

## 10.6 TeleCampus Student Numbers Survey

### Editorial

As DEOS-L moderator, one of the more frequent questions that I hear is "How many students are out there taking courses at a distance, anyway?" This has proved to be a very difficult question to answer, as the authors of this month's article found out. The authors circumspectly state that the response rate to their survey was too low to be generalizable, but some interesting facets do come to light, including the sensitive nature of enrollment data at the institutions surveyed.

The fragmented nature of distance education support and administration at most institutions is indicated by the difficulties these institutions experience in finding the right person(s) who could answer the survey questions. Some of the underlying presumptions made by the authors were that each institution to be contacted had a designated person for distance education courses; that this person could be located--either by perusing the institution's Web site beforehand, or by referral once the survey is received; and that this person would find the survey sufficiently useful to participate in it. The authors found themselves wrong on all counts.

Another interesting facet of this research is the insight that it provides into issues of compilation and maintenance of the databases underlying "distance learning portals." These are sites that offer a comprehensive list of courses offered by various institutions so students in search of a course or a degree program will have the convenience of having a "one-stop shop." Despite the best will in the world, TeleEducation New Brunswick discovered many deficiencies in their database and that constant update is necessary in this fast-moving field.

Mauri Collins

Editor

### TeleCampus Student Numbers Survey

Julia Innes

TeleEducation NB, Office of Learning Technologies Project

500 Beaverbrook Court, Fredericton, NB E3B 5H1

Rory McGreal

TeleEducation NB, Office of Learning Technologies Project

500 Beaverbrook Court, Fredericton, NB E3B 5H1

Anthony Roberts

TeleEducation NB, Office of Learning Technologies Project

500 Beaverbrook Court, Fredericton, NB E3B 5H1

## **Abstract**

A survey was undertaken to determine the number of students worldwide currently studying online. The authors surveyed 520 institutions whose courses are listed in the TeleCampus Online Course Directory to obtain enrollment data on a representative sample of 789 online courses randomly selected from a database of 17,000 courses. The primary objective was to determine how many students were enrolled in certain courses at an institution. However, the objective was not attained, due to the wide range of enrollment figures received and a poor response rate (14 percent). Although the poor response rate does not justify generalization of the data, the results may be of some use to researchers as a best guess.

## **Background**

This research project, sponsored by the Office of Learning Technologies, Human Resource Development Canada as part of a Workplace Learning Technologies Project, was designed to conduct research on online courses and provide information and tools that allow the learning technology practitioner and buyer community to access learning opportunities and improve the quality, productivity, and implementation of learning technologies in the workplace. The project included researching online courses in the TeleCampus Online Course Directory (<http://telecampus.edu>) and scanning the marketplace to provide data of use to online course developers and deliverers, particularly for those involved in workplace training. Tremblay (2000) compares the TeleCampus directory with more than twenty other course directories available online. The TeleCampus directory was rated first overall.

The TeleCampus Course Directory provides a full-text search engine for users who can also search using subject categories or institution. The category list allows users to search by subject (e.g., biology, architecture, classics, computer technology). For more detailed information, users are directed to the institutional Web sites. Other search features (such as searching by program, level, and country/state/province) are being built in. The aim is to keep the database as simple and useful as possible. The database has been built on an open architecture to allow additional fields to be added as needed.

The database fields in the TeleCampus Online Course Directory conform to The Institute of Electrical and Electronics Engineers, Inc. (IEEE) 1484 Learning Object Metadata standards, which are mapped to the Dublin Core and International Organization for Standardization (ISO) Standard 11179 as developed by Educause's Instructional Management System. The IEEE has formed a Learning Technology Standards Committee (LTSC), which includes ARIADNE (Alliance for Remote Instruction Authoring and Distribution Networks of Europe). LTSC has joined with the new ISO/IEC (International Electrotechnical Commission) subcommittee on Learning Technology (called SC36) to further the international development and acceptance of IEEE's Learning Technology (1484) standards. LTSC is working closely with SC36 on the development of common IEEE and ISO/IEC standards for the Internet and other forms of electronic transmission (Farance and McCabe 1999).

## **Project Rationale**

This study was intended to determine the number of students worldwide currently studying online. Institutions whose courses are listed in the TeleCampus Online Course Directory were surveyed for enrollment figures for a representative sample of nearly 800 online courses.

### **Project Expectations**

Journalists frequently ask TeleEducation NB staff how many students are enrolled in online courses. The question is usually limited geographically, and the journalist usually wants an order of magnitude ("thousands", "millions", etc.). TeleEducation NB felt that its database was eminently suited to extracting this valuable information and producing the world's first enrollment figures based in reality. In the area of distance education, nearly all the figures that circulate are guesses or extrapolations.

### **Goals and Objectives**

The survey allowed TeleCampus staff to come to terms with two independent issues:

- a. describing the types of errors likely to occur in the directory and appropriate solutions to these errors; and
- b. the likelihood of institutions with five or more courses in the directory to respond to a survey about those courses.

The survey had two objectives: locating and confirming the contact person for distance education within each institution, and finding how many students were enrolled in certain courses at an institution. The first objective was attained in that eighteen respondents made corrections to the contact information, and it may be presumed that those who made no changes nonetheless had verified the information.

