The Use of Online Surveys in Capturing Large-Scale Data

Jan Stenton & Joanne Pascoe
Griffith University

Surveys are regularly used as a tool to capture data for research in the social sciences. Advancements in Information and Communication Technologies, including the rapid growth of Internet usage, have encouraged the development of online and email surveys as a cost and time effective research tool. This paper will explore the design and implementation issues of using a Griffith University developed online survey maker for a DEST funded Queensland-wide, cross-sectorial study involving primary school principals and learning support teachers and the provision of education services to students with learning difficulties. It will discuss national and international issues and trends regarding current use of online surveys for research in order to inform the design, construction and implementation of surveys for the pilot study in the current ongoing project.

Introduction
Electronic technology, which has grown at a phenomenal rate, is a potentially powerful tool for education researchers. Both the Internet and email can be the delivery vehicle for large-scale surveys. The purpose of this paper is to examine issues related to using online and email surveys for data collection in a large, statewide cross-sectorial study.

Literature review
Since it first became available to the public, acceptance and use of the Internet has grown more rapidly than that of any previous technology, including radio and television (Duffy, 2002). The interactive nature of the Internet gives it the potential to be a highly useful tool and the Internet's potential for academic and applied research has recently begun to be acknowledged and assessed (Sheehan & Hoy, 1999). Researchers use both online based surveys to study large groups of online users and email surveys to study smaller, more homogenous online user groups. Increasingly sophisticated software and increased computer power has enabled web survey designers to create complex surveys that utilize html tables, multiple colours, motion and other advanced features to appeal to respondents and encourage them to provide answers (Dillman, Tortora, Conradt, & Bowker, 1998). Computer assisted data collection methods are increasingly replacing paper-and-pen methods of survey data collection (de Leeuw & Nicholls, 1996). Online surveys are surveys presented as web pages and the respondent enters the data online.
Advantages of online surveys
Sheehan & Hoy (1999) suggest various reasons why web-based research is a useful tool for academic researchers. This research method is capable of collecting broad-based data from individuals all over the world, limited only by access to networked computers. Individuals self-select to respond to surveys and the information they are frequently asked to provide includes demographic data and other types of purchase, psychographic and opinion data. Surveys can be made attractive, interesting and inviting to respondents using the graphic powers of current programming languages. Questions can be designed to adapt to a respondent's responses (adaptive questioning).

Web-based polls can be far reaching and have been found to have the potential to generate a high number of responses very quickly. According to Duffy (2002) web-based research is useful for access to geographically and culturally diverse populations as well as for difficult to find populations. Apart from the high start-up costs for equipment and web page design, the implementation of an online survey has no postage, paper, nor data entry costs. Also data analysis costs are low when this is undertaken electronically following direct transfer of responses to analysis software. A further cost saving is a reduction in the need for interviewers and this also leads to survey responses being free from the potential for interviewer introduced bias.

Web-based surveys appear to have the option of providing anonymity to respondents and this should affect response rates positively. Response rates can also be increased by the use of enticements. Particularly prevalent these days is the chance of winning prizes as an incentive for potential respondents to participate in and complete online surveys.

Limitations of online surveys
Online surveys are not a complete panacea for all of the ills of traditional survey data collection. Using the latest version programming features in a survey may actually require significant transmission and processing time by respondent browsers and even cause overload in older browsers to the point that the survey may not download or might crash when the survey is only partially completed. This would "cause response rates to be lower than those obtained by simpler more conventional designs, the opposite of their intended effect" (Dillman et al., 1998).

Some respondents will be more attracted to an electronic survey than others. Respondents will not all use the same browser, so they may not all see the images and text in the same manner. This can be a particular problem for researchers who look for potential respondents through newsgroups and links to other web pages etc. Sheehan & Hoy (1999) maintain that some limitations in the use of online surveys include generalisability issues. The self-select nature of online surveys will affect their generalisability. It is almost impossible to develop response rates to web page-based surveys based on survey "hits" as it isn't possible to tell how many "hits" are revisits. Due to the anonymity of open online surveys, questions can be raised about the accuracy of respondent information and identification. Another issue is that online surveys frequently allow for multiple attempts from respondents, as well as from respondents outside of the population of interest, which could lead to biased results.
Email surveys

Email can be a suitable method for collecting data from a large population and experiences many of the same advantages and disadvantages as a data collection method as the online surveys though several issues pertain uniquely to the email survey. An email survey involves either writing the survey questions in the body of the email or attaching the survey to an email possibly as a Word or Excel document. Using email as a survey data collection method has the potential to minimize some of the issues associated with collecting data using web-based pages. Numerous people these days have email accounts, however, a lack of a national directory of email addresses could be a problem for researchers in obtaining email addresses for appropriate respondents. Online search engines might provide a solution to this issue depending on the population required.

