An International Survey of Research in Statistical Education

Flavia Jolliffe
University of Greenwich, Computing and Mathematical Sciences
Old Royal Naval College, Park Row
London SE10 9LS, UK
F.R.Jolliffe@gre.ac.uk

1. Introduction

Statistical education research is an emerging discipline, and researchers in this area do not yet have a clear identity (Jolliffe 1998). There are relatively few outlets for publication and presentation of research activities and results so that researchers tend to feel isolated, and finding out what others are doing is partly a matter of chance. In consequence important research findings do not always get as widely disseminated as they deserve, and unbeknown to one another researchers could be doing similar studies, but without the advantages offered by discussion and comparison of methodology and comparison and pooling of results.

Clearly an easily accessible and widely available database giving details of research and researchers in statistical education would be a valuable resource for both current and future researchers in this field, and for those who teach statistics. An international survey to find out who the statistical education researchers are, and details of their past, current and planned statistical education research could provide the initial input to such a database. The Learning and Teaching Support Network (LTSN) Centre in Mathematics, Statistics and OR felt that an international survey of this nature would be useful and productive for the teaching profession in statistics and OR (Davies 2000) and is funding the survey discussed in this paper. By indicating the directions which research is taking and where research is needed, the results of the survey might well influence the future of statistics education research.

This is an ambitious project, presenting several challenges. It might be argued that in itself it constitutes statistical education research. As with any such study, the different stages involved are inter-related, and consideration has to be given to its aims, the target population, the sample design and sampling frame, the design and method of administration of the questionnaire, and to the presentation and dissemination of results. As it is primarily a fact-finding exercise, questions of inference are of secondary interest.

2. The nature of the database

As the main aim of the survey is to provide initial input to a database, the contents of the database and the ways in which this might be used help determine many other aspects of the study. Society as a whole is fast becoming internet dependent, and although there are currently parts of the world where access is difficult or impossible, it is thought that the majority of those interested in statistical education research have the facility to obtain information from the World Wide Web (WWW). This is certainly the case for those based in the UK, for whom the LTSN has been set up. This points to placing a database of statistical education research and researchers on the Web, with all the advantages that offers over a printed version, such as ease of searching for items, and flexibility in making changes.

Ways in which users might want to search a database on statistical education research include searches by the topics (both general and specific) of the research, by the educational stage with which the research is concerned, by the research methods used, and by the name of the researcher. Each of these search areas can be broken down into categories, "research topic" having by far the largest number. Users might also want to search for specific products such as software or assessment instruments arising from research. Searches would be made via a user interface on a Web page. As a result of a search, users are likely to require contact details of researchers and publication details of the research or products. It should be emphasised that no evaluations of the quality of the research are envisaged, nor descriptions of the research apart from key words. The database would be a kind of directory. It would complement bibliographies on related topics such as Sahai et al (1996) and Garfield et al (2000), and would serve a different purpose from the newsletters of the Statistical Education Research Group within IASE which

include summaries of publications (http://www.ugr.es/~batanero/sergroup.htm).

Consideration of users' likely requirements suggests that the information collected would best be stored in a relational database having two main tables, one for researchers and one for publications, and these would be linked through a table containing only pairs of unique researcher and unique publication identification codes. Publication has to be interpreted broadly to cover both hard copy and electronic versions of research activities. To avoid repetition of data such as an address of an institution common to more than one researcher, or details of conference proceedings containing several papers, the data structure would be simplified by storing data relating to items of this nature in separate files. The design would allow for the addition of details of new researchers and new research, and of details relating to past research as they come to light, and would allow updating, for example of changes of address. Historical data would be retained.

3. Contributors to the survey

The target population of the survey is anyone living at the time of the survey who is currently undertaking or planning statistical education research or who has done such research in the past. A broad definition of research would be taken.

There are no ready-made lists of statistical education researchers although there are lists of persons with an interest in statistical education, for example members of the International Association for Statistical Education (IASE) and participants in the International Conferences on Teaching Statistics (ICOTS), and these are a useful starting point. There are also email discussion lists and lists of subscribers to journals such as the on-line Journal of Statistical Education, and to Teaching Statistics. Most of these lists are biased towards the English speaking world, and as few of them are public permission to use them would have to be obtained. The majority of those who are visible in doing research in statistical education are university teachers or researchers, and it is relatively easy to access lists of these, although such lists are often out of date. Many school teachers probably engage in research-like activities, although they do not bring their work to the attention of the larger community of statistical education researchers. Finding members of this elusive but important minority group is not easy.

Taking into account the lists that are available and the fact that unknown researchers might be considered to be a rare population, a combination of snowballing and of screening lists likely to contain names of statistical education researchers would be an appropriate method of building up a list of members of the target population (Kalton and Anderson 1986). In snowballing, known researchers in statistical education would be asked for names and contact details of other statistical education researchers. Further researchers discovered in this way would then be asked for names and contact details of other researchers. The process would be repeated until the number of additional researchers discovered is negligible. In order to make the implementation of this snowballing as smooth as possible only those on email will be asked for names of other researchers.