The second objective was not attained. The survey was short (one page), with multiple-choice, yes/no and short-answer questions. As the only incentive to respond was the opportunity to view the survey results, much effort was put into making it as easy to complete and non-intrusive as possible. In retrospect, it could have been half as long. Future surveys could be only four or five questions long.

### **Barriers**

The most important unknown factor in this study was the reaction from institutions to requests to reveal sensitive data. No survey of this nature appears to have been conducted previously. Statistics Canada surveys adult education once every four years, but did not separate online education from home in its 1994 study (the results of the 1998 survey have not been published). No equivalent survey for US institutions was found. The lack of previous models made the survey methodology much more difficult to mount. In particular, the lack of upper and lower limits for certain variables in the study made it impossible to determine how large the sample size needed to be.

Language was another complicating factor. All communications were translated into French

and Spanish, as experience with institutions over the course of the previous year had shown reluctance for non-English institutions to communicate in English. Twelve copies of the survey were sent in French and three in Spanish. Six responses from French institutions were received, indicating a much higher response rate (50%) than for English institutions (34%). None were received from Spanish institutions.

Non-response rates for surveys have been increasing over the past decade. It was anticipated in planning the survey methodology that certain categories of institutions would not have sufficient results to be compared and analyzed with other categories.

## **Benefits**

The benefits of this project included raising TeleEducation NB's profile in the international distance education community and market. One hoped-for benefit was increased or continued funding for the TeleCampus Online Course Directory based on the final report of this survey. Another was to establish an annual survey which will provide reliable statistics on student enrollment and profile the status of distance education on the Web over time. Given the difficulties in obtaining data, these benefits could be revised to cover more modest expectations, e.g. profiling Canadian online courses, or online courses in a particular sector (traditional universities, community colleges, or the private sector).

The eventual commercialization of the directory depends in large part on TeleEducation NB's ability to show value-added for the information it has gathered over the past two years. Its short-term goal is to become a recognized name in distance education circles, and create a niche as a respected, neutral authority in establishing links between students, courses, and institutions. This survey was a preliminary step to see if institutions were willing to trust TeleEducation NB and use its data as the world authority on this subject.

## **Methodology**

*Schedule* The study was initially designed to be mailed out in late September 1999. Due to the availability of the directory in the third week of September, the sampling was redone by hand over a two-week period. The surveys were mailed out instead in the second week of October. A prenotification was done by e-mail during the week of October 3-8 to determine the preferred response method of respondents, and to lower postage costs. In addition, this offered respondents from overseas the opportunity to respond more quickly. A total of 62 requests for electronic surveys were made, generating 38 responses.

Mail-outs included a preaddressed, postage-paid envelope. The cover letter that accompanied the survey informed respondents that the results of the survey would be available on TeleCampus' Web site as of early December, and that late responses would be incorporated into the results on a continual basis after October 31.

*Assumptions* The underlying assumptions of this survey were that each institution to be contacted had a designated person for distance education courses; that this person could be located, either beforehand by perusing their Web site, or by referral once the survey is received; and that this person would find this survey sufficiently useful to participate in it. The

literature mentions survey fatigue as a major factor in non-response rates (Pourjalali and Kimbrell 1994). It was hoped that the novelty of this survey would motivate professionals in this sector to participate in order to have access to its results.

Results suggest that at least two of these assumptions were false. Institutions did not always have a designated person for distance education, and an appropriate contact person at each institution was not easily identifiable, either from the institution's Web site or by internal referral. Because it is not known how many surveys actually reached an appropriate person, the third assumption cannot be assessed. However, it is clear from the low response rate that few of those who received the survey found it sufficiently interesting to respond. It follows that perhaps little or no effort to locate the appropriate respondent was made. Further study is required to determine whether better targeting of the survey would increase the overall response rates.

The sample was drawn only from institutions with five or more courses in the database, and made it theoretically possible to estimate the worldwide population of students taking online courses in a statistically valid manner. The sample was stratified by type of institution in an attempt to maximize statistical validity. In the event that the brute results were insufficient, this approach allowed the researcher the possibility of analyzing the data by educational sector (university, community college, private company, school, etc.).

### **Telecampus Online Course and Program Taxonomy**

It is important to distinguish between the target population of all online courses currently offered worldwide and the survey population of online courses in the TeleCampus directory. No complete list of online courses worldwide exists, nor is there any means of determining how many courses might exist that have not been listed in existing course directories. The TeleCampus directory was considered to be a reasonable substitute for the target population because it was among the largest online course directories worldwide. Also, its narrow definition of "online course"--i.e., one with no on-site requirements or visits, which may be taken entirely at a distance--was deemed useful in eliminating courses with hybrid characteristics, e.g. short on-site visits, on-site exams, and no e-mail contact with professor.

An online course is defined as one that can be followed completely online. This does not mean that all course materials need be online. Books, CD-ROMs, video and audio tapes, laboratory materials, and other items could be shipped out directly to students. Examinations may be taken at local institutions or testing centers. The database would exclude courses with no online component as well as those that require residency.

On a continuum, the courses and programs that are advertised online continue (McGreal 2000):

Courses with no or limited distance availability:

1. Classroom-based courses with no online features;
2. Classroom-based courses with some materials available online;

3. Classroom-based courses that integrate online materials;
4. On-campus courses that are online but are not available to distant students;
5. On-campus courses that offer limited access to students at a distance (often limited to one region);
6. Teleconferencing courses where students must participate from specified learning centers.

