Challenges for Conducting a Computing Sciences National Study

Marguerite S. Giguette¹ and Karen C. Davis²

Abstract – Conducting a national study in the computing sciences is a process that requires detailed planning and a consistent effort by researchers to accomplish the study’s goals. This paper discusses the many challenges that were addressed when organizing and implementing a large study involving 50 colleges and universities from across the United States. Issues affecting both the coordinating institution, Xavier University of LA, and one of the participating institutions, the University of Cincinnati, are discussed. The issues include the survey administration details, confidentiality and security of the survey data, evaluation of survey software packages, recruitment and retention of the participating institutions, the processes for Institutional Review Board approvals, and recruitment and compensation of students. The lessons that were learned by the researchers at Xavier are organized into a suggested methodology for conducting surveys of this scope.

Index Terms – Culture, Ethnicity, Gender, National Study

INTRODUCTION

In 2003, Xavier University of Louisiana was awarded a research grant from the National Science Foundation to study gender, ethnic, and cultural differences in the computing sciences. This interdisciplinary research project includes an extensive nationwide study of computing sciences and non-computing sciences students from 50 institutions across the United States.

This paper discusses the process required to conduct such an extensive study. From Xavier’s perspective, the process includes the recruitment and retention of faculty at each of the institutions who serve as liaisons for the project, Institutional Review Board approval at each of the institutions, and solving the survey implementation issues that simplify the administration of the survey at the various institutions. In addition, the faculty at each of the participating institutions must also understand their institution’s Institutional Review Board process and obtain the necessary approvals, recruit students for the study, and coordinate how the students who complete the survey will be compensated. The effort required to accomplish these tasks is not trivial, and moreover, the nature of the tasks is not typically part of the training undertaken by a computing sciences researcher. An earlier multi-institutional study by McCracken et al. [5] focused on assessing programming competency after completion of an introductory programming course. In terms of implementing the study, their results identified the difficulties of differing student backgrounds, the importance of having a coordinator for the whole project, and the importance of conducting a trial run of the study. However, McCracken et al. do not consider the recruitment, planning, and contact management with other institutions, recruitment and compensation for students, and Institutional Review Board issues. They also had significantly fewer participating institutions.

Since students from various disciplines will be surveyed, it is important that there be interdisciplinary participation in all phases of the project. The project team includes principal investigators from computer science and psychology, and research partners and advisory board members from computer science, psychology, sociology, and education. The interdisciplinary team effort is unlike previous studies since both computer scientists and social scientists are working together in every phase of the project including the planning, development, and implementation of the survey as well as the analysis of the data. In addition, previous research studies focusing on gender issues in the computing sciences typically use smaller samples of students from one to five institutions. No study as extensive as the one described here has ever been conducted in the computing area.

ISSUES RELATED TO INITIATING THE SURVEY PROCESS

The main purpose of the research project is to examine issues related to gender, ethnicity, and culture in the computing sciences. The study’s research goals are to: “(1) Scientifically investigate by gender, ethnicity, and culture the under-representation of women in Computer Science (CS), Computer Information Systems (CIS), and Computer Engineering (CE) – collectively referred to as the computing sciences, (2) Increase the knowledge-base of gender, ethnic, and cultural models, creating newer models that lead to a more equitable and inviting CS, CIS, or CE education, and (3) Increase national capacity in the field of gender diversity and education in the computing sciences by encouraging new researchers and research-oriented education practitioners [3].”