Benefits of email surveys

An advantage of using email is that it is possible to increase response rates by tracking responses and sending reminders only to respondents who have not responded (Cohen, 1988). An interesting feature of email responses is that they can be returned using one of several methods, such as email, fax or post. This flexibility may appeal to some respondents as they have the benefit of completing the survey at their own pace and convenience.

Limitations of email surveys

Based on a study that analysed response rates to email surveys, Sheehan (2001) suggests that because response rates to email surveys have significantly decreased since 1986 they might have become obsolete. There are various reasons for this. Sheehan & Hoy (1999) identified two major limitations unique to email. The first one is that respondents may consider unsolicited surveys inappropriate, intrusive and not in keeping with current online culture. Information overload as a result of receiving numerous emails is a current problem and it is likely these respondents will delete the survey without responding to it (Sheehan, 2001).

Secondly, it appears likely that email addresses may become out-of-date quickly. It is suggested that it is difficult to compensate respondents for participating in an email survey but that a compensation programme may be an option for obtaining a higher response rate. Enticement using a reward (for example, a financial reward or a holiday) would increase the cost of both email and online surveys. It is suggested that the respondents tend to regard email invitations for survey response as "SPAM" or junk email (Topp & Pawloski, 2002). An incentive for completing electronic surveys could include using an authority figure of interest to the respondents to send the email invitation.

Anonymity can be a further problem if encryption technology is not used, as the respondent's email address automatically attaches to an email response. Confidentiality can only be guaranteed through confidentiality assurances.

Specific use of the Internet for data collection

In one of only a few documented studies that has used an electronic questionnaire for data collection, Mrayyan (2004) reports that the availability of listservs, valid emails,
viruses and familiarity with the Internet and its applications were the major technical issues they encountered in a research project examining the role that nurse managers have in enhancing hospital staff nurses' autonomy. Electronic questionnaires were used to collect data from nurses who participated in 23 nursing listservs in the United States of America, Canada and the United Kingdom. The study reported a non-deliverable rate of 23% for the electronic questionnaires (sent via email) due to a variety of reasons including invalid email addresses and changes in Internet providers. As a result it was difficult to compute an accurate response rate but there was an estimated response rate of 10% from a large and diverse population (6,000 addresses were collected). Discussion of this study suggests that use of the Internet limits the research samples to those respondents who have access and the knowledge to use the Internet. However, it also suggests that "unless some researchers take the risk of conducting electronic data collection, the promises of web-based research will remain unfulfilled" (Mrayyan, 2004, p. 328).

Piloting SurveyMaker

Background to the study

A study in Queensland funded by the Australian Government Department of Education, Science and Training is being undertaken in 2004 to 2006. It is a large, field-based mapping of provision of services to primary school students with learning difficulties in literacy and/or numeracy. A random sample of 500 primary schools statewide across all education sectors will be surveyed, with a target final response rate of at least 300 schools. Of the 500 schools, there are 373 Education Queensland (state), 70 Queensland Catholic and 57 Independent schools. These numbers reflect the proportion of schools in each sector as well as representing schools from urban to rural/isolated and across a wide range of cultural compositions.

Part of the project involves the collection of information from principals and learning support teachers in the identified schools. An online survey was decided upon as possibly the most appropriate and economical form of data collection for several reasons. They include the large sample size required for a statewide, cross-sectorial study; the anticipated high response rate expected by using an online survey; and the availability of timely reports made possible with computerized data collation. A pilot study using the online surveys was undertaken across sixteen schools and this paper will discuss the development of these online surveys, with particular concentration on the survey for the school principals.