These represent the majority of courses presently advertised. They would not be included in the database because they require residency, either on campus or near an accessible learning center.

### **Correspondence Courses (Print, Audio and Video Tapes, Software)**

7. Print-based correspondence courses using the postal system with minimal student support.
8. Print-based correspondence courses with continuing access to a tutor by telephone.

These also would not be included, as they are not available online.

All of the following class of courses would be included in the online course database:

### **Basic Online Courses**

9. Print-based correspondence courses also using e-mail for tutor access;
10. Correspondence courses with course content available online in electronic format.
11. CBT- based self-study courses with access online to an instructor.
12. CBT- based self-study courses with no instructor.

These courses represent the most primitive form of online courses.

### **Computer-Mediated Conferencing (CMC)**

These courses often include texts and sometimes include audio/video tapes, and computer software possibly including Computer-Based Training courseware. Students may need to download and install client-side software to participate.

13. Courses that use e-mail for submission of assignments and private tutoring and e-mail lists or listservs for discussions and tutoring;
14. Courses that use CMC software for discussions as well as e-mail for submission of assignments and private tutoring;
15. CMC courses with all content, audio/video, and software available online; students can

download and print out content or read it online.

### **Hypermedia on the World Wide Web**

These courses can be followed online on the Web. They take advantage of the links to other relevant sites using subject trails and other techniques. These courses can be either text-based, and so available to students with older computers and/or low bandwidth connections, or they can include graphics and animations that require a more powerful computer and higher bandwidth.

16. Courses that use hypertext links and have all necessary course materials online.

17. Courses with hypermedia links with multimedia using Java, Shockwave, Quicktime, or other applications.

Online courses need to be further divided into those for which the course materials can be accessed using a standard web browser (Netscape or Internet Explorer) and those that require students to download and/or install a client-side application. This can be quite troublesome for students. Some plug-ins like Shockwave and Quicktime are standard and might be considered as normally available. Others, however, are nonstandard and can be problematic.

Various permutations of the taxonomy described are possible. For example, any multimedia Web course might also have a text component. A CMC course could have some hypermedia links and a workbook. A correspondence course could incorporate some multimedia courseware on a CD-ROM with hyperlinks to the Web. Because of this, it is very difficult to provide links to real examples of these different course types as they are constantly changing.

In order to limit the size of a database and make it more useful to online students, this taxonomy limits the online course selection to those courses and programs that are universally available online--that is, those that can be taken from anywhere in the world where there is an Internet connection. If a student is in Johannesburg, Rio de Janeiro, or Kuala Lumpur and has a reliable web connection (and the price of tuition), he or she should be able to access every course in any user-friendly database. So, the TeleCampus directory includes only courses that can be taken on the Internet from anywhere with no residence requirements or need to attend sessions at a physical location.

### **Survey Methodology**

One of the initial confusions in setting up the survey methodology was whether the target population was composed of institutions, or of courses. After discussions with a statistician, it was decided that what was being measured was the number of online courses worldwide, independent of the number of institutions offering the courses. Thus, the initial, institution-based sample (consisting of every third course for each institution with five or more courses) was rejected, and a second sample of randomly chosen courses was created.

A representative sample of 789 courses was taken from 17,335 courses listed in the TeleCampus directory in early October 1999. First, courses from institutions with fewer than five courses were removed from the selection list. This was done in order to eliminate free courses, and those offered by persons or institutions whose main purpose was not education. It was ascertained that most of these courses belonged to institution entries with fewer than five courses. Out of a total of 894 institutions, this left 14,518 courses in the list, offered by 520 institutions.

It was determined that, for a population of over 5,000 items, an absolute minimum of 377 valid responses were needed (Krejcie and Morgan 1970). Given limited financial and human resources, it was decided to aim for a 60 percent return rate, and factor in a 10 percent error rate (to cover courses no longer offered, broken URLs, etc.). Therefore, it was decided to sample 800 courses. In fact, the final figure was only 789, due to errors in the totals reported by the database. These errors were the result of a system change--deleted courses were reported as active for a period of about one month (September 1999). The errors were not discovered until late in the selection of random courses. While efforts were made to correct this by adding more courses and selecting additional institutions, this was not always possible, depending on the number of institutions available in each category of institution (see institution types below).

To improve statistical relevance, it was decided to stratify the sample according to institution type. The distribution of sampled courses was as follows:

- \* Association: 15 courses (2%)
- \* Community college: 185 courses (23%)
- \* Graduate: 304 courses (39%)
- \* Other: 8 courses (1%)
- \* Private: 166 courses (21%)
- \* School: 23 courses (3%)
- \* Undergraduate: 87 courses (11%)

The majority of the questions included on the survey were mainly for camouflage, i.e. the most important information was the last question about enrollment figures for particular courses. The preceding questions were used to put the respondent at ease, based on evidence in the literature that difficult questions should come near the end of a questionnaire. Some useful data on pricing, and the average life span of courses, could be extracted from this data for future use.

The use of letterhead and a signature on each survey was designed to give the survey a professional yet personalized appearance. The decision to do a postal mail-out was based on previous contacts with institutions, whereby many appeared to prefer paper to electronic

contact--possibly because paper could be filed away and dealt with later. It was felt that a paper mail-out carried more credibility and reduced the chances of the survey being thrown away (e-mail is easy to delete).