Research has compared the gender differences at Historically Black Colleges and Universities (HBCUs) versus those at non-HBCUs [4]. Lopez and Schulte found that there is an extremely small difference between the number of African American women and African American men who

¹ Marguerite S. Giguette, Xavier University of Louisiana, Computer Sciences and Computer Engineering Department, New Orleans, LA, mgiguette@xula.edu
² Karen C. Davis, University of Cincinnati, Electrical & Computer Engineering and Computer Science Department, Cincinnati, OH, karen.davis@uc.edu

0-7803-9077-6/05/$20.00 © 2005 IEEE. October 19 – 22, 2005, Indianapolis, IN 35th ASEE/IEEE Frontiers in Education Conference T3D-13
receive BS degrees in computer science overall; however, more African American women receive BS degrees in computer science at HBCUs than their male counterparts [4]. This data analysis by Lopez and Schulte leads to questions that go beyond the sole issue of gender and include ethnicity and culture. Some of these questions are: Why are there proportionally more African American women than men studying computing sciences disciplines than their white counterparts? Why does the small number of HBCUs contribute a large percentage of the BS degrees in computer science to African American women? What is it about the culture of the African American community that attracts African American women in the computing sciences in greater proportion than white females? In addition, since the study is large and involves students from various computing and other disciplines, it has the potential to validate as well as expand conclusions made in previous studies.

The major theoretical focus is the Social Cognitive Career Theory (SCCT). SCCT attempts to explain the process by which people develop interests in a certain area, make choices based on these interests, and perform at various levels both academically and in their occupation [2]. One of the tenets of SCCT is that if people believe that they are effective in accomplishing tasks for a particular interest, then they will pursue that particular interest [2]. Previous studies [1, 2] that use SCCT as their theoretical basis focus on engineering students. These studies are limited to a much smaller number of schools and were not national in scope.

I. Institutional and Student Composition

In order to accomplish the project’s goals, a three-year survey of students from various disciplines and institutional types (e.g., research vs. teaching, historically black (HBCU) vs. non-historically black, etc.) is conducted at 50 institutions of higher education from across the United States. Of the 50 institutions, 25 are HBCUs and 25 are not HBCUs.

At each institution, 100 students are surveyed: 70 from the computing sciences and 30 non-computing sciences students. The computing sciences students are a part of the longitudinal portion of the study, and will therefore be contacted in the future years of the study.

In order to facilitate the implementation of the survey, two faculty members at each of the institutions, one from the computing sciences and the other from another discipline, have been recruited to serve as Higher Education Partners. In order to facilitate a smooth implementation at each institution, it is also an advantage if one of the Higher Education Partners has previous experience with their institution’s Institutional Review Board (Human Subjects Committee).

Recruiting faculty from all 50 institutions was a long and detailed process. Historically Black Colleges and Universities are concentrated in the southeastern portion of the United States. A Jeffersonian apportionment method, which was originally used from 1790-1830 to allocate the number of representatives to the U.S. House of Representatives from each state [6], was used to select the number of HBCUs from each of the 15 states with HBCUs. For each of these states, a non-HBCU institution was selected. The remaining 10 non-HBCU institutions were selected to ensure representation from various geographic regions of the United States.

II. Challenges When Recruiting and Engaging Higher Education Partners

The Xavier researchers made a list of possible institutions based on the above distribution. In some cases, the researchers had made previous contact with potential Higher Education Partners; however, in most cases, institution websites were visited in order to locate faculty contacts. The computing sciences faculty members were contacted first since these faculty serve as the main contact for the study at each institution. Therefore, wherever possible, online vitae were examined to see if faculty had an interest in research similar to that of the study or if faculty were interested in women’s issues in the computing sciences. All the potential Higher Education Partners were then contacted individually by email.

The email communication explained the basic goals of the study and gave an overall explanation of the study’s various components. In addition, the email message requested that the faculty member identify and recruit an additional faculty member outside of the computing sciences to serve as the institution’s second Higher Education Partner.

The initial communication resulted in many responses that had to be answered individually since many of the follow-up questions were typically specific questions that proved extremely time-consuming to answer. In addition, it was important to remind the faculty contacts that a second faculty contact was needed.

