The development of an online adaptive questionnaire for health education in Taiwan

Chien Chou*, Yi-Fan Chang, Yi-Ying Jiang

Institute of Communication Studies, National Chiao Tung University, Hsinchu, Taiwan

Received 1 July 1999; accepted 4 March 2000

Abstract

This paper explores the feasibility and usability of using computer networks, in particular the World Wide Web, as educational survey research instruments to collect students' opinions on proposed health education courseware. Also covered are the review of computer-assisted survey techniques, and the comparisons among different survey methods. The development and the evaluation of a Web-based adaptive questionnaire for health education are presented. Research issues are also discussed. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Internet; World Wide Web; Networked survey; Adaptive questionnaire; Health education

1. Introduction

Among the various social science research methods, surveys are perhaps the most frequently used to collect information from remote and distributed subjects (Babbie, 1992). Lavrakas (1993) states that survey methods are a collection of techniques of which the most typical purpose is to provide precise estimates of the prevalence of some variables of interest. Gay (1981) claims surveys are one of the ‘self-report’ approaches in educational research to collect standardized, quantifiable information from all members of a population or sample. In an educational context, surveys can be used to collect information from school administrators, teachers, students, or parents about their opinions on or reactions to a variety of topics such as

* Corresponding author. Tel.: +886-3-5731808; fax: +886-3-5727143.
E-mail address: cchou@cc.nctu.edu.tw (C. Chou).
Surveys can be distributed through a variety of communication technologies. More traditional methods include using regular mail, telephones, and fax machines. As computer technology improves rapidly, however, the computer-assisted survey information collection (CASIC) approach is becoming more feasible and popular (Couper & Nicholls, 1996). This approach includes disk-by-mail (DBM), computer assisted personal interviewing (CAPI), computer assisted telephone interviewing (CATI), and so on. Along with the emergence and continued growth of computer networking technology, additional distribution methods are now available for educational researchers. Computer networks can distribute survey questionnaires via four major methods: e-mail, news systems, bulletin board systems (BBSs) and the World Wide Web (WWW or ‘the Web’) (Chou, 1997). With so many survey distribution approaches and options, how can educational researchers choose the best method(s) to employ?

In order to soon develop an online health education curriculum at National Chiao Tung University (NCTU) in northern Taiwan, a survey was conducted to collect students’ opinions on health education, network usage, and the proposed online delivery. In searching for a feasible and reliable survey distribution method to collect students’ opinions, both traditional and computer-assisted methods were reviewed and compared. Each method has its own advantages and disadvantages. Among the many network applications, the Web is one of the most popular because of its user-friendly interface, equal and instant access to available information resources, and use across different computer platforms. Therefore, a Web-based survey questionnaire was conducted. Moreover, in order to further make full use of the powerful computer capabilities of branching, storing, and calculation of data, an ‘adaptive questionnaire’ was developed during this research.

This paper reports on the authors’ search for suitable survey methods to collect data to answer the research questions. It explores the feasibility of using computer networks in general, and the Web in particular, as educational survey research instruments. This paper first reviews the computer-assisted survey distribution methods and compares the advantages and disadvantages of each. The second part of this paper describes the development of a Web-based adaptive survey system and its questionnaire on health education. The evaluation process and results are reported. Conclusions and future work are also discussed.

2. Review of survey research methods

2.1. Computer-assisted survey research

Computers have been used to conduct social science survey research in many ways. In the construction of survey questions, computers help researchers draft, revise, and format questionnaires through their word-processing abilities. During questionnaire distribution and data collection, the technology and techniques of DBM, CATI, and CAPI help reach respondents as well as collect and store data for further analysis. Saris (1991) called the use of computers in survey distribution computer-assisted data collection (CADAC), and pointed out that this helps researchers skip the printing, mailing, coding, and data entry procedures.
Another term for CADAC is computer-assisted survey information collection (CASIC) (Couper & Nicholls, 1996). In this section, DBM, CATI, and CAPI are briefly described.