The equipment required for this survey included spreadsheet programs for the two researchers (Paradox and Excel), computers, printers, fax machines, Word, WordPerfect, and a Web site for publishing results. All of these were readily available at both locations. However, it was decided to publish results on the Nova Southeastern Web site as well, because the researcher had left her position with TeleEducation NB before completing the project.

Institutions whose courses appeared in the random list were contacted and asked to complete a one-page survey and update contact information if required (see Appendices A, B, and C). The following formats were offered for the survey: generic PDF form on the TeleCampus Web site; Word, WordPerfect, RTF, ASCII, and PDF by e-mail; fax and prepaid envelope. Besides an intended positive effect on response rates, this effort was to determine preferred channels of communication for future surveys. Follow-up e-mails, faxes and telephone calls were planned, but the low response rate and changes in staffing made this impractical.

The goal of this exercise was to locate the right contact people at each institution for the next survey and to determine their willingness to participate in such a survey. The literature suggests that the more contacts that are made, the higher the response rate.

The use of TeleEducation NB letterhead (paper and envelopes) was intended to add prestige and an element of peer support to the survey. The literature is equivocal with regard to the impact of prepaid envelopes, but these were considered to be common courtesy and were likely to increase response rates.

It was anticipated that institutions would be extremely reluctant to divulge class sizes for online courses. As a result, the survey was anonymous. Originally, qualitative information was to be sought where quantitative data was not forthcoming; this is still a possibility, but it would have to be undertaken as a separate project. One device used to determine inconsistency in responses was the use of verification questions. For example, one respondent responded "no" to the question "Do you track enrollment for this course?", then provided a current enrollment figure. The reverse situation was more frequent, but both situations indicate that either the questions were unclear, or respondents reacted differently to similar questions phrased with greater or lesser specificity.

Green paper was used, based on literature stating that the use of colored paper could help response rates. Initially, the cover letter and survey form(s) were to be printed on double-sided paper. However, this created a difficulty with preserving anonymity. Instead, the cover letter was printed separately on letterhead, and survey forms were printed where possible on double-sided paper, although institutions with three courses received three separate sheets. There was a tradeoff between the maximum weight allowed for postage, and the time to do double-sided printing (this was done by professional staff, not clerical staff).

One unforeseen difficulty was with the use of a home printer, which ran out of ink late in the printing process. The last survey forms were therefore much paler and did not look as professional as those printed earlier. A pretest of the survey form was planned, but did not occur, mainly because the researcher was not on-site in Fredericton and did not have access to distance education experts who might have vetted the form.

Because of time restraints and the very high rate of non-response, follow-ups were not conducted to try to complete the minimal statistical requirements. Another area not pursued was a detailed analysis of the characteristics of non-respondents. This was deemed more appropriate for a separate study.

## Results

A total of 96 institutions were accounted for, out of the 299 institutions that were solicited for the survey. Of these, 89 completed the survey form, while seven surveys were returned to sender.

Responses were received for 171 courses out of 789 solicited (22% response rate). Of these, 60 percent (102) contained usable enrollment figures, while 40 percent (69) contained invalid or missing responses.

Out of 118 valid responses, seven responses of "zero" enrollment could not be calculated statistically. Therefore, with only 111 out of 789 possible responses, the valid response rate was only 14%. Because of the low response rate, certain suspicious responses were included that needed to be verified by telephone or e-mail. These involved figures rounded to the nearest "ten" (e.g. 10, 20, 30).

As the methodology required a minimum of 377 valid responses before any statistical manipulations could be considered, and only 27 percent of that target figure was reached, any statistical analyses of the data must be suspect. Also, given the very low response rate, planned follow-ups by fax, e-mail, telephone and mail were abandoned. In fact, given the wide range of enrollment figures retrieved (between 1 and 929), a much higher number of valid responses would have been required to compensate. It is probable that close to 100 percent of the courses would need to be surveyed in order for the results to be valid.

By country, 69 out of 225 US institutions responded for a total of 103 courses with valid responses. The return rate for Canadian institutions was much lower than expected, although proportional to the US return rate, and to the percentage of Canadian courses in the directory (18%). Thirteen out of 34 Canadian institutions responded with valid responses for 17 out of 103 possible courses. Seven out of 28 non-North American institutions replied with valid responses for four out of 75 possible courses.

Survey questions not dealing with enrollment figures were not compiled; they were included only as "camouflage." The enrollment question, because of its importance, was placed at the end of a series of routine questions to increase the chances that it would be answered. This strategy did not appear to work well--respondents who answered other questions still left the sensitive enrollment question unanswered.

No strong preferences for a particular delivery format were apparent--every format option was chosen by more than one respondent (i.e. PDF, RTF, ASCII, WordPerfect, Word, and regular mail). No respondent asked for fax, but this medium was used for returns. Familiarity and convenience appeared to be the overriding factors in respondents' choice of format. Although the researcher worked from photocopies of the returned surveys, and not the originals, none of the faded ink copies were returned, indicating that poor ink quality might be a determinant for nonresponse.

## **Enrollment Figures**

Although extreme caution must be used in extrapolating or generalizing from the data collected, the statistical information below can be taken as "a best guess", until a more thorough data collection project is undertaken. The number of respondents can be sufficient for a reasonable guess at the average number of students if the outliers are removed. The analyses below are based on removing the courses with 0 or 1 student and the two courses with the largest number of students.