As time progressed, some of the schools that were initially contacted declined the invitation to participate. Consequently, the researchers continued to look at the list of possible institutions and selected other schools to invite; the process of contacting these schools began as before. Many people did not respond at all to the initial email which proved difficult since it was not clear why there was no response; some of the possible causes were: Did the faculty member not receive the email? Were they not interested? Were they checking with colleagues? These faculty members were sent email once again specifically asking for a “yes” or “no” response. If no response was received, a new institution was selected from that particular state. At times, phone calls were used when more clarification was needed. The process continued for a number of months.

A database was created to store contact and institution information. The database was critical in assisting the Xavier researchers in keeping track of which institutions and individuals were contacted. Because of the large number of people who were contacted, copies of all email correspondence were also kept to ensure that previous communication could be reexamined as the project progressed.

It was learned early in the process that once a commitment was made by the faculty contact, it is important to keep the Higher Education Partners informed as much as
possible. This had the benefits of keeping the Partners interested in the project and also accomplishing the third goal of the study which is to “increase national capacity in the field of gender diversity and education in the computing sciences by encouraging new researchers and research-oriented education practitioners [3].” It is important to balance the amount of communication so that the Higher Education Partners remain engaged in the process, yet not overwhelmed by constant communication.

As mentioned previously, the Higher Education Partners from the computing sciences were contacted first. Once a computing sciences faculty member made a commitment, he/she was asked to identify a colleague from a non-computing sciences discipline to assist in the project. This proved to be simple in some cases, but more difficult in others. At some institutions, it was difficult for the Higher Education Partner to find an additional interested faculty member. At other institutions, the computing sciences faculty member would send the name of a potential non-computing sciences Higher Education Partner for the Xavier researchers to contact. In the latter case, the Xavier researchers would then use email to invite the potential Higher Education Partner to participate in the project; some replied quickly, while others had to be contacted a number of times.

As the time approached for the study to begin, all the previously committed Higher Education Partners were contacted to ensure their continued participation. Most of the Partners replied; however, some had to be contacted by email and eventually by telephone.

In order to recognize the contribution and work of the Higher Education Partners, their names and institution affiliations are listed on the study’s website, http://webusers.xula.edu/tlopez/PROACT/GED.htm; the names were listed after having received permission from the Partners. The Higher Education Partners also receive a small honorarium for their participation. In addition, the Partners will be given access to the collected data.

The most important aspect of recruiting and retaining the Higher Education Partners is to make sure that all aspects of the research project go as smoothly as possible. It is the responsibility of the Xavier researchers to assist the Higher Education Partners in all aspects of the study.

**Survey Implementation Challenges**

A research project of this size and scope presents a number of implementation issues that must be addressed. These include the acquisition of Institutional Review Board approval at the various participating institutions, the administration of the survey instrument, and the payment of survey participants.

I. Institutional Review Board Approvals

Since the study involves human subjects, approval of the appropriate committees on each of the participating campuses must be obtained, as required by the Department of Health and Human Services Regulations for Protection of Human Subjects found in Title 45 of the Code of Federal Regulations, Part 46 (45 CFR 46) [7]. The process requires either approval of the institution’s Institutional Review Board (IRB) or Human Subjects Committee. Failure to comply with the federal law can result in loss of funding, closing of research programs, and sanctions against the individual faculty members involved, including sizable fines.

Each institution’s IRB documentation requirements are similar, but all vary slightly. Required information includes the survey instruments, informed consent forms, recruitment flyers, and information about the implementation of the study and the confidentiality of the data.

The Xavier investigators have relied on the Higher Education Partners to seek IRB approval at their institution, and have attempted to make the IRB process as easy as possible at each of the participating institutions. The Xavier researchers sent copies of all of Xavier’s IRB materials and IRB approval to each of the Higher Education Partners. Since Xavier had already received IRB approval, an expedited review of the materials by each IRB was possible and made at many of the participating institutions. In addition, whenever needed, the Xavier investigators have responded to specific questions raised by the IRB at various institutions through the Higher Education Partners.