The instrument – SurveyMaker

The programme "SurveyMaker" was used as the data collection instrument. It is an online survey tool developed and supported at Griffith University. It was originally designed to conduct evaluations of teaching and learning environments in Higher Education. Using SurveyMaker, university lecturers could send surveys to their students about course content and delivery and the student responses were received, data analysed and a report formulated quickly and efficiently. Data entry is done in the process of completing the survey and this programme can be set up to email frequent reminders to the students to complete the survey. SurveyMaker has been extended as a research tool because it has a number of useful features.
Using the SurveyMaker Wizard, it is possible to quickly and expertly build an online evaluation or survey. Formatting the online survey, emailing respondents and the administration of the survey are all automated. The online Report Generator can analyse the data and instantly generate a report with tables and graphic charts. There is no limit to either the number of questions or the length of the survey. Once the survey has been put online it is possible to overview its status at any time with snapshot views of completions and response rates. Response rates can be improved with computer-generated email invitations, follow-up and acknowledgements at specified time points.

SurveyMaker has a large range of question formats including multiple-choice questions, Likert scales, binary questions (true/false; yes/no), ranking, rating and prioritising, multiple response and free text responses. It is also possible to combine two or three of these question types into one combined question. It is possible to customize the path a respondent takes through the survey by adding skip logic on Boolean questions. This can eliminate unnecessary confusion by skipping non-applicable questions and is a user-friendly way to reduce confusion and dropouts that jeopardise the validity of the survey. Responses to essential questions can be made mandatory by specifying on a question-by-question basis the questions that require an answer. This helps to encourage data quality.

**Report generation**

Computer generated html reports in a variety of graphic outputs are available in straightforward data reports or more structured reports in which introductions and discussions can be entered into a pro forma. Data can be accessed and reports viewed and generated as the data come in, with final versions available after the survey closing date. Data files can be downloaded at any time for importing into other statistical analysis software packages for further data manipulation. If there is a need for respondents to complete paper-based surveys SurveyMaker allows the designer to setup Data Processors who will enter the data into the system so that all results can be collated.

A variety of reports can be generated including question-by-question reports, which show the statistical data for each question as a table or bar graph. Alternatively all questions and data for a survey can be displayed on the one page. Accompanying text can be added as an introduction and conclusion to the report and the table and/or graph views for each question can be displayed. The programme also generates the statistical data for the survey. It is possible to share results with other people who can view the report online without having access to the data or the account.

**The online survey for principals**

A survey will be designed and sent to the principals in the study giving them the opportunity to respond to the following major questions addressed in the project.

In relation to students with learning difficulties in literacy and/or numeracy:

- What are current intervention approaches and what is known about their impact/effectiveness?
- What are the accounts of assistance for target students in policy and in practice?
The survey focuses on:

- Modes of literacy and numeracy intervention;
- Perception or evidence of effectiveness;
- Assessment information collected relating to effectiveness for individual students and student progression;
- In-school and out-of-school support for literacy and numeracy;
- Characteristics of students identified as experiencing learning difficulties.

The online principal survey is also one of the strategies being applied to locate school case study sites by inviting schools to self identify if they believe they have exemplary programs. Table 1 outlines the focus of the questions in the online survey for the principals in the pilot schools.

<table>
<thead>
<tr>
<th>QUESTIONS</th>
<th>FOCUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Current size of the student population by year and gender.</td>
</tr>
<tr>
<td>2 – 3</td>
<td>Assessment information used to identify students with learning difficulties.</td>
</tr>
<tr>
<td>4 – 7</td>
<td>School-wide data on students by year, gender and literacy/numeracy difficulties.</td>
</tr>
<tr>
<td>8</td>
<td>Response to DEST definition of students with learning difficulties.</td>
</tr>
<tr>
<td>9 – 15</td>
<td>Interventions – type and where they occur.</td>
</tr>
<tr>
<td>16</td>
<td>Resource allocations learning support personnel.</td>
</tr>
<tr>
<td>17 – 20</td>
<td>Out-of-school support including tutoring and parental.</td>
</tr>
<tr>
<td>21 – 24</td>
<td>Likert scale – monitoring and tracking student progress.</td>
</tr>
<tr>
<td>24</td>
<td>Collaboration between staff across schools.</td>
</tr>
<tr>
<td>25</td>
<td>Open question – hindrances and helps</td>
</tr>
<tr>
<td>26 – 27</td>
<td>Effectiveness of support for students and identification of schools for case study</td>
</tr>
</tbody>
</table>