The results show that the average online course is delivered to 33 students, whereas the median is only 19. The average community college course contains 27.4 students, whereas the average university course is significantly higher at 38.9 students. The medians of the two systems is similar (community colleges 19; universities 19.5). The extremely sparse data on online courses given by associations, private companies, and schools make any guessing about them specifically impossible. Their figures are, however, incorporated into the totals (See Appendix D).

## **Discussion**

The TeleCampus survey used many of the most effective survey techniques: green paper, commemorative stamps, hand signatures, prenotification preceding the survey (repeated contact). More contact could have been made through follow-up. University sponsorship would have been difficult to negotiate, and points to the need for partnerships with professional associations or a major research university in any future survey. When deciding on the length of the survey, the researcher examined only the length of a single form, not the total length of each tailored copy (up to 17 forms). Part of the difficulty in establishing the survey methodology was how to minimize the number of courses to a reasonable number that institutions would want to fill out. The form's contents are shown in Appendix C.

Examination of returns, and especially nonreturns, suggests that length may have determined whether or not a survey mailing was answered. However, the pattern is not clear. For example, eleven institutions were asked to respond to surveys about six courses. Of the two replies, one had no valid responses and the backsides of the three pages appeared to have been missed altogether; the other had three valid replies. Conversely, the highest number of courses for which a survey form was returned was thirteen, with twelve valid responses. Out of fourteen surveys covering double-digit courses (10 to 17), three responded, with valid responses for seven, two, and twelve courses.

The numbers are too small to reveal any significant patterns, but factors other than length

seem to be affecting some respondents. In general, the determinant length for returns versus nonreturns appears to have been between four and six courses. The higher the number of courses, the more likely the number of invalid responses, often because courses were inactive or no longer offered.

Positive points included the fact that only one institution requested that it be removed from the survey. While only a systematic survey of nonrespondents could determine all the reasons for nonresponse, it would appear that the tone of the cover letter was sufficiently professional to avoid "flames" and not cause serious offense to potential participants.

Those who were interested responded very quickly (within a week, regardless of format). This might be taken to mean that tailoring each survey to respondents' preferences paid off with rapid turnaround. Because the researcher was not located at the head office, it is not known whether prepaid envelopes were used.

There were some clear contradictions in answers--e.g., "Do you track enrollment for this course?" (Answer: yes), followed by "If yes, what is the current enrollment?" (Answer: Don't know); "Are you currently offering [this course]?" (Answer: No), followed later in the survey with the disclosure of an enrollment figure. Such contradictions point to confusion on the part of respondents, probably due to a lack of clarity in the survey design.

The electronic cover letter is viewable in Appendix A. A major omission involved the mail-out cover letter (See Appendix B). The wrong electronic version was used to print the paper copies, resulting in the omission of two introductory paragraphs explaining what TeleEducation NB did and why the survey was being conducted. Also, the use of a home printer resulted in poor ink quality on the final survey copies printed. The bulky nature of surveys to institutions with five or more courses selected was reflected in the very poor return rate for these institutions. There were difficulties in obtaining prepaid US postage (a trip to Maine was required). Prepaid postage resulted in wasted money, given the poor return rate. It was difficult to determine when to abandon electronic delivery, and conduct a postal mail-out--researchers waited two weeks, which was probably too long. Finally, logistical difficulties occurred because the survey was conducted at a distance from the head office, where support staff were available to receive faxes, postal mail, and telephone inquiries.

Much of the data is suspect--e.g., figures are rounded up to the nearest ten, with figures of 10, 20, 30, 40 students per course reported. The reasoning of respondents who chose to round off figures needs to be probed. Was this yet another manifestation of their unease in revealing sensitive information, or did they believe the figures to be close enough to be of use? Only a telephone interview would adequately determine the cause for this behavior, at least by revealing voice inflections, hesitations, or other signs of state of mind in responding to questions.

Secondly, much work remains to account for and eliminate sources of errors. For example, several respondents neglected to answer questions printed on the back sheets of the survey pages. Apparently they did not realize that sheets were double-sided. They would have to be contacted individually to confirm this hypothesis, and asked to complete the back sides.

The most serious impediment to using data gathered via this survey is proving its accuracy. One method for future surveys is to gather data via a third party agency, such as a reputable polling firm. Even channeling survey questions via an association could be precarious. The president of the Association of Community Colleges of Canada believes that member institutions are extremely secretive about their online course offerings and are unwilling to discuss difficulties or techniques with fellow members (personal conversation, October 1998). It can be speculated that, in this competitive environment, institutions need more than verbal assurances of anonymity to disclose sensitive data to a third party.

### **Nonvalid Results**

"Not offered" responses came from eleven institutions for one or more of their courses. This can be assumed to be due to prenotification for these courses on the institution's Web site. Indexers are instructed to index courses in advance of their start dates. Future surveys should search by start date and remove courses that have not yet commenced. However, institutions do not uniformly indicate which courses are under development, nor do indexers necessarily take the time to verify under each course description when indexing large sites.

"Don't know" appeared to be used as well to avoid answering the enrollment question--i.e., respondents answered affirmatively to the question of whether enrollment figures were kept, but answered "Don't know" to the next question, which requested an exact figure.