II. Administration of the Survey Instrument

Traditional survey instruments are administered on paper to a group of subjects. The response forms are then scanned and analyzed. Since this study involves a large number of participants at 50 different institutions, it was important that the implementation logistics of the research be streamlined as much as possible. In addition, the cost in both money and time of creating and shipping the survey instrument was also considered. In order to address these concerns, a web-based survey is used. The administration of the survey thus becomes easier for both the Xavier researchers and also for the Higher Education Partners.

The use of a web-based survey better ensures the integrity of the data since answer sheets do not have to be scanned and open-ended questions are entered into the database as typed on the web by the survey participants. The web-based survey also simplifies the work of the Higher Education Partners. This type of implementation does not require each institution to schedule and coordinate the actual administration of the survey. The Higher Education Partners can now recruit students, explain and have the students sign the informed consent, and provide student participants with the survey URL. The students can then complete the survey on their own time.

It is important that the software used for the web survey be able to handle a variety of question types, including single selection, multiple selection, short answer, and open-ended. The ability to restrict access to the survey was also important. An additional needed feature was the inclusion of automatic email follow-up in order to contact students for the longitudinal portion of the study. With the variety of required features, the researchers decided to select a commercial product. After an extensive review, Perseus SurveySolutions Enterprise (http://www.perseus.com) was chosen.
To insure the security of the data, a web-based implementation required that Xavier install a secure Apache web server. In addition, a random student number is generated for each student participant; a student receives one of these numbers when he/she signs the informed consent form. The number received by the student ensures that only those students who have completed the informed consent form participate in the survey; it also ensures that each student completes the survey only once.

The web-based survey provides the Xavier researchers with the capability to view the data as it is submitted; thus providing them with the opportunity to monitor the number of participants at each institution as the survey progresses. This helps the researchers identify each institution’s success in recruiting students. If potential problems are discovered, then the Higher Education Partners are contacted.

III. Paying the Participants

Well before the survey began, Xavier sent a check to each of the computing sciences Higher Education Partners in order to pay each of the survey participants a $10 stipend. It is important to determine exactly how the payee should be written; if it is not correct, there may be problems for the fiscal office at some of the institutions.

In addition, each of the computing sciences Higher Education Partners should be informed that he/she should work with his/her institution’s fiscal office to determine how the students should be paid. In some cases, the fiscal office requires that the payments be dispensed at the bursar’s office while others allow the Higher Education Partner to disperse the stipend as long as he/she maintains proper receipts.

Once again, it was important for the Xavier researchers to make this aspect of the survey implementation go as smoothly as possible. When a student submits the survey, he/she will receive an email message that contains an unsigned voucher. The voucher indicates that the survey has been submitted. The students are instructed to take the voucher to the Higher Education Partner for further instructions. At this point, the Higher Education Partner signs the voucher and informs the student how he/she can receive payment.

PARTICIPATING INSTITUTION PERSPECTIVE

One of the participating institutions, the University of Cincinnati, has two Higher Education Partners. One is a faculty member in the Electrical & Computer Engineering and Computer Science (ECECS) Department and the other is a faculty member in the Sociology Department. Experiences as a participating institution are described below.

I. IRB Approval

For most computing science researchers, applying for IRB approval is an infrequent experience, so an overview of the documentation is given here in order to illustrate the process. The University of Cincinnati has two Institutional Review Boards for research involving human subjects; one is for medical research and the other is for social and behavioral science research (IRB-S). This project falls under the latter category and required the following documentation:

1. An application form (5 pages) with a one page appendix for each investigator (2 pages),
2. A protocol checklist (9 pages),
3. Data collection materials: two surveys (36 and 25 pages, respectively, provided by Xavier and used without modification),
4. Recruitment materials: flyers (2 pages),
5. Informed consent document (1 page),
6. A copy of the federal grant application which was provided by Xavier (17 pages), and
7. References (5 pages).