DBM is a CASIC method which allows respondents to use their own personal computers to complete and mail back questionnaires contained on disks (Saltzman, 1993). Wimmer and Dominick (1994) said that DBM surveys are essentially the same as the typical self-administered mail survey. Both procedures involve problem definition, questionnaire design, and pre-testing. The differences between regular mail surveys and DBM surveys, however, are the study type, sample selection, the use of computers, and so on. DBM surveys clearly require respondents to have a computer or at least computer access to answer the questionnaire on the disk. Therefore, the sample population is limited.

CATI uses computers to help interviewers and their supervisors perform the basic data-collection tasks of telephone interview surveys (Nicholls, 1988). In a typical application, the interviewer is seated at a computer wearing a telephone headset, reads the questions displayed on the screen to the respondent, and enters responses on the computer keyboard. House and Nicholls (1988) consider that the basic objectives for CATI questionnaires are to collect survey data, meet interviewer needs, and ensure program correctness, portability and documentation. Nicholls and Groves (1986) discuss the advantages of CATI to facilitate or expedite telephone surveys, enhance and control survey data quality, and permit new types of surveys not possible with paper-and-pencil methods. CATI is widely used in commercial research, at universities, and by government agencies (Spaeth, 1990). In short, CATI is the computerization of traditional telephone surveys.

More recently, CAPI has been introduced as another example of computerization of survey interview methods. In CAPI, the interviewer takes a laptop computer to the respondent’s residence. Similar to the way the computer is used in CATI, questions appear on the screen and the interviewer reads the questions to the respondents and keys in their answers (Saris, 1991). CAPI is usually used in large-scale survey studies, such as in the Labour Force Survey with 25,000 households in England (Sainsbury, Ditch & Hutton, 1993). Sainsbury et al. claim that CAPI can improve data quality by providing customized questions, performing mathematical calculations on-site, checking for inadmissible or inconsistent responses, and eliminating entry errors. When compared with paper surveys, CAPI generates both extra costs and savings. The extra cost involves the equipment investment, software development and/or paper questionnaire conversions to a computer version. The savings include the elimination of data entry and coding costs. CAPI also saves time by eliminating a separate data entry step. Thus, the time between the end of fieldwork and the supply of a clean data set is reduced.

2.2. Computer network-assisted survey research

Advances in computer networking technology allow stand-alone computers to be equipped with powerful communication abilities, thus providing additional survey distribution channels for educational researchers. Computer-network surveys can be used in two ways. The first is to supplement other methods such as mail surveys, phone surveys, and fax surveys. The same questionnaires are distributed by phone, regular mail, or networks. No matter what the transmission method, target respondents are free to choose their own reply methods, (such as in Schuldt & Totten, 1994). The second is to be used as the sole vehicle for distributing
questionnaires, that is, researchers need not make phone calls or provide paper questionnaires for some respondents, nor leave regular mailing addresses or fax numbers for respondents to return questionnaires. In this case, all target respondents must be network users, or must at least have access to and be able to use computer networks (Anderson & Gansneder, 1995).

Chou (1997) discussed the advantages and disadvantages of computer-network surveys. The first advantage is that the cost of delivering network surveys is far lower than the cost of using regular mail; if researchers are already connected to the Internet or have an existing e-mail system, charges are based only on transmission time (Anderson & Gansneder, 1995). Second, the delivery of network surveys is usually faster than that of mail surveys (Mehta & Sivadas, 1995; Oppermann, 1995). Third, the delivery of network surveys is more certain. Most e-mail systems provide for immediate return of undeliverable mail and thus help researchers identify incorrect or currently invalid addresses, enabling use of replacement samples if necessary. Fourth, network surveys relieve some time constraints on researchers as compared to phone surveys (Parker, 1992). Fifth, network surveys allow more user flexibility in responding, as in regular mail or fax survey. Sixth, networks provide a convenient medium for target respondents to communicate with researchers (such as in Oppermann, 1995; Fisher, Margolis & Resnick, 1996). Seventh, network surveys do not use paper, or at least they use less paper, so they are environmentally preferable (Saltzman, 1993; Parker, 1992).