Fifteen institutions left the enrollment figure question blank, although they responded to other questions. This is the most serious nonresponse case, and it requires further investigation. There was no pattern by type of institution--almost every category of institution is represented in the nonresponse group.

Contrary to expectations, only three respondents chose "Don't know", while the remaining two nonvalid responses were from webmasters who used the number of server hits for the enrollment figures (one reported "20,000", the other "1000+").

Although responses to other questions were not compiled, it is worth noting that not a single respondent completed the question "Why was it discontinued? Check all which apply." This may be due to two reasons: first, almost all institutions that replied offered at least one course. It may be inferred that those to whom this question applied did not respond to the survey; second, those responding to the survey may not be in administrative positions that have access to sensitive data. This may provide collateral evidence for the low enrollment figure response rate. If respondents were not privy to administrative data, then the database of contact names needs to be revamped, and considerable effort expended to locate the appropriate respondents within each organization. This in turn implies a revamping of the survey approach, perhaps exchanging e-mail and paper versions by telephone contact via a reputable survey firm. Alternatively, focus groups organized by professional associations could provide contextual data and ideas for approaches that appeal to high-level administrators.

An estimate of two months' work is required to contact nonrespondents (including those who returned a survey with a missing or invalid enrollment figure). Nonrespondents represented

66 percent of potential respondents (196 out of 299).

## Conclusions

The response rate of this survey does not justify generalization of its results. However, it does provide an idea of the extremely wide range of enrollment sizes for online classes. In that sense, the survey may establish outer limits, which may or may not be confirmed by subsequent studies. It also can provide some sense of trouble areas, i.e., those areas for which quantitative analysis is impractical because ranges are too broad, response rates are too low, or numbers of respondents are too low, but which may be studied using qualitative methods or some other technique.

This study indicates that the current size of the database may be too small, and the range of enrollment sizes too broad, to determine any meaningful projection of course enrollment in the near future. Secondly, the prediction that institutions would be uncomfortable in divulging enrollment statistics, even with safeguards for anonymity and under the authority of a neutral third-party agency, appears likely. The high rate of item nonresponse and invalid item response for the enrollment question suggests that this information is viewed as being sensitive. More research is required to determine if this is in fact the case, and if other barriers to full disclosure also operate in the survey context.

As a result of the lack of comparable data, the results of this study must be treated with circumspection. They can be validated only after a second iteration, when comparisons with this year's data can be made, or by comparison with similar studies based on similar methodologies. Sources of suspect data, biases, and noise may be identifiable only after two or more surveys have been conducted and norms have been established in this area (e.g. course enrollments which are abnormally high). It may be that very few norms exist, and that the data will reveal few consistent patterns over time. Future studies may benefit by being much narrower in scope, or if broken down by sector, subject area, type of institution, or some other characteristic. Combining results from studies in different sectors may prove to be the most reliable method for obtaining online course enrollment figures on a regional, countrywide, or international basis.

The results indicate that the survey was too long and complicated for institutions with over four courses, resulting in frequent omissions and survey fatigue. The most remote institutions from Canada were the least likely to respond, making efforts to reach overseas participants especially vital in the future.

From an administrative viewpoint, the effort required to send out various electronic formats was considerable for correspondingly few results. The researcher considerably underestimated the amount of personal involvement required to reassure participants, explain the survey, resend earlier versions of software, etc. However, the literature clearly points to the need for even more investment in human resources for personalized follow-ups, particularly by telephone.

Based on previous studies, future surveys should extend the survey period to at least four months, and increase the number of staff involved to at least three full-time persons (one

researcher, an assistant for telephone follow-ups, and one clerical person to handle various e-mail and paper formats, as well as fax follow-ups).

A possible explanation for the poor response rate is the lack of cohesiveness among disparate online course providers. Their perception of themselves with respect to online education appears to vary enormously. The newness of the online education field, the fact that it is still often found at the periphery of traditional academia, the aggressive marketing of private firms in this market, as well as the blurring of lines between universities, colleges, and private companies all work against a coherent approach to surveying online educators. Further research should focus on whether a generic approach is possible, or if future surveys should be tailored to each sector.

The error rate found in the current data drawn from TeleCampus was found to be unacceptably high. Contact information for each institution had been completed by a summer student, and was accurate as of July 1999. Yet, the researcher and her assistant found it necessary to verify and correct addresses for a significant portion of these institutions. The problems appear to have arisen because of incomplete or absent address and contact name information on the distance education department Web sites of many institutions. Surveys sent to seven institutions covering thirteen courses were returned to sender. The cost involved in gathering this information was high in terms of employee time--it took two employees a month to obtain correct address and contact information for 400 institutions.

A considerable initial effort in obtaining reliable contact information is therefore required, followed by frequent updates. It was remarked that, in many instances, having the correct address did not suffice in reaching the correct person to answer the survey. Those who responded by e-mail appear to have forwarded the survey much more quickly and efficiently to the right individual. Examination of the trail of forwarding indicates that the person reading the initial e-mail did not know to whom to send the survey--sometimes as many as five people were contacted before someone agreed to respond to the survey.

Paraprofessional staff did not take time to transcribe address and contact information for institutions. This is not surprising, as they were paid on a per-course basis and were expected to index three courses per hour. Also, the database structure was not conducive to navigating between the course index form and the institutional entry form. Staff were further confused by the difference between "Contact name" within a course record and "Contact name" in the institutional record. The first was usually the course professor or the admissions office of the institution, while the second was often the director or communications officer for the distance education department.