Five copies of over 100 pages of documentation were submitted to the IRB-S. One month later, email was received that denied the request because 14 discrepancies were discovered. As a result, the University of Cincinnati was unable to participate in the study pilot conducted in the spring of 2004. Corrections were accepted via email that enabled the University of Cincinnati to participate in the study starting in the fall of 2004. IRB requirements vary greatly from institution to institution; the UC experience appears to be at the more difficult end of the spectrum and therefore illustrates maximal challenges that a CS researcher attempting to initiate such a study might encounter.

In addition to the sheer volume of paperwork, there were some unanticipated complications that warrant mentioning here. These details may not seem unusual to social and behavioral science researchers, but they may be unexpected to a computing sciences researcher who is unaccustomed to procedural details required for performing research with human subjects.

1. A subject must be allowed to quit at any time, for any reason. This has implications for the online implementation of the survey; should a student be allowed to submit a blank survey and still receive compensation for it? This was ultimately resolved by emailing vouchers for completed surveys only. A student may opt to quit at any time, but compensation is only given to students who complete the survey.
2. Subjects must have contact information for local persons to talk to if they have any discomfort or concerns arising from participation in the survey. Contact information for both a computing sciences and non-computing sciences faculty member are included in the informed consent form; each student retains a copy of this form after submitting the signed version.
3. The consent form must be prepared at the appropriate reading level (12th grade education in this case). Fortunately, Microsoft Word has a built-in tool that assesses reading level and indicates where wording is not appropriate.
4. Preserving anonymity of subjects while providing compensation to them and performing longitudinal studies is difficult. The approach used here is to provide a unique login to a student for a secure website when he or she returns the signed consent form; no mapping of user...
logins to consent forms is kept, so there is no connection between a student’s survey and their name at the University of Cincinnati. At Xavier, there will be a record of the student’s email address and their survey because a voucher is sent via email when the survey is submitted. In addition, since the computing sciences students will be a part of the longitudinal aspect of the three-year study, these students will also be asked for a cell phone number. This information is requested to assist in contacting these students in future years. Only two Xavier researchers have access to the information and the records are kept in a locked filing cabinet.

II. Recruiting Students

Students were recruited by the Higher Education Partners by making personal appearances in first and second year classes taken by computing sciences majors and non-majors. In particular, the ECECS Department offers about 8 sections of its introductory course each fall, and there are numerous sections of Introduction to Sociology that were targeted. In the initial class meeting, informed consent forms were distributed, followed by handing out survey login IDs when signed forms were returned. In a follow-up class meeting, the faculty member returned to sign vouchers for students who had completed the survey. This seemed to increase student participation.

III. Working with the Local Fiscal Office

Some institutions allow one check to be sent to the bursar’s office from Xavier and students may pick up their cash or individual checks from that office. The University of Cincinnati is unfortunately not one of them (it has no central office to perform this kind of service.) The ECECS Department’s business manager dispensed compensation. She had a locked drawer with 100 $10 bills, and students claimed them by submitting an email voucher signed by one of the faculty investigators. One option to avoid the necessity of so much cash in one person’s office is to buy bookstore gift certificates, but it was thought that cash would be more appealing to students. The solution still involves a lot of student traffic for both the faculty and the staff, but it is difficult to avoid.

SUGGESTED METHODOLOGY

The process used in this research project has fostered the development of a methodology that should be used when conducting a national study. Key components of the process are outlined in Figure 1 and are discussed in detail in this section.

The smooth implementation of a national study requires that the researchers make a number of survey administration decisions at least 12 to 18 months before the study begins. The questions that must be addressed are:

1. How many and what institutions are needed for the study?
2. Will the study be longitudinal?
3. Will the survey be taken on paper or will it be web-based?
4. Will the participants be paid?
5. Will the contacts be paid?
6. How will the data be secured?