Still, network surveys have some disadvantages. The first is that the target respondents must be network subscribers. Anderson and Gansneder (1995) reported that respondents may underrepresent infrequent users and people who perceive the e-mail system as difficult to access. Second, network surveys are prone to technological problems, such as incompatibilities between the researchers’ and users’ e-mail systems. Third, network surveys require researchers to have adequate computer-network knowledge and techniques, not only for transmitting the electronic questionnaires but also for collecting and checking returned data. Fourth, some human factors such as target respondents’ typing skills, computer anxiety, or inexperience with networks may contribute to a response bias (Schuldt & Totten, 1994; Oppermann, 1995).

In brief, compared with regular mail, phone, or fax surveys, the primary advantages of network surveys include lower costs, faster turnaround times, greater user-response flexibility, a convenient medium for extra information, and saved paper. In contrast, constraints on networks surveys include limitations on target respondents, possible technological problems, the need for researchers to have adequate network knowledge, and various human factors. Understanding the nature of network surveys and their inherent limitations will help educational researchers make better decisions in choosing among network survey methods.

3. Development of a Web-based adaptive questionnaire

3.1. Objectives and questions of the survey study

In order to develop an online health education courseware at NCTU in the near future, learner analysis was conducted by disseminating a survey questionnaire to target students. Some researchers such as Babbie (1992) and Fowler (1988) point out that the first and most important requirement of any survey study is a clear set of objectives. The objectives of the
present study were set as: “collect information about target students’ experiences in past computer and network usage, their opinions on health education, commentaries on proposed online delivery and presentation, and reactions to our on-line survey questionnaires.”.

The second step for conducting survey studies is to create questions based on the objectives by using word-processing software. The authors’ questionnaire consisted of five sets of items with a total of 59 questions. The first set consisted of 20 questions in a Likert-scale format; students were required to read the statements and indicate the extent of their agreement or disagreement with one of the options on a 5-point scale: A (agree), SA (strongly agree), U (undecided), D (disagree), or SD (strongly disagree). A factor analysis done on the same questions in a paper-and-pencil version revealed four principal components in the questionnaire: the needs for health information, the reason for looking for health information, channels for searching for health information, and sources of health information. Reliability (Cronbach’s) of the questionnaire was 0.83.

The second set consisted of questions 21–28 which asked students whether they know or have already visited some Web sites on health education. This set of questions is adaptive because the following question(s) will depend on the previous question’s answer. This set is in a net structure (see Fig. 1) in which the branching paths diverge after the lead items (questions 21–23, and 26), and converge on some items (questions 27 and 29 in the third set) within the structure. The pilot test indicated that this set of questions was quite complicated and somehow difficult for students to follow even when graphically presented on paper. The third set of questions, items 29–34, asked students about their opinions and wishes for a future Web-based health education courseware. The fourth set of questions (35–45) assessed students’ reactions to the authors’ on-line survey questionnaires. The fifth set of questions (46–59) assessed students’ demographic information and their past experiences in computer and network usage.

3.2. Choosing among the survey methods

As discussed above, a variety of survey methods can be used to distribute questionnaires. For this study, the authors crossed out fax surveys because fax machines are not popular fixtures in students’ laboratories or dormitories. Phone surveys, CATI and CAPI were not feasible because they constrain the researchers’ time to conduct the study and require a high investment in laptop computers. Since the questionnaire was not very long (only 59 questions), occupying less than 100K of disk space, the DBM method would be a waste and subject to computer virus infection. Between the final two candidate methods, paper surveys and network surveys, the authors chose the network surveys for the above-mentioned advantages and the following two specific reasons:

1. The Internet and its applications such BBS, e-mail, and the Web have already become important parts of most NCTU students’ campus lives nowadays. The university uses networks to pass public and private administration information to students; students use networks to register for courses, access digital libraries’ reference materials, and receive online gradebooks.
2. As Bugbee (1996) states, if what is learned or being tested is done from a computer, then it
Fig. 1. The structure of questions 21–28 in the questionnaire.
is more appropriate to assess it by computer. In this case, since the future health education
courseware would be presented on computer networks, it was felt it was more appropriate
to assess students’ opinions by using computer networks.