The online survey for learning support teachers

The survey for learning support teachers was developed in response to the same two concerns that were the basis of the principals' survey. An early version of this survey was used in parallel with the development of further questions designed to be more specific to the role of the learning support teacher. The survey focuses on modes of intervention; perception or evidence of effectiveness; assessment information collected relating to effectiveness for individual students and student progression as they relate to the role of the learning support teacher. Issues of in-school and out-of-school literacy practices and representation of gender make-up of students identified as experiencing learning difficulties are also explored from their perspective. Another aspect of this survey is to
give learning support teachers the opportunity to present exemplary practices from their school in order to locate school case study sites for in depth investigation. A similar process to the one used for the development of the principals' survey was used to develop this survey. Table 2 provides the focus of the questions in the survey for learning support teachers.

**Table 2**

Focus of questions for learning support teachers' survey

<table>
<thead>
<tr>
<th>PART</th>
<th>QUESTIONS</th>
<th>FOCUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part A</td>
<td>1 – 2</td>
<td>Modes of support:</td>
</tr>
<tr>
<td></td>
<td>5 questions</td>
<td>Numbers /Year level / Gender / Literacy / Numeracy</td>
</tr>
<tr>
<td></td>
<td>3 – 4</td>
<td>Involvement with other people</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Definition of learning support</td>
</tr>
<tr>
<td>Part B</td>
<td>1 – 2</td>
<td>Demographics</td>
</tr>
<tr>
<td></td>
<td>2 – 6</td>
<td>Learning support teacher:</td>
</tr>
<tr>
<td></td>
<td>23 questions</td>
<td>Qualifications / Experience / Work details</td>
</tr>
<tr>
<td></td>
<td>7 – 9</td>
<td>Identification / Allocation of learning support</td>
</tr>
<tr>
<td></td>
<td>10 – 15</td>
<td>Practice of learning support:</td>
</tr>
<tr>
<td></td>
<td>16 – 19</td>
<td>Collaborative partnerships:</td>
</tr>
<tr>
<td></td>
<td>20 – 23</td>
<td>Intrinsic factors related to learning problems:</td>
</tr>
</tbody>
</table>

**Implementation of the surveys - pilot stage**

There has been a collaborative process of developing and trialing the pilot survey involving the researchers and sixteen principals with the assistance of a deputy principal from Education Queensland, who is working on a similar, but much smaller study. As well, the project's advisory committee has viewed the survey for comment and advice and comments from the three sectors and other relevant parties was incorporated into the survey.

The importance of piloting the online survey was to trial the technology in school situations; identify strong reactions against certain questions; be notified of important issues omitted from the survey; to find relevant options for multiple choice questions; as well as to ensure that the terminology is up to date and suitable across the three school sectors. The advantage of Surveymaker in this process was the access to instant and cumulative reporting as well as the ease of distribution of the survey. A response rate of 100% was obtained from the pilot study.
General observations of online survey methodology for this application – Strengths

An online survey was decided upon by the research team as an effective way to reach a wide and large sample of target respondents. Initial perceived advantages of the program Surveymaker were the availability and technical support to Griffith University researchers and its professional finished format. The program proved relatively easy to master and results of the pilot survey impressed researchers with the instant reporting in a variety of useful combinations of data and graphics.

The pilot survey assisted the team to reflect upon current terminology in the field, cover the breadth of options possible in multiple choice answers and add questions requested by pilot participants. Once responses to the pilot were discussed with the researchers, it was easy to duplicate the survey in Surveymaker to create a new version.

Automated instant reporting of results is a great advantage of the online data entry. Data is collected in a format that can be moved to SPSS and Excel programs for further statistical analysis. Timely reporting of feedback is considered important to both the university ethics process and for goodwill towards cooperative research participants. Because the survey is one of the methods of selecting case study schools, the ability to respond immediately to incoming data in order to involve identified participants in a school case study is an advantage.

Limitations

Several issues arose in the use of the online survey in the pilot stage that required consideration by the research team. These limitations will be discussed then solutions proposed. The first issue of concern in online/email-based surveys is the ease of deletion and the competition for attention amongst growing numbers of incoming emails. As the pilot schools were eager to participate in the survey this was not an issue at this time but it is anticipated that it may become an issue when the surveys are sent to the whole sample.