The initial list will be drawn up as an amalgamation of lists thought to contain a high proportion of researchers in statistical education. Some people will be on more than one list, but sorting names on the combined list into alphabetical order, relatively easy to do with lists in electronic form, will make it easier to spot and eliminate duplicates. Misspelt names and matches of name where there are different addresses are likely causes of failure to match. The next stage would be to contact those on this list to identify the statistical education researchers. This initial contact would be a good opportunity to publicise the survey and to get agreement to participate, and should help increase the response rate (Witmer et al 1999). Screening lists of school teachers would be much more troublesome, due to the difficulty of obtaining lists, and the small proportion out of a large number of teachers likely to be statistical education researchers, and would be impractical on a world-wide basis. It is hoped that a mixture of snowballing and targeted announcements would result in some expansion of an initial list of teachers thought to be researchers, but there is a risk of bias in this and of non-inclusion of some groups.

Two obvious problems with snowballing are that those researchers who are well known will be on the initial list, and will be named by many other researchers, so that this procedure will be rather wasteful, and those researchers who are isolated and known to only a few are fairly unlikely to be discovered by the procedure. One way of avoiding the naming of those who are

already known to be researchers in statistical education is to circulate a list of these when asking for names of other researchers. Checking and dropping duplicates would be a similar process to that done when drawing up the initial list. Once researchers have been identified, a procedure for contacting them at later dates would be developed in order to update their entries on the database.

4. Questionnaire design and administration

As pointed out in section 2, most of the people interested in statistical education research, of whom statistical education researchers are a subset, are likely to have access to the internet which suggests that the internet is an appropriate medium for collection of information. It tends to be both quicker and cheaper than the more traditional methods of collecting information by interviewers or by post, and enables easy conversion to data files. To some extent use of the internet for conducting surveys is still in its infancy, but it is attracting interest in the market research industry, although little has yet been published regarding internet-based survey methodology (Ohsumi and Yoshimura 1999, Witmer el al 1999).

The initial decision is between a Web-based survey and an email survey. Given that initial contact with potential respondents is planned to be by email and that additional effort is needed to visit a Web page, an email survey would a suitable next step. A further choice is between including the questionnaire as a plain text message or as an attachment. The former tend to look dull, whereas questionnaires sent as attachments can look more professional and convert more easily to database format. As the survey is under the auspices of the LTSN the questionnaire will be in English although this could mean the exclusion of some groups of respondents. The design would correspond to databank entries.

The information required from respondents will be mainly factual – contact details, and a list of what they consider to be their statistical education research publications, classified according to a predefined list of categories. They will also be asked about their future research plans, and their agreement to providing updates will be sought. It would be interesting to ask them also what they thought constituted research in statistical education. A short questionnaire asking mainly for information which is already in the public domain might be expected to have a high response rate, especially as it is likely that researchers would want to be included on the database and to provide updates without undue prompting as it would be in their interests to publicise their work. However, the time needed to input responses could be considerable. In any case there is a tendency these days to be overwhelmed by the volume of incoming emails, and busy people are just as likely to postpone responding to email questionnaires or to ignore them as they are in the case of paper ones. In a survey of members of the Teaching-statistics email list, many of whom are statistical education researchers, the response rate was about 27% (Green and Fuller 1999), which is on the low side.

5. The survey and the future of statistical education research

The internet has made the world a smaller place, and has made it more feasible for researchers in statistical education to be an interactive community. It is in fact essential that researchers take full advantage of the newest technological developments. Statistical education in itself is undergoing substantial and rapid changes and it is important that research is done into the effects of these changes and that those concerned with providing education and assessing the outcomes are familiar with the research. By making publication details of the research readily accessible, the proposed survey and the resulting database will have an important contribution to make as regards the future of statistical education and of statistical education research.

REFERENCES

Davies, N. (2000) Promoting effective methods of learning and teaching statistics and operational research. MathStats&OR, 2, 2-3.

Garfield, J., Gal, I. and Chance, B. (2000) Bibliography on assessment. In Statistical Education Research Newsletter. http://www.ugr.es/~batanero/sergroup.htm

Green, D. and Fuller, M. (1999) Analysis of the Teaching-statistics mailbase list. Teaching Statistics, 21, 2, 56-59.

Jolliffe, F.R. (1998) What is research in statistical education? In Proceedings of the Fifth International Conference on Teaching of Statistics (eds L. Pereira-Mendoza, L.S.Kea, T.W. Kee, W-K Wong), 801-806. International Association for Statistical Education. Singapore.

Kalton, G. and Anderson, D.W. (1986) Sampling rare populations. Journal of the Royal Statistical Society Series A, 149, 65-82.

Ohsumi, N. and Yoshimura, O. (1999) The online survey in Japan, an evaluation of emerging methodologies. In ISI Invited papers proceedings, IASS topics, Helsinki August 10-18 1999, 117-133.

Sahai, H., Khurshid, A. and Misra, S. C. (1996) A second bibliography on the teaching of probability and statistics. Journal of Statistical Education, 4, 3. http://www.amstat.org/publications/jse/v4n3/sahai.html

Witmer, D. F., Colman, R.W., Katzman, S. L. (1999) From paper-and-pencil to screen-and-keyboard. Toward a methodology for survey research on the Internet. In Doing Internet research: critical issues and methods for examining the Net (ed. S. Jones), Sage publications, Thousand Oaks, London, New Delhi.

RÉSUMÉ

Les chercheurs en didactique de statistiques trouveraient utile une banque de données concernant leurs recherche. On y trouverait, par example, les adresses des autres chercheurs dans la même discipline, et les détails des publications. L'auteur examine l'obtention d'information par internet pour une enquête internationale, et propose que la banque de données soit placée sur le Web.