As Chou (1997) points out, computer network surveys can be distributed via four
major network methods: e-mail, news systems, bulletin board systems and the Web. The
first is the e-mail system that enables its users to exchange documents or messages in
electronic form. Researchers must have exact e-mail addresses for all target respondents
and then send them questionnaires in an e-mail letter. The second is the news system in
which researchers post text-formatted questionnaires and invite students to fill them in and
send them back to the researchers’ e-mail addresses. The third is the BBS which allows
students to log-in to read and leave messages. Surveys are distributed through the BBS in
ways similar to those used for the news system. The fourth is to use the Web, an
Internet hypertext-distributed information retrieval system. Researchers prepare the
questionnaire as so-called home pages or Web pages, and invite students to fill them in
and send them back to the researchers.

Comparisons of the four network survey methods can be made from different
perspectives (Chou, 1997). The first is the message format required to present questions, e-
mail, BBS, and news usually transmit only simple text letters or questionnaires, but
graphic or audio files can also be attached. The Web can present text, graphics,
animation, video, and audio, and ‘adaptive questions’ to reduce the number and
complexity of questions presented to users. Second, from a researcher’s standpoint,
different network methods demand different amounts of work before and after
questionnaire transmission. Similar to regular mail surveys, e-mail surveys require the
researchers to manually enter respondents’ e-mail addresses one by one into questionnaire
forms or mailing-list files. By contrast, news or BBS surveys require posting only once for
each respective news group or BBS, and Web surveys require putting an appropriate
folder or directory on a Web server only once. To send out and collect responses from
Web questionnaires, researchers must use the ‘form’ tool and develop common gateway
interface (CGI) programming to gather information. In contrast, questionnaires using the
e-mail, BBS, or news systems are returned in e-mail format. The responses can then be
checked and coded by hand or by computer programs developed by the researchers.

Third, in a sampling perspective, probability sampling can be done for e-mail survey studies,
but not for news, BBS, or Web survey studies. Parker (1992) points out that since e-mail
surveys require exact e-mail addresses for all respondents, they are appropriate for finite target
populations from which samples are drawn. If the researcher can identify a population for
which everyone has a known and valid e-mail address, and the researcher has access to those
addresses, then a true random sample can be drawn, as in the study by Anderson and

On the other hand, because of the anarchical, distributed, and geographically heterogeneous
characteristics of today’s computer networks, it is almost impossible to define the population
of network users, and difficult even to define the population of one particular news group, a set
of BBS users, or visitors to a particular Web site (Chou, 1997). Without a clearly defined
sampling frame, probability sampling cannot be done for news, BBS, or Web surveys.
Therefore, sample problems and self-selection biases may confound these surveys if researchers want to generalize research results.

In this case of health education survey study, probability sampling was required in order to generalize the survey results and the Web’s branching feature was needed to present adaptive questions. Therefore, a hybrid method was adopted in which the authors drew probability samples from their target students e-mail account addresses, sent them e-mail letters explaining the research purpose, and asked them to check the researchers’ Web site ‘http address.’ If target respondents used regular e-mail software to read their e-mails, they had to use the Web browsers such as Netscape or Internet Explorer (IE) to access and answer the questionnaire. If target respondent had already used these browsers to read their e-mails, they could just point and click the http address, as a ‘hotkey’ in the authors’ letter, then the browser would access the questionnaire automatically for them.

3.3. The development of adaptive questionnaires

In order to use the computer network’s branching capabilities and reduce the number of questions and complexity of the questionnaire, adaptive questioning techniques were used in the Web survey system. Unlike past adaptive tests based on item–response theory (IRT) (Load, 1980; Bejar, 1983; Jacobs & Chase, 1992), the authors’ adaptive questionnaires used CGI programming to control presentation. Adaptive questioning uses the answers to certain questions to determine the next series of questions and to skip unrelated questions. In this case, questions 21–23, and 26 were adaptive questions in which student had to make yes/no choices, and then the next appropriate question(s) would be presented. Students answering ‘yes’ to question 21 were taken to question 22, while students answering ‘no’ were routed to complete questions 25 and 26. The process is used frequently in most surveys; however, Web questionnaires can be programmed so that questions 25 and 26 would not be shown at all to the students answering ‘yes’ to question 21. In other words, if a student said ‘no’ to question 21, she or he would not have the chance to mistakenly answer questions 25 and 26. Pitkow and Recker (1994) consider the Web survey to be probably the easiest for target respondents to use and respond to when compared with other methods of network surveys.

All survey questions, whether they are adaptive ones or not, were stored in a question bank. The first set of 20 questions, the third set of 6 questions, the fourth set of 11 questions, and the fifth set of 14 questions were placed on one Web page each for a total of five web pages. Although the page covered more than one screen, students could use scroll bars to control questionnaire presentation. Fig. 2 is a sample questionnaire screen for part of the first set of questions. The adaptive questions in the second set were presented one by one on individual Web pages. For students, the diamond-shaped symbols in Fig. 3 represent the adaptive questions for which they had to make yes/no answers. Researchers developed the CGI linkage functions in visual basic language to judge students’ yes/no answers in order to trigger or not trigger the next series of questions. Fig. 4 is the screen for question 21. Please note that the entire questionnaire is in Chinese; this English version is for demonstration purpose only.

3.4. The evaluation of a Web-based adaptive questionnaire

In addition to the system developers’ self-testing (α-test), a formative evaluation was conducted to examine the usability and the presentation of the health education survey system
and questionnaire. Two major evaluation approaches were adopted: expert-based and user-based (Sweeney, Maguire & Shackel, 1993). One experienced computer-interface expert, one survey methodology expert, and one network technology expert were invited to assess the aspects of interface, format, presentation, branching functions, and programming techniques of the system. The evaluation method used was thinkaloud. Flagg (1990) points out that the thinkaloud or talkaloud method was used to obtain information from users about their moment-to-moment processing of the program while they were using it. Nielsen (1993) believes thinkaloud may be the single-most valuable usability evaluation method. In the present study, three experts reviewed the system and questionnaire while continuously thinking out loud, so researchers could understand not only what they said about the system and interface, but also why they were doing what they did. Data gathered from the three thinkaloud reviews provided valuable expert opinions, and the system, interface, question presentation, and programming were then revised accordingly.

A pilot test of the paper-based version was conducted by 95 volunteer sample students and on-line version by 15 students. The purpose of the two pilot tests was to help researchers refine the questions, identify possible technical problems such as unidentified characters or incorrect formatting appearing on the target students’ computer screen after transmission, and test the accuracy of the CGI programs and the codebook.

![Figure 2. A sample questionnaire screen for the first set of questions.](image.png)
A sample of 150 undergraduate students was drawn from NCTU. The sampling procedures involved the drawing of stratified samples from one class of five colleges (Electrical Engineering and Computer Science, Engineering, Science, Management, and Humanities and Social Sciences). A total of 76 valid questionnaires were answered and returned. Since the present research focus is on students' reactions to the on-line questionnaire, only their answers to the fourth set of questions are reported here. About 64% (49 students) indicated they liked this kind of questionnaire, 68% (52 students) thought it was easy to read, 57% (43 students) said it was faster to answer than the paper-and-pencil questionnaire, and 79% (60 students) viewed network questionnaires as more efficient than paper-and-pencil ones. When asked whether network questionnaires increased their anxiety, about 74% (56) answered no, 11% (8 students) answered yes, while 15% (12 students) were undecided. When asked whether network questionnaires were more likely to tire them than paper-and-pencil testing, about 38% (29 students) said yes, 50% (38 students) said no, and 12% (9 students) were undecided. If asked to choose between a network questionnaire and a paper questionnaire, about 53% (40 students) preferred network version, 15% (12 students) preferred the paper-and-pencil version, while 32% (24 students) expressed no favorite.

Fig. 3. The design of the adaptive questionnaire; diamond shapes indicate adaptive questions.
When asked whether the screen interfered with their answering process, 71% (54 students) answered no, 24% (18 students) said yes, and 5% (4 students) had no opinion. When asked whether the typing speed significantly influenced their completion time, 59% (45 students) said no, 36% (27 students) said yes, and 5% (4 students) expressed no particular answer. These results contradict the present researcher’s current study (Chou, 2000) in which a Web-based test was conducted. In that study, about 67% of 54 students said computerized testing was more likely to tire them than paper-and-pencil testing, and about half the students said the typing speed significantly influenced their completion times. The difference between these two studies may come from the ‘testing effect,’ that is, being tested made students feel more tired and under pressure. In a survey, however, students’ answers were not graded and no long open-ended answers were required. Students basically performed ‘point and click’ to answer the questionnaire. Therefore, they did not think the distribution method (whether paper or computer screen) interfered with their answering, and typing speed was not of a great concern.

In short, students liked the network questionnaire. They thought it was faster, more efficient, easier to read, and did not increase their anxiety. More than half of the students’ surveyed would like to choose network questionnaires in the future.

Fig. 4. The screen for question 21; clicking ‘yes’ or ‘no’ will trigger different questions to follow.
4. Discussions and conclusions

The development and evaluation of Web-based testing have raised several research and utilization issues. First of all, the successful team-approach is required. When computer network techniques are used to construct an on-line questionnaire system like the one reported above, the technical components and research methodology components should be considered indispensable. Katz and Conrad (1996) point out that developing CASIC instruments is often resource-intensive and time-consuming, requiring the cooperation of computer programming experts and questionnaires design experts. Network technologies provide the system framework and functional blocks upon which questionnaires can be developed and delivered. They also help record and analyze students’ inputs. On the other hand, new questionnaire design strategies and presentations are needed to use the full capabilities of network technologies, such as the colorful presentation and adaptive questions in the present study. In addition, empirically-derived information on system functionality, user-interface design, and students’ reactions to the on-line delivery are needed to guide the quantitative and qualitative development of educational research applications of computer network technologies.

Other issues concern the privacy and safety of network surveys. In this study, no sensitive questions regarding students’ private health lives were asked. Only one question asked was whether they would ask their friends if they had some sensitive health problems such as sexually-transmitted disease, out-of-wedlock pregnancy, etc. The purpose of this specific question was on the information source, not on their existing or possible health problems. Even so, the present study followed some existing guidelines, as suggested by Anderson and Gansneder (1995), including informing participants of the nature of the research, making their participation voluntary, and keeping their data confidential. The authors understand that survey research is usually an intrusion into people’s lives, and network surveys represent intrusions into people’s network environments. How to balance the researchers’ needs to inquire and the individual’s need for privacy is itself both a science and an art, and an issue for future research.

Other research issues concern the design of Web-based questionnaires. As discussed above, the Web is an Internet multimedia information retrieval system which provides more message formats to present questions. Although few academic surveys today incorporate multimedia messages, some on-line surveys for entertainment purposes have often already used graphics, voice, or video clips. For example, as mentioned by Chou (1997), in a Web survey conducted by ESPN, users were asked to click on photo icons to activate video clips and vote for the ‘Outrageous Play of the Year.’ In an ‘on-line Pageant,’ clicking the photo icon of any participant activated a video clip and a vocal introduction of that person. Self-examination of psychological traits on networks is also popular among young adults. After clicking on a favorite picture, feedback tells you what type of person you are. Should educational survey research studies incorporate such message formats in the future? Does this design influence respondents’ responses? How can one design the Web questionnaire so that it provides respondents with new survey experiences not previously possible? These are research questions waiting to be answered.

This paper has discussed the comparisons among a variety of survey methods, considering their advantages and disadvantages. It reported the development of computer networks as
innovative research instruments. As the world becomes more related to computer-networks, so too must surveying. Even though regular mail, phone, fax surveys are here to stay for still some time, computer-assisted survey information collection techniques will keep improving. Researchers such as Schuldt and Totten (1994) and Walsh, Kiesler, Sproull and Hesse (1992) have claimed that network surveys will become the standard data collection method in the 21st century and be used routinely. The authors’ exciting and positive experience in the development and utilization of a Web-based health education survey supported their optimistic claim.

Acknowledgements

This study is supported by the National Science Council, Taiwan, R.O.C., under Project NSC87-2511-S-009-007-ICL.

References