Some limitations were more to do with the characteristics of the target respondents than with the Surveymaker program. The greatest difficulty was in the requirement to cover a broad range of situations and terminology in one survey without allowing the wording to become uselessly generic. This was due partly to the inclusion of three different education sectors in the sample as they each have specific terms etc that they use. It was preferable that the one survey be appropriate across all three education sectors and the rural/urban situations. Difficulties arose in trying to cover broad options but also make the survey specific enough to provide useful information to the project. Using SurveyMaker it would have been easy to duplicate the survey for each of the three sectors and change the terminology to suit each sector. The team decided against this as it would restrict the ability to combine all the responses into the larger sample size. Furthermore, it was difficult devising easy to collate numerical-type responses or data for the information required. The alternative, in some cases, was to provide open responses using text boxes, but this will require coding and data entry by the researchers.

There are limitations in the program in formatting issues. For example, there are no "Other" or "Please specify" boxes available for multiple-choice questions. Other issues
are more trivial, such as the absence of a spellcheck programme, but this becomes a frustration for the researchers when combined with the small print. It makes it difficult to thoroughly check SurveyMaker for errors. A useful way around this was to copy the SurveyMaker text into an empty Word document and check it for errors there.

It is likely that the principals are not the holders of all of the information required to complete their survey and they may have to refer to learning support teachers and registrars in order to do it. This became a problem due to the limitation of the survey timing-out if the program is left unattended for thirty minutes. It was also an issue because the school principals often lacked a large enough block of time in which to complete the task. Limited access to technology in some schools also proved to be a problem and the researchers are considering making a print or email version of the surveys available in an effort to overcome the situation where access is restricted to only those respondents with suitable technology. During the pilot stage it was found that the function allowing the respondents’ to print their own responses for their records was not working satisfactorily, although SurveyMaker programmers are working to improve this situation. There are limitations with technology and information technology support available to some participants. In the case of learning support teachers, not all have an email address or access to a computer. They may have to use principals' computer and that would possibly compromise their confidentiality.

**Design solutions**

The pilot stage, sent to sixteen schools, revealed a number of problems with the use of an online survey. The technology in some schools was not adequate for the online survey. Specifically, Education Queensland school computers are set with high security levels. Some school computers incorrectly branded SurveyMaker as a game and blocked access to it. Levels of technology support vary greatly between schools, so it is anticipated that technical difficulties with accessing the survey might lower response rates. In response to this a Word version of the survey was made and it became an email survey allowing the participants an alternative means of entering and returning the data. Two of the design benefits of the Word version of the survey are that it is easier to construct new questions in the more familiar program of Word and that the language can be spell checked before being copied into Surveymaker. Benefits to the principals in the pilot stage were that it allowed them to consult with various staff members for answers beyond their knowledge without risking the online survey timing out.

Due to the large variety of respondent circumstances, it was not possible to devise all the survey questions as easy-to-collate, numerical type responses. The alternative for some questions was to have open response boxes, which will require coding and data entry on the part of the researchers. In some cases it was realised that the depth of questioning was more appropriate to interviews or case studies than to a large-scale online survey. There is a limit to the scope of questions possible in an online survey.

Given than the targeted participants are extremely busy, once the email survey is filled out, another more junior staff member can enter the data and the emailed Word version of the survey also remains as a record of their responses. In the case of the learning support teachers using the principal's computer, they must be assured in the
accompanying letter that the principal is unable to access their responses once "submit"
is pressed at the completion of the survey.

Due to the commitment of the pilot schools to this project, the response rate to the
online survey was 100%. It is difficult to determine if such a successful response is
possible when the online survey is sent to the whole sample but given the limitations
discussed it is unlikely.

Conclusion

Online and email surveys are a cost and time effective research tool to capture large-scale
data for research in the social sciences. This paper has explored some design and
construction issues related to the use of a Griffith University developed online survey
maker for a DEST funded Queensland-wide, cross-sectorial study involving primary
school principals and learning support teachers on the provision of education services to
school students with learning difficulties. The observations from the use of SurveyMaker
at the pilot stage in this ongoing study tend to mirror the general findings of the
literature's exploration of national and international trends regarding current use of
online surveys for research.

References

Erlbaum Associates.


design on response rates of Web surveys. Paper presented at the Joint Statistical Meetings, Dallas,
asa98ppr.pdf.


45(3), 326–336.

issue2002/sheehan.html.

Sheehan, K. B., & Hoy, M. G. (1999). Using e-mail to survey Internet users in the United States: