issues related to capturing costs for this type of service, and relevant tasks and activities to consider when assigning costs to reference services.\(^\text{21}\) R. David Lankes, Charles McClure, and Melissa Gross articulate a set of methodologies that may be used when assigning costs to reference services and define a set of suggested cost categories.\(^\text{22}\) Gross, McClure, and Lankes continue this work, and present three measures that they identify as the most crucial data points that any cost study of reference service must identify: the total cost of providing digital reference service, the cost of digital reference service as a percent of the total reference
budget, and the cost of reference as a percent of the total library or organizational budget.\textsuperscript{23} Neil Kaske and Julie Arnold, motivated by similar questions to the study presented here, identify some of the costs involved in providing a virtual service and make some simple calculations of the total cost of the service to a library.\textsuperscript{24} The only cost model template for virtual reference service that has been created prior to the one presented in this article, however, was developed by Matthew Marsteller and Susan Ware.\textsuperscript{25} Their model provides information regarding both traditional best practice measures and newer, digitally oriented measures. During an early iteration of the model, they utilized "block of time" estimation, which estimates the number of hours provided by the relevant personnel to the service without breaking down any of that time for activities not directly related to reference service such as checking e-mail or preparing educational materials. Acting upon Abels' critique of then current costing methodologies,\textsuperscript{26} they created a second version that removes extraneous time from cost calculations.

Finally, another interesting approach to cost modeling is presented by Claire Hulme, who conducted a cost effectiveness study to compare two library services within a hospital library.\textsuperscript{27} Although Hulme's paper provides excellent background information on how to design a cost study and collect cost information, as a case study, it offers no ready-to-use model that a library could adopt for its own purposes. In developing the model presented here, we use a methodology similar to that described by Hulme but also provide a practical template for libraries to use to capture their virtual reference costs.

Background on Cost Models

The assessment of cost by an organization can serve a number of different purposes, and the design of the cost model will be determined by the primary purpose that it is meant to serve. Broadly speaking, cost models are distinguished as serving cost-benefit, cost-effectiveness, service enhancement, or budgetary purposes. In each of these cases, an organization needs to articulate accurately the full costs of engaging in a particular process; thus, a fully formulated cost model should be able to serve any one of these purposes. Most ILS literature on costs is concerned only with cost-benefit and cost-effectiveness analyses, though the other two are important and deserve wider use in ILS.

Cost-benefit models are attempts to compare the benefits of a venture to the costs, recognizing that any benefits or costs over time are subject to the time value of money. Cost-benefit models can be difficult to construct in the non-profit sector of the economy, and libraries may be considered to be non-profits since they do not typically generate revenue through sales of their services. Benefits to a non-profit may be financially tangible (such as another institution paying for an expense like software licensing fees for which the library now does not have to pay), but there may also be intangible benefits (for example, increased patron satisfaction). A cost-benefit model should quantify as many benefits and costs as possible and should also present intangible benefits and costs in a manner that allows them to be properly examined side-by-side with the purely financial analysis. Cost-benefit models usually differentiate between up-front costs and ongoing costs since up-front costs are often paid from different accounts than ongoing costs. Since the value of money changes over time due to inflationary pressures and interest rate fluctuations, the impact of interest rate compounding can cause ongoing
costs to have an inordinate impact on the cost-benefit ratio of some projects.

Alternatively, other projects may find that the impact of a large up-front investment has the primary impact on the cost-benefit ratio. It is generally impossible to make comparisons between choosing a large up-front cost versus a long period of payments until after one conducts a cost-benefit assessment that includes assumptions about expected interest rates and performs discounting of future streams of money to determine their present value.

A library considering the implementation of a virtual reference service should consider undertaking a cost-benefit analysis prior to finalizing its implementation decision. As with all library operations, relatively fixed budgets imply that offering a new service may well reduce the expenditures available for other service offerings. Opportunity costs abound in library management with a key question frequently being—What must I give up of this to offer more of that? The cost-benefit analysis forces library personnel to explicitly voice the benefits of the new service while enumerating all of its costs and where these costs include the potential lost value that occurs when another service area must reduce some activities in order to reallocate funding to the proposed new service area. Indeed, such an analysis may well require comparing the value expected to be received by one category of patrons to the potential loss of value to another category of patrons. Still, these value decisions are a hard fact of life that all libraries must make at all times anyway, and cost-benefit analysis helps to make these decisions more rational by providing the facts.

Cost-effectiveness models, like cost-benefit models, require a clear delineation of the costs of engaging in an activity or set of activities. Unlike cost-benefit analyses, however, the focus is not on comparing the total costs of an endeavor to its total benefits. Rather, the focus of cost-effectiveness analysis is upon assessing the efficiency or effectiveness of the outcome. Lancaster notes that "cost-effectiveness analyses can be thought of as studies of the costs associated with alternative strategies for achieving a particular level of effectiveness." Such a study essentially asks whether the final outcome was achieved as inexpensively as it could have been or if alternative ways of doing the job would have been less expensive. Another way to say the same thing is to ask how effective were the results given the amount of money spent?

A cost-effectiveness model requires two primary components. First, it requires a definition of "effectiveness." In practice, this implies articulating all benefits provided by the activities being assessed and developing a set of standards for defining "more" and "less" effective. For example, if a virtual reference transaction involves follow-up work that is completed by e-mail, one must first decide whether the entire communication with the patron is one reference transaction or two. If the answer is "two," does the first component meet quality requirements even though the question wasn’t answered at that time? If the answer is "one," how much weight is given to each component in order to consider the transaction to have met quality standards? Second, a cost-effectiveness model requires an assessment of all costs at an appropriate level of granularity. Continuing the same example, the costs of a virtual reference transaction include the librarian's time spent answering the question via chat and all other media, a parsing of network and other technical fees associated with that time period, and any other associated costs such as for telephone calls and paper and printing.
With respect to the cost effectiveness of virtual reference one might ask whether the outcomes achieved are worth the costs involved to achieve those outcomes. For example, Arnold and Kaske examined the number of correct responses to chat reference questions and found that the quality was much higher than expected—in the range of 90+ percent, as opposed to earlier studies of in-person reference transactions that found only about a 55 percent accuracy level. They note both the possibility that methodological differences between the earlier and the newer studies may account for this difference and the possibility that in-person reference transaction accuracy rates may also have improved since the earlier studies took place. More research could clarify what is leading to such significant increases in accuracy. However, with information about accuracy rates, one is in the position to ask whether the expense incurred to reach these levels of accuracy is justified. After all, chat reference requires up-front infrastructural expenses and staff training expenses as well as the ongoing expenses of maintaining the service, and these expenses alternatively could be used for potential in-person reference service improvements. In addition, costs will vary depending upon whether one chooses a collaborative arrangement through services such as Questionpoint or a simpler IM-based service using a tool such as America Online Instant Messenger. Cost effectiveness analysis in combination with an accuracy study will allow one to determine which combination of in-person and chat reference services (using which technical applications) will provide the desired level of accuracy at minimum cost.

The issue of "desired level of quality" raises some of the truly difficult aspects of cost-effectiveness studies, however. In general, the most difficult area of a cost-effectiveness study is not the articulation of the costs. Although some costs are, indeed, notoriously difficult to assign to particular services, the total costs to the library can generally be ascertained. Some formula, however arbitrary, can be applied to allocate the costs among the different library services once the benefits are defined and effectiveness criteria applied to them. The most difficult area of a cost-effectiveness study is the articulation of the benefits, which are used to define what is meant by "effective" use of funds spent.

A service assessment is a type of cost-effectiveness analysis. Service assessment seeks to understand, for some activity or set of activities, whether there are more efficient ways to achieve the goals of those activities. Rather than being focused primarily upon a number of alternatives for achieving a particular goal, however, a service assessment focuses upon whether the already chosen process used to reach a particular goal can be made more efficient. In other words, although a cost-effectiveness study will endeavor to determine, for example, whether outsourcing virtual reference service is more cost-effective than internally staffing the service, a service assessment will take whatever service method has already been chosen and will seek to understand whether the processes used in that service method are being performed as cost-effectively (that is, as efficiently) as possible. To perform a service assessment, costs need to be defined and collected at a very granular level so that the specific costs of all activities associated with providing a service can be related back to the processes of which they are a part. This, of course, requires a detailed understanding of the workflows involved in providing a service, as well as methods for collecting cost data. With such information in hand, however, a service assessment can be a very powerful management tool, one that will
even allow "tests" of different processes to occur, allowing continual process improvement without disrupting ongoing work.

Budgeting will occur even in absence of a cost model, but accurate assignment of costs and benefits is necessary to maintain fiscal control of an organization. A good budget will highlight all sources of costs and revenues and will accurately predict future spending and revenue streams. A fully developed cost model enables more accurate budgeting and easier management of the budget by ensuring that the necessary spending streams have been well considered and comprehensively outlined. With such a cost model, populating a budget becomes a relatively straightforward issue of summing up line items to fit whatever categorization scheme an organization uses for its costs.

**Development of the Model**

The cost model described here has been designed to be capable of serving any, or all, of the four goals just described. It expands upon Marsteller and Ware's model by including more of the relevant service costs, such as telecommunications costs, consortial fees, training, and so on, and by allowing the library to estimate actual hours spent on relevant activities. It also takes into account the flow of expenditures over time and discounts future expenditures according to a library-selected discount rate, making our model also appropriate for multi-year cost assessments.

**Methodological Considerations**

A decision about the implementation, continuation, or discontinuation of a virtual reference service (or indeed any library service) is a major one and has significant impacts on the library's budget, staff, and resources. This decision must, therefore, be based upon a rational comparison of the expected costs of implementing and maintaining that service and the benefits that will accrue to the library and its patrons due to the service. Assessing these costs and benefits involves articulating the actual benefits that virtual reference services provide to the library's users (for example, research assistance or user education), enumerating the activities and resources required to provide these services, and determining the extent to which the use of these resources affects the overall costs of providing the service.

It sounds simple in the abstract; just figure out what you provide that patrons want, what resources go into creating it, and how much those resources cost. In practice, however, one runs up against the perennial problem of costing in a service-oriented environment. Because the resources that are used to provide services are often applied to multiple non-separable activities, the costs of the various services are very difficult to differentiate. For example, Internet access is frequently utilized in reference interactions to search the library's online catalog or subscription databases. Internet access is generally provided throughout the library, however, and not just to the reference desk. Internet access may be utilized for interlibrary loan requests, copy cataloging, and a host of other library services, as well as by patrons at public access computers and staff at their desk computers. How can one assess the "amount" of Internet access that is utilized by the reference service as opposed to any other library function? For another example, consider the ultimate library resource—the reference librarian. A reference librarian may
answer reference questions in person, answer questions by e-mail and chat, may develop educational materials, and may plan the next staff meeting, all within the span of a few hours at the same desk during any given day. The librarian may have no exact hour-by-hour daily plan but instead may allocate his or her time on the basis of a wide range of factors, from relatively unpredictable patron volume to the imminence of upcoming meetings and class sessions. How does one assess the resources required for reference transactions as opposed to class preparation or meeting planning?

Carol Spencer noted that two separate problems exist here. The first is the difficulty of estimating the "fragmented time" of reference librarians, and the second is the difficulty of defining the specific units of work, such as "reference transaction," in a manner that is widely acceptable, complete, and easily measurable. Unfortunately, challenges such as breaking down costs to a useful and usable level of granularity for controlling operations, assessing the benefits and costs of offering various levels and types of services to patrons, and justifying these offerings often seem so overwhelming that this analysis simply does not happen. John Edward Evans notes that "despite years of repeated calls for cost data on services, the response of the library profession has been poor."

However, this avoidance technique is akin to throwing out the baby with the bath water. By foregoing the effort of determining the costs of providing various services, a library is unable to assess accurately what services are provided to patrons, the extent to which patrons value these services, and how much it costs to provide each service. Without the latter information, one cannot make a rational assessment of which services provide the most bang for the buck, particularly during periods of budget crisis such as we face today in the library environment. Lacking this data threatens the long-term viability of all services.

As noted above, most current library ROI studies focus on the benefits that libraries provide to society at large. Although the value of these studies to the library profession is great, it must be recognized that, when virtual reference services are being considered for implementation or continuation, this decision is usually made at the library level. This is certainly the case when virtual reference services are being considered for discontinuation. The 2006 NCES Academic Library Survey shows that, in the 2006 fiscal year, 72 percent of academic libraries provided online reference service, whereas a 1999 study reports that 96 percent of member libraries of the Association of Research Libraries provided online reference service. Given the prevalence of this service, the lack of studies of the costs of this service is quite surprising and highlights the importance to the profession of understanding the cost structures of these highly popular services. This present study was conducted in order to provide a tool that will allow this major gap to be filled.

The reasons why few cost studies have been performed at the level of analysis of the individual library are varied, but Evans notes that the single most prevalent reason given in studies is staff resistance to conducting cost studies. Among specific reasons for this resistance are fear that one's performance is being measured, an anti-measurement culture in intellectual environments, distrust of data, and concern that the cost study will play such a small part in decision-making that it will have been a waste of time.

Another probable reason for the lack of cost studies is the difficulty of capturing many of the relevant virtual reference costs at a suitably granular level or even in a quantitative manner at all. Cost allocation is contentious for desk reference services, and
this issue is even more complex for virtual reference services. It seems obvious that the time spent by librarians staffing the service should be charged to the service. Since virtual reference services are generally staffed in shifts, this cost would be some fraction of an individual librarian's total salary, calculated according to the amount of time the librarian spends staffing the service. Librarians staffing reference services, however, may perform other tasks during the service's idle times, making this calculation considerably more difficult. It seems equally obvious that some portion of the budget for computers and network access should be borne by the service since without those it would be impossible to operate a virtual reference service. Computers and network access are, however, used by many other library functions—searching the online catalog and databases, interlibrary loan, course reserves, e-mail, and a host of others—so, clearly, the costs of computers and network access cannot be borne entirely by the virtual reference service. What portion of these costs is appropriate to charge the virtual reference service, specifically, rather than to other tasks and services?

The issue of cost allocation becomes even more complex when considering the cost of information resources used by the service. Some researchers argue that the entire cost of the library's physical reference collection should be borne by the reference desk service; others argue that none of the cost of the reference collection should be borne by the reference service and that the cost of all collections should be borne by a separate fund. Still others, such as Samuel Rothstein, argue that the cost of the reference collection should be split among several library departments. As Rothstein suggests, much of the use of the print reference collection is by users other than reference librarians and for purposes other than answering patrons' reference questions. This is equally true for electronic materials; libraries subscribe to many online information sources that are used by the reference service as well as by other library departments and by library patrons. Should the cost of these online sources be borne by the virtual reference service alone, and, if not, how should it be apportioned? Ultimately, the answers to these questions are dependent upon the particular accounting climate of the individual libraries and must be answered by them, based upon their accepted policies and procedures for costing their services.

**Identifying Cost Drivers and Assigning Costs to Drivers**

Engaging in a cost study is costly in its own right, and one will not engage in this enterprise without a motivation for doing so. However, once the goal and scope of the cost study have been clarified, delineation of all the costs within the scope of the project is required. In practice, this amounts to deciding which line items from the model need to be populated, and this involves listing all the cost drivers of one's virtual reference service and the activities and resources that are required to produce each of those drivers. The cost drivers are the outputs that cause changes in the overall costs. For example, virtual reference transactions will drive costs because providing more transactions costs more than providing fewer transactions. This is, in part, because increases in virtual reference transactions have an impact on the number of librarian FTEs needed to staff the service, thus increasing the proportion of librarian salaries that are to be attributed to the service.
To determine which activities and resources go into "creating" the cost drivers and what quantity of these resources and activities are required involves a detailed understanding of the virtual reference workflows. Recalling an earlier question, for example, is a virtual reference transaction only composed of the initial time spent online with the patron, or does it also include all follow-up work required to answer the patron's question, even if the process is finalized using a different medium such as e-mail? Whatever answer is chosen, this will determine the components that need to be costed to produce a transaction. (One guideline that should be considered for making this determination revolves around the questions of what particular steps will make the patron feel they have received what they are looking for, and what level of granularity does one need to break down these steps to understand how much it costs to create a satisfied patron.) Then the expected number of such transactions can be multiplied by the cost per transaction to come up with a total cost per given time period.

After cost drivers are determined and costs associated with these drivers are defined, the library needs to determine which of these costs are direct costs, which are indirect costs, and which are general (administrative) costs. Direct costs are those which can easily be tied to an individual product or service. For example, the cost of the virtual reference software license can be attributed directly to the virtual reference service. Indirect costs can be tied to two or more outputs, but not to all outputs. For example, librarian wages and salaries may span both virtual reference and in-person reference services. Technical personnel wages and salaries likewise may be attributed across several departments. Finally, general or administrative costs cannot be tied to any particular product or service. For example, the utility or facilities costs associated with the building that houses the virtual reference service are often associated with all library services and products.

In general, the direct costs will be the easiest to find and measure. The question to ask in order to identify the direct costs is—Is this cost entirely attributable to the virtual reference service? If so, then it need only be appropriately assigned to the chosen timeframe (for example, license fees per year) and populated directly into the model.

Indirect costs may be more difficult to measure. For example, if the particular librarians staffing the service and the hours they staff the service are known, one could use their actual salaries to estimate total wages and salaries for the service. However, since librarians may work on activities that are not directly related to the virtual reference service while they are staffing the service, their hours need to be allocated appropriately to all the cost drivers (such as services) that they provide, and determining the allocation formula can sometimes be challenging. This is where shadowing or work logs may be utilized. Another difficulty with determining indirect costs occurs when an activity crosses departmental borders, as is often the case. For example, suppose it is determined that more online texts or journals will be needed if the virtual reference service is implemented. In such a case, acquisitions and technical service personnel will be needed in order to ensure that these materials are provided for this service. One must determine how to appropriately allocate the costs in such a case, often a politically charged undertaking.

Overhead costs are relatively easy to delineate but can be very difficult to allocate among the various library services. A decision regarding such allocation is imperative, however, since a portion of these costs allow the virtual reference service to continue.
Perhaps the trickiest costs about which to make decisions are capital costs that were incurred in a previous year but for which the library is still subject to ongoing payments. The general practice is to leave the initial up-front cost of such resources out of the study if the resources were purchased prior to the decision to implement the virtual reference service. They are sunk costs and irrelevant to the reference service itself. If these resources are used for the provision of virtual reference services, however, any ongoing payments to them need to be allocated appropriately to the reference service.

This model attempts to give a highly granular list of likely areas in which relevant costs will accrue. Each individual library can determine the level of detail best suited for its particular costing needs.

Collecting Cost Data

Once all relevant costs have been identified, data collection is necessary. This will often be the most time consuming aspect of the cost assessment. Generally, estimates and assumptions need to be made (and documented) in order to assign the appropriate portion of indirect and overhead costs to the virtual reference service, and this will require discussion among the staff and documentation of decisions and their rationale for later justification of the accuracy of the model.