Separating the maintenance of course and institution information would facilitate the maintenance greatly. The use of a marketing student to search out and enter a marketing contact for each institution was, in retrospect, a good approach. However, contact information is TeleCampus' most valuable asset. A summer student should not be solely responsible for determining the "best" contact person. For future data mining to take place, institutional information must be impeccable. The experience from this survey indicates that

the lack of an individual's name can result in a significant delay in locating the appropriate individual, or even in the rejection of the survey altogether.

### **Data Quality**

The principal difficulties with the survey were

1. reaching the appropriate respondents and
2. the inclusion of courses which had been discontinued, were not yet available, were temporarily deactivated, or, in some cases, had never existed.

After the first month, there were indications of the extent of faulty course data, the survey was altered to make reference to only one data field in the database, course title. Contact name and address information taken from the data was carefully verified and corrected.

However, the poor return rate indicates that even information transcribed from institutional Web sites is not sufficiently reliable for the undertaking of a massive survey. It is recommended that print directories, association membership directories, and telephone verification take place before undertaking a similar venture. The time spent in establishing a contact person could be recompensed by a significantly higher return rate.

This study sheds some light on the enormous data maintenance problems inherent in managing a database of the magnitude of the TeleCampus Online Course Directory. Training for paraprofessional staff is important. With an error rate well above 10%, it is clear that TeleCampus can benefit from the model of human-indexed systems, such as Yahoo, by investing in staff training. Technical solutions, such as spellcheckers and URL checkers, do not address issues of currency of information. Furthermore, remote agents to glean metadata from participating institutions are not yet available.

Even when some form of automatic indexing appears, it will not resolve the most sensitive issues involved in indexing online information. These include human error in copying verbatim from Web sites, the desire of many institutions to remove selected information from the database (e.g. courses whose release date has been delayed), and fluctuations in pricing.

Data quality issues can be resolved over the long term by assigning human resources to verify, update, and maintain institutional contact information, in addition to course metadata. Research demonstrates that poor data quality can have a major negative impact on databases (Redman 1998).

As noted, the sparse figures provided by a small subset of the original sampling cannot be used for scientific analyses or extrapolation. However, using them for a reasonable best guess, one can surmise that the large difference between the average (33) and the median (19) suggests a wide disparity in student numbers among online courses. Many courses are given to very small numbers of students, whereas other courses have very large enrollments. Assuming that the TeleCampus Online Course Database contains at least 60

percent of the total number of courses available online, a reasonable extrapolated guess might put the total number of students presently studying online at a minimum of 935 000. This would be based on an average of 33 students in  $17,000 / 60\% = 28,333$  courses. Please note again that this is a best guess figure rather than a scientific analysis. There are more than 250,000 students taking courses from SmartForce Mentoring alone, so the issue becomes very problematic.

The U.S. Department of Education (US Department of Education National Center for Education Statistics 1999) estimates that there are approximately 1.7 million students in post-secondary distance education in the USA, of which about 60% are enrolled in asynchronous classes. If we assume that these are primarily online courses, then the figures suggest about 1 million students study online in the USA. Adding another 25% for the rest of the world and for other types of courses make 1.25 million. In comparing these figures with those of the TeleCampus survey, a reasonable guess is that between 935,000 and 1.25 million students studied online in the fall of 1999.

## References

Farance, F., and K. McCabe. 1999. New international subcommittee advances IEEE standards work on learning technology. Available online at <http://standards.ieee.org/announcements/newintsub.html>. Updated November 22, 1999. Accessed December 3, 1999. E-mail: frank@farance.com & k.mccabe@ieee.org Author's Web Site: <http://standards.ieee.org/>

Krejcie, R. V., and D. W. Morgan. 1970. Determining sample size for research activities. *Educational and Psychological Measurement* 30:607-610.

McGreal, R. 2000. TeleCampus Online Course Directory. In *Perspectives in web course management* B. Mann, ed. Canadian Scholar's Press.

Pourjalali, H. and J. Kimbrell. 1994. Effects of four instrumental variables on survey response. *Psychological Reports* 75:895-898.

Redman, T. C. February 1998. The Impact of poor data quality on the typical enterprise. *Communications of the ACM*. 41(2), 79-82.

Tremblay, E. March 31, 2000. *Distance education database evaluation: Final report*. Fredericton: TeleEducation NB. Available online at <http://teleeducation.nb.ca/media/reports.shtml>.

US Department of Education National Center for Education Statistics. (1999, December). Distance education at post-secondary education institutions 1997-1998 [NCES 2000-013]. Washington, DC: Office of Educational Research and Improvement, by Laurie Lewis, Kyle Snow, Elizabeth Farris, Douglas Levin, and Bernie Green, project officer. Available online at <http://nces.ed.gov>.

+++++

## Appendix A. Cover Letter for Electronic Survey

Dear ~2~,

You may not be aware that your institution's online courses are listed with the TeleCampus Online Course Directory (<http://courses.telecampus.edu>). A non-profit government-funded service of the New Brunswick Department of Education in Canada, the directory links students to your courses or website directly. It contains nearly 18,000 courses from 30 countries in 10 languages and is free of charge.

To help all of us determine the growing needs in the field of distance education, we are conducting a short survey of student enrollment worldwide. One or more of your courses was chosen at random from the TeleCampus Directory. We would greatly appreciate it if you would respond to ten (10) brief questions regarding enrollment for each course.

The survey will be mailed to you next week and may be returned by fax or pre-addressed envelope. It will also be available in PDF format on our website. If you would prefer to answer by email, please reply to this message, indicating your preferred format below:

1. ASCII (body of message)
2. PDF
3. Word
4. WordPerfect
5. Rich Text Format
6. Other (please specify)

The results will be posted on our website by early December. Many thanks in advance for your participation,

Julia Innes

Database Librarian

TeleEducation NB

500 Beaverbrook Court

Fredericton, NB E3B 5H1

[julia@teleeducation.nb.ca](mailto:julia@teleeducation.nb.ca)

(506) 444-4234

Fax: (506) 444-4232

+++++

## Appendix B. Cover Letter for Postal Survey

October 14, 1999

Dear '1<sup>a</sup>,

Thank you for taking the time out of your busy day to respond to our survey.

The course(s) listed below were chosen at random from the TeleCampus Directory. We would very much appreciate it if you could complete and return the survey by October 31st, 1999. Responses received after this date will be incorporated into the study on an ongoing basis.

Please note that this information will be held in the strictest of confidence and that all information for public viewing regarding the survey will be anonymous. Of course, we will provide you with the results of the survey at your request, and this information will also be available on-line at <http://courses.telecampus.edu>.

Also, to better serve you and your students, we have included your institution's information below. If there are corrections, we would greatly appreciate your most current information. Please correct as necessary:

Contact name: '1<sup>a</sup>

Institution Name: '2<sup>a</sup>

Address: '3<sup>a</sup>

City: '4<sup>a</sup>

State/Province: '5<sup>a</sup>

Zip/postal code: '6<sup>a</sup>

Telephone: '7<sup>a</sup>

Fax: '8<sup>a</sup>

Email: '9<sup>a</sup>

Once again, we genuinely appreciate your assistance,

Julia Innes

Database Librarian

TeleEducation NB

500 Beaverbrook Court

Fredericton, NB E3B 5H1

Julia@teleeducation.nb.ca

(506) 444-4234

Fax: (506) 444-4232

+++++

**Appendix C. Survey (Word Format)**

SECTION A

COURSE TITLE: \_\_\_\_\_

When was this course first offered?

19\_\_\_\_\_

Winter\_\_\_\_\_

Summer\_\_\_\_\_

Fall\_\_\_\_\_

Month\_\_\_\_\_

Don't know\_\_\_\_\_

What is the highest fee you charge for this course? (Some institutions charge more for international or out-of-state students)

(\$)\_\_\_\_\_ Currency: \_\_\_\_\_

Free\_\_\_\_\_

Don't know \_\_\_\_\_

Are you currently offering it?

Y\_\_\_\_\_

N\_\_\_\_\_

Don't know \_\_\_\_

If yes, how long has it been offered? (Please fill in one)

\_\_\_\_\_ months

\_\_\_\_\_ years

\_\_\_\_\_ semesters

\_\_\_\_\_ Don't know

If no, when was this course discontinued?

19\_\_\_\_

Winter\_\_\_\_

Summer\_\_\_\_

Fall\_\_\_\_

Month\_\_\_\_\_

Don't know\_\_\_\_

Why was it discontinued? Check all which apply:

a) insufficient enrollment

b. b) professor no longer available

c. c) not cost-effective

d) material outdated

d. e) material reorganized into new course(s)

e. f) don't know

g) other (please specify): \_\_\_\_\_

Do you intend to offer it again?

Y\_\_\_\_

N\_\_\_\_

Available continuously\_\_\_\_

Don't know\_\_\_\_

NOTE: For discontinued courses, go directly to Section B.

### FREE COURSES

Would you be able to estimate the number of students who make use of this material on your site at the present time?

Number of students \_\_\_\_\_

Don't know \_\_\_\_\_

Not being offered \_\_\_\_\_

NOTE: Go directly to Section B.

### FOR-FEE COURSES

Do you track enrollment for this course?

Y \_\_\_\_\_

N \_\_\_\_\_

Don't know \_\_\_\_\_

If yes, what is the current enrollment?

\_\_\_\_\_

Don't know \_\_\_\_\_

### SECTION B

Please feel free to make additional comments on the reverse of this page.

Would you be willing to be contacted regarding your institution's involvement in offering online courses?

Y \_\_\_\_\_

N \_\_\_\_\_

If yes, what is your preferred method of contact?

Telephone \_\_\_\_\_

Email \_\_\_\_\_

If you would like to be contacted when the results are posted, please fill in your email

address here: \_\_\_\_\_

PLEASE RETURN TO: Julia Innes, Julia@teleeducation.nb.ca; Fax: (506) 444-4232.

Thank you very much for you time! The results will be posted on  
http://courses.telecampus.edu.

+++++

**Appendix D. Data Analysis (Best Guess)**

Total Students for Community Colleges: 1235

Average: 27.44

Median: 19

Total Students for Universities: 1791

Average: 38.93

Median: 19.5

Total Students Overall: 3499

Average: 33.01

Median: 19

[Top of Page](#)