If the researchers decide that the survey will be web-based, then survey software tools must be evaluated at least one year before the study begins. Researchers can choose a commercial product or write the necessary package themselves; trade-offs between these two options must be weighed when making this decision. If the commercial product option is chosen, then various packages must be evaluated. The evaluation must ensure that the software can accommodate the different types of survey questions that will be used, e.g., single selection, multiple selection, short answer, and open-ended. In addition, if the research involves a longitudinal study, then it is critical that the software have an email follow-up tool that can be used.

A web-based survey requires that researchers have data security measures in place as well as develop a mechanism to distribute the URL of the survey. To ensure that the survey is only accessed by participants who have given their informed consent, survey logins are required and must also be distributed.

Figure 1. Overview of Survey Methodology

As discussed earlier, researchers are required to obtain approval from the institution’s Institutional Review Board (IRB). The requirements of the IRB vary from institution to institution and it is imperative that the researchers work with their school’s IRB chair to ensure that that process goes smoothly. This step should begin at least 12 months before.
the study begins. Typical documents needed by an IRB are: general study implementation information, the survey instrument(s), the informed consent form(s), the recruitment flyers, and data confidentiality information.

At least 6 to 12 months before the study begins, researchers should begin recruiting individual contacts at each participating institution. Recruiting individual contacts is very time-consuming and requires periodic email and telephone follow-up since many of those contacted do not respond in a timely manner. In the situation when a school declines to participate, it is important to have a list ready with additional schools and individuals that can be requested to participate.

During the recruitment of schools and contacts, it is also important develop and maintain a database that stores information about the institutions and contacts. Information that should be stored includes addresses, phone numbers, and email addresses of the faculty contact person as well as institution characteristics specific to the study, e.g., HBCU or non-HBCU.

IRB information should be sent to each participating school at least 6 months prior to the start of the study. It is important to emphasize to each of the institutions’ contact persons that the IRB approval process should begin as soon as possible. This will ensure that the beginning date of the survey is not delayed.

As the start of the study approaches, it is vital that each institution’s contact person is kept informed of the progress of the study and any changes in the process. In addition, these contacts should be recognized in some way, such as a listing on a web site.

If the study requires that the participants are paid, each institution’s contact must be aware of this and encouraged to make the necessary arrangements on their campuses for these payments. Each school’s process will be different and it is important for the researchers to remain flexible when paying the participants; however, the researchers must also ensure that appropriate payment records are retained for reporting purposes. Ideally, if the host institution sends a single check to each participating institution, then participants can be compensated using funds from that check.

During the study, the researchers must monitor the web-based implementation to make sure there are no problems and quickly respond to questions from the various participating institutions.

At the study’s conclusion, the data must be formatted for statistical analysis. If the institutional contacts are paid an honorarium, then arrangements must be made for payment.

The methodology needed to conduct a national study requires detailed planning and consistent communication with the participating institutions. This process should begin at least 18 months before the study begins.

CONCLUSIONS
This paper discusses the implementation of a national study that focuses on gender, ethnic, and cultural issues in the computing sciences and ensuing issues and difficulties. Since computer scientists are not typically trained in research involving human subjects across multiple institutions, the necessary implementation details often prove to be time-consuming and nontrivial. Conducting a national study of this nature requires development of processes that consider survey administration details, confidentiality and security of the survey data, evaluation of survey software packages, recruitment and retention of the participating institutions, Institutional Review Board approvals, recruitment and compensation of students, and survey deployment and data collection. A suggested methodology to conduct a national study in the computing sciences was presented.

ACKNOWLEDGMENTS
This material is based upon work supported in part by the National Science Foundation under Grant No. HRD-0332790, Microsoft Inc., and Apogen Technologies, Inc. Any opinions, findings, and conclusions or recommendations expressed herein are those of the authors and do not necessarily reflect the views of the National Science Foundation, Microsoft, or Apogen Technologies.

The authors would like to thank Antonio M. Lopez, Jr., Lisa J. Schulte, and Sarah Beth Estes for their leadership and collaboration in the study.

REFERENCES