The most accurate way to assess the nature of all costs is via process analysis, such as shadowing or even detailed task logs kept for a short period of time. Although this type of endeavor is a short-term interruption, it can provide such detailed information about proportions of work time spent on various activities that it generally is worthwhile, especially for a library that is conducting a cost-effectiveness or service assessment study.

Some libraries, however, decide not to engage in this type of detailed analysis, often due to a shortage of available resources to conduct such a cost study. If that is the case for a particular library, the fallback method is to use the current budget. In this case, these budget documents will need to be analyzed to determine which budgetary items include resources necessary to provide virtual reference service. Then, assumptions regarding what proportion of these total budgetary costs are applicable to the reference service must be applied.

One can easily see why this type of gross costing method tends to be much less accurate than the micro-costing method that relies upon workflow analysis. In fact, using a budget to populate the cost model introduces inherent inaccuracies into the cost study, and, from a finance perspective, really would not be considered a proper cost study. This is because most budgets are the result of a complex bargaining process, in which the budgeting officer or administrator develops an initial budget that represents his or her "most desired" outcome, perhaps tempered by some realistic assessment of what is actually likely to be received. Sometimes some "padding" may be added to requested budgets because the administrator knows that the full amount requested will most likely not be received or because some extra is added for contingencies. Once the budget is approved, there is pressure to spend all the money allocated since future budgets are usually reduced if it is not all spent. This leads to the tendency to inflate costs by spending money on desired resources that may not, strictly speaking, be absolutely necessary for the provision of current services. (Although, it is true, these "excess"
expenditures can often allow an increase to the number of services otherwise available.) As Michael George writes in his book *Lean Six Sigma for Service*, a well-known and respected work on service: "Empirical data have shown that the cost of services are inflated by 30–80 percent waste—that is, the processes are riddled with activities that add no value from the perspective of the customer."\(^{45}\)

Although an assessment of the past costs of producing the planned services often occurs during the budget creation process, the time for constructing the budget is often short, and it is almost unheard of for a full assessment of the cost efficiency of the past year's services to take place while developing the next year's budget numbers. As a result, past costs generally do not reflect true costs of service production; rather, they reflect a relatively inefficient cost of production. Moreover, after the final budget is determined, a reshuffling of internal work priorities will usually occur when it is recognized that all the desired activities cannot happen within the actual budget allocated. In other words, when future costs are estimated on the basis of past budgets and operations, they tend to lead to an often hazy assessment of what resources will be needed to do the job in the future. Furthermore, because this assessment is based upon current, not optimal, work processes, the costs of producing a given service are usually overestimated when using this costing technique. Questions of the effectiveness or quality of service provision are usually treated as a completely separate endeavor, a "management" project or "clean up" effort, and are not tied to the budgeting process itself. As a result, there may be no firm link between the true costs of producing a service and the actual budget that is designed.

Although a library may well recognize that using budgets to estimate the costs of providing a service can be misleading at best, resource restrictions may still lead one to perform this analysis on the logic that some information is better than none and that beginning to include the process of service costing is a step forward, even if the "ideal" cost assessment cannot be performed at that point in time. In fact, just introducing cost assessments into one's operations is a huge step forward for many libraries and can lead to increased knowledge of improved ways of providing services to the patron. The caveats introduced in this article, therefore, should be not be taken as a warning, simply as information to help one interpret one's final cost assessment. Workflow analysis is time consuming and expensive, and one's reasons for conducting the cost study may not justify such a resource intensive process. That does not mean that some form of cost assessment should not be undertaken. In either case, increased control over operations can be gained via a cost assessment.

Once the assumptions regarding the allocation of costs have been made, all that remains is to populate the line item detail into the model. The summation across categories and years occurs automatically, and the roll-up to the multi-year calculation of total costs and net present value of costs also occurs automatically. This will allow a library to modify assumptions and determine how the multi-year cost calculations are affected by these changing assumptions.

*Data Sources*

Two sources of data were used to populate this cost model: data from The National Center for Education Statistics' (NCES) Library Statistics Program, nces.ed.gov/surveys/libraries/, and data from the budgets of individual libraries. The
NCES' Library Statistics Program collects data on academic, public, school, and state libraries across the United States. (As of 2007, however, the Public Libraries and State Library Agencies Surveys are being implemented by the Institute of Museum and Library Services; see harvester.census.gov/imls/publib.asp.) These data include such variables as the library's expenditures on staff and paraprofessional salaries and benefits; expenditures on books, serials, and other materials; expenditures on computer hardware and software; and a wide range of other data points. These data are somewhat dated; at the time of this study, the most recent data on academic libraries were from fiscal year 2006 and the most recent data on public libraries were from fiscal year 2005. Despite the age, however, these data were still used. Although the specific sums allocated in individual libraries' budgets may change over the span of years, the categories into which libraries allocate expenditures change more slowly.

The most recent available data on library's expenditures were collected from libraries' actual budgets. It was a minor surprise for the researchers that reference services in libraries often do not maintain their own budgets; rather, the reference service (and most other library divisions, for that matter) is wrapped up into the library's budget as a whole. This is stated not to suggest that the reference service and other library divisions do not create their budgets independently. In the final library budgets, however, these divisions are not uniquely identifiable as line items. All expenditures for salaries, for example, are often pooled across all library divisions, so it is difficult to identify only those salaries that are unique to the reference service.

Budget data for this study were collected from one large academic library and one large urban public library. Both of these libraries were part of larger library systems. The academic library was the main research library on a university campus with several other more specialized libraries, and the public library was the main branch of a city-wide system with several branches. These libraries were both in North Carolina, where the researchers are located, and both libraries were amenable to the idea of identifying the cost of the reference service.

An obvious limitation of the data collected for this project is that the two libraries that provided budgets are both large and comparatively well-funded libraries. As such, these libraries may have specific costs, and even entire categories of costs, that smaller libraries may not have. This was, in fact, one of the reasons why large libraries were selected for this phase of this project. It was necessary to analyze budgets that contain as many costs and cost categories as possible in order to develop a cost model that is as exhaustive as possible. This cost model can then be used by smaller libraries with fewer costs by simply ignoring irrelevant costs and cost categories.

Another limitation of this data collection is that the libraries that provided budgets were self-selecting. This was unfortunately necessary due to the sensitive nature of budgetary data. Another minor surprise for the researchers was that many libraries were unwilling to share their budget data for this project. The budgets of public libraries and public educational institutions are, of course, public information and are often available on the city's or the university's Web site. Budgets are rarely perfectly transparent to those outside the institution, however, and it was necessary for the researchers to contact the two libraries to clarify the scope of some items on the budgets. As such, having the cooperation of individual managers in these libraries was invaluable. It was the authors' intention from the outset of this project, however, to develop a cost model that any library...
could use unassisted, and the authors believe that this has turned out to be the case. In other words, libraries unwilling to share budget data need not do so but may still benefit from the improved cost assessments that this model could allow them.

Structure of the Model

The model developed for this project is composed of three worksheets in one Microsoft Excel workbook (version 97-2003). The model was developed in Excel because most libraries will have access to that application and because of its relative ease of use. Because the model is in Excel, an operational version cannot be reproduced here. The Excel file containing the model can be found online at <PLACE MUSE URL HERE>.

The three worksheets of which the model is composed are as follows: (1) start-up costs, (2) ongoing costs, and (3) a summary worksheet (called "NPV of Costs") that provides roll-up sums of the costs from the two other worksheets. If a library is beginning to cost a reference service that has already been implemented and is now fully operational, the "start-up costs" worksheet can be ignored entirely. "NPV" means net present value, which is the value of all costs, present and future, expressed in terms of the present value of "today's dollars." The summary worksheet provides the totals by year for each of the categories of costs provided in the model, as shown in figure 1. These categories of costs were based on those described by McClure et al., with some extensions to the categories and their line item details based upon common cost accounting practices. The category totals are taken directly from the other two worksheets' corresponding totals. Total per year and category should always be identical to the total for that category's year on the corresponding detail worksheet (such as start-up costs or ongoing costs). In addition, at the bottom of the summary worksheet are two important lines, the "discount rate" and the "present value of costs." The discount rate should be populated with an appropriate interest rate for the library in order to represent its best estimate of the expected annual growth in overall costs, which will include both the inflationary expectations for labor and the inflationary expectations for other resources such as supplies, reference related materials, licenses and service contracts, and so on. Because of this, it will be some weighted average of the expected annual growth rate in the costs the library incurs to provide reference services. Once the discount rate has been chosen, the present value of costs is automatically calculated from the totals by year for the start-up and ongoing costs. The first year's costs (both start-up and ongoing) are not discounted since they already represent today's dollar. All future years are appropriately discounted.

<INSERT Figure 1: Summary Information and Present Value of Costs ABOUT HERE>

Although some cost models will allow an organization to choose separate interest rates for labor and resource costs, this model assumes that the library will average these potentially different inflation rates to use an average inflation rate. For example, perhaps the growth rate in labor costs is only 3 percent but the growth rate in other resources is 12 percent. The library would compare the percentage of all costs for labor and the percentage of all costs for other resources and create a weighted average that
represents an overall inflationary rate. (The weight for each component would represent its relative proportion within the library budget.) This is the rate at which one would expect total virtual reference costs to increase over time, even if the number of librarian FTEs or expected quantity of resources used were to remain constant over time. A future improvement to this cost model could allow for different inflation rates for different costs, although for most purposes the current level of granularity in inflation rates will work perfectly well.

Since the model is prospective, that is, it assumes that the study begins at time zero and that discounting begins one year from that time, if past costs are included, one would need to represent today's costs as inflated costs. For example, if the study inputs costs from two years ago as year one costs, one will need to use the inflation rate over the past two years to inflate today's dollars when inputting them into the year-three column. A future improvement to this model could be to allow the first year to be chosen by the library and to automatically perform that calculation.

The structure of the "start-up" tab (figure 2) and the "ongoing" tab (figure 3) are identical, including the same categories and line items. The only difference is that start-up costs are assumed to be those costs incurred during the implementation of the virtual reference service, whereas ongoing costs represent the costs from the launch of the service onward. It is possible, therefore, that the ongoing costs for year one will be zero if it takes more than one year to implement the service. It was assumed that an implementation would take between one to two years, so the ongoing costs are discounted starting in the second year. In other words, the time value of money (that is, the discount rate) begins to be taken into account at the end of the first year and is calculated as a yearly estimate. Although it would have been possible to track costs on a monthly or quarterly basis, most libraries create their budgets annually; and, therefore, annual tracking should be sufficient. Additionally, this model calculates present value over a five-year timeframe, a common enough present value assessment for smaller scale technical and process implementations.

Limitations of the Model

A major difficulty with conducting cost studies in the realm of library reference services lies in getting the data. Even when one knows which categories of costs are applicable to virtual reference services, the accounting practices in most libraries make accurate costing of individual services difficult to achieve. Jennifer Ellis-Newman has noted that one source of this problem lies in the tendency of libraries to rely upon accounting structures that roughly mimic those used in manufacturing organizations over a decade ago. Rising costs and increasing competition led the manufacturing sector to develop "new tools and techniques for measuring and allocating costs, while in the process gaining a better understanding of costs and cost behavior." Many manufacturing organizations have, therefore, moved away from traditional fund-based systems and toward activity-based costing, which relies upon determining the actual services (or
goods) produced and then allocating costs toward those services. Libraries have not followed suit, however. Their more traditional accounting systems focus on the budget, deciding how to allocate resources based on staff workloads, "with little consideration given to the actual cost of providing services." The model presented here assumes that populating the model will occur after a reflection on the cost drivers by the virtual reference service. Costs will be allocated based upon applying direct costs to those drivers and spreading the indirect and overhead costs by evaluating the proportion of those services provided to other services provided within the library. This should create a costing approach that is more closely aligned with the actual provision of services rather than upon traditional workload customs.

In general, the decision regarding what proportion of a reference collection to charge to the reference desk is a highly arbitrary one. Again, however, if one has clearly delineated all the services provided by the library, has defined those services as cost drivers, and has linked the work processes and resources to those drivers, one will be able to assess what proportions of overhead costs should be assigned to which drivers much more accurately. It should be noted, however, that Robin Cooper and Robert Kaplan, who first advocated activity-based costing, recommended that excess capacity not be allocated directly to services. This would imply that, if an organization purchased additional hardware in order to have a buffer for future growth, this excess capacity should not be included in the current service provision costing.

Another difficulty associated with assessing costs is related to the difficulty of predicting reference transactions. In order to accurately cost a service over time, it must be possible to make good estimates of future flows. This will allow a full understanding of capacity and resource requirements. Currently, no good methods for predicting future flows exist for reference services. There is, however, documentation of regular fluctuations in the use of virtual reference services corresponding to the academic calendar. Additionally, transportation models allow call centers to appropriately staff their facilities and to fine tune their staffing decisions on an ongoing basis. Implementing such a transportation model in the virtual reference arena, where similar ebbs and flows in chat transactions are known to occur, may help a library better predict its staffing needs and, thus, its staffing costs for the service.

Some potential enhancements to this model are possible for the future. For example, the current model does not include risk assessment techniques or templates, although a contingency figure is included for the library to add such an estimate. The risks to a library of providing virtual reference service are not hazards to life and limb, of course, but rather are hazards that may interfere with the successful implementation of the service. A future version of the model could include templates that highlight potential risks of providing this service and means for assessing their likelihood and cost impacts. This is much in the spirit and practice of the relatively recently published DRAMBORA trusted digital repository toolkit, which offers a risk toolkit for digital repositories to use to mitigate the risks inherent in digital preservation. Although the categories of risk faced by virtual reference services may differ somewhat from those existing in the digital curation arena, the methodology employed by the Digital Curation Center is a sound one to use for identifying risks and may provide some very useful tips for a library to begin building its own risk assessment program.
Finally, no formal mechanism for performing sensitivity analysis has been included within the model. Sensitivity analysis is largely unfamiliar in the library arena. Briefly, it enables one to see how varying the resource allocation that goes into producing one's service will lead to variations in the output of the cost model. For example, a sensitivity analysis could identify the causes of the different returns from different investment strategies. Likewise, it could show, for example, how costs and number of reference transactions that a library is capable of providing would change if it were decided to use paraprofessionals to answer the most basic reference questions so that the library could reallocate more senior librarians to the virtual reference service or if it were decided to reduce the number of in-house librarians on the virtual service and use outsourced question answering instead. In fact, this type of analysis could be used to assess the most appropriate level of outsourcing to use by comparing questions answered and costs associated with various proportions of outsourced services over time. How much of an impact would these changes have on the number of reference transactions that could be provided and the cost of providing them? (This, of course, is not meant to indicate that quality of answers is irrelevant—a responsible cost assessment will always include potential and expected changes in quality and will explicitly take them into account along with any benefit-cost assessments.) Within this model, libraries could manually change their line item estimates and see the overall impact to the present value of costs on the summary tab; however, they must track those changes themselves. This type of analysis was not included in this version because we desired to keep the model simple to populate and easy to understand. A possible enhancement to the model could, however, be an additional tab that automatically performs sensitivity analysis on interest rates, labor costs, or other important costs.

Conclusion

The scope of this project was the development of a cost model that could be used by a library to assess the costs of providing virtual reference service. In order to identify the full set of costs to a library of providing virtual reference service, however, future research is still needed. For example, gross costing methods (that is, comparison to actual budget numbers) were used to engage in a high-level validation of cost categories. A more detailed (workflow-based) empirical analysis is needed to continue to validate the model's category and line item appropriateness. This would allow one to begin to understand how the costs of providing virtual reference through consortia may differ for different types and sizes of libraries. For example, do academic libraries show significantly different cost structures than do public libraries? Do full university research libraries show different cost structures than do community college libraries or independent research libraries? It would also be useful to examine different consortial structures to understand whether similar library categories within these different consortial structures exhibit significant cost differences when other variables are held constant. In short, a body of empirical data needs to be developed that will allow the costs of virtual reference in different libraries to be compared with accuracy and statistical significance.  

This project was the first stage in a two-stage effort to develop cost models both for individual libraries participating in a collaborative service and for the collaborative as
a whole. The second stage will involve incorporating the individual library costs into a larger model that will capture overall consortial costs. The methodology of this endeavor is likely to be somewhat different than what is presented here since economies of scale are likely to be involved and some costs may be borne by a larger funding agency and not by any individual library. If this hypothesis is true, it would suggest that incentives to provide virtual reference services are likewise skewed because the individual parties that make decisions regarding what types of virtual reference services to offer and "how much" to offer will not take these costs into account within their own private decision-making. Another challenge of this second stage will be to capture the different types of consortial agreements that occur and to model them appropriately. It is likely that a more broadly statistically oriented approach will be needed to undertake this effort.

The issue of virtual reference costs is an important one. More and more libraries feel a pressure to provide online services to their patrons, and very little is currently known about the cost structures of these services. Are libraries likely, over time, to move away from consortial services provided with commercial chat reference applications and toward more individual instant messenger-based services? What are the benefits to the library members of a consortium, and what are the costs? The answers to these questions will provide us crucial knowledge about what to expect in the increasingly online library of the future.

Acknowledgements

This material is based upon work supported by the OCLC/ALISE Library and Information Science Research Grant Program (LISRGP).

Lori Eakin is Ph.D. student, School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC; she may be contacted via e-mail at: lorraine.eakin@unc.edu.

Jeffrey Pomerantz is Assistant Professor, School of Information and Library Science, University of North Carolina at Chapel Hill, Chapel Hill, NC; he may be contacted via e-mail at: pomerantz@unc.edu

Notes

3 American Library Association, "Library Funding News," American Library Association,
http://www.ala.org/Template.cfm?Section=libraryfunding&Template=/ContentManagement/ContentDisplay.cfm&ContentID=139034 (accessed October 2, 2008).


18 Abels.
26 Abels.

31 Marsteller and Ware.


35 Horowitz, Flanagan, and Helman; Radford and Kern.


38 Evans, 37.

39 Ibid., 37–38.

40 Murfin.

41 Abels, Kantor, and Saracevic.


46 Charles R. McClure et al., *Statistics, Measures and Quality Standards*.


48 Ibid.

49 Ibid.


51 Ibid.
54 Of course, one significant component of the overall assessment of virtual reference service that was not part of the scope of this project is that of the benefits or value of the service. Items such as patron satisfaction, librarian satisfaction, value to the community, and so forth deserve their own analysis and integration with the cost information provided by this model. For an example of a model that focuses heavily upon the benefits of an information technology project, see espida, *espida Handbook: Expressing Project Costs and Benefits in a Systematic Way for Investment in Information and IT* (Glasgow: University of Glasgow, 2007), https://dspace.gla.ac.uk/handle/1905/691(accessed October 2, 2008).