SAMPLE SURVEYS VIA ELECTRONIC MAIL: A COMPREHENSIVE PERSPECTIVE

Zeki Simsek
Research Assistant at Bogazici University, Istanbul, Turkey.

RESUMO
Apesar de os questionários enviados pelo correio, as entrevistas realizadas pessoalmente e as por telefone serem os métodos principais de pesquisa, tem havido aumento na utilização de e-mail como um meio de coleta de dados. No entanto poucas pesquisas publicadas no Ocidente em geral e na Turquia em particular avaliaram a técnica de pesquisas por e-mail a partir da perspectiva pura de pesquisa. Na tentativa de desenvolver um instrumento para avaliar o e-mail como um meio de coleta de dados, o propósito deste estudo é explorar a técnica de pesquisa por e-mail a partir de ângulos complementares. Para alcançar esse objetivo, são discutidas a representatividade da amostra, a qualidade dos dados, as taxas de resposta e as vantagens e as desvantagens da pesquisa por e-mail.

ABSTRACT
Although postal questionnaires, personal interviewing, and telephone interviewing are the main methods of survey-based research, there is an increasing use of e-mail as a data collection medium. However, little, if any, published Western research in general and that of Turkish in particular have investigated e-mail survey technique from pure survey research perspective. Attempting to develop a framework to assess e-mail as a data collection mean, the purpose of this study is to explore e-mail-based questionnaire technique from complementary angles. To this goal, sample representativeness, data quality, response rates, and advantages and disadvantages of e-mail surveying are discussed.

PALAVRAS-CHAVE
Método de pesquisa, e-mail, coleta de dados.

KEY WORDS
Survey method, e-mail, data collection.
INTRODUCTION

Primary data, that is data generated exclusively for the research project at hand, can be obtained by either performing an experiment or conducting a survey. The survey method of gathering data relies mainly on the questioning of a representative sample of individuals to elicit particular information, and survey methods are usually classified by mode of administration as telephone interviewing, one-to-one interviews, mall-intercept interviews, mail questionnaires, and so on. Even though data generated through experimentation are considered better and despite the well-known inherent weaknesses of survey research, during the last several decades, gathering data by conducting surveys has been more widespread mainly for such reasons as costs and difficulties involved with carrying out experiments. If applied properly, surveys allow researchers collect inexpensive, quick, and effective data with less effort due to simple administration of questionnaires. Use of sample surveys has been so pervasive that most knowledge of societal trends and lifestyles comes from studies using this method of data gathering (Synodinos & Brennan, 1988), and the survey approach has been by far the most common method of primary data collection in marketing research (Malhotra, 1993, p. 189). This growth of survey research is related to the simple idea that to find out what people think, one should ask them (Zikmund, 1994, p. 171).

Albeit sketchy, the history of surveys can be traced back to thousand of years. However, mail questionnaire, personal interviewing, and telephone interviewing had for a long time been the only available methods to collect survey information prior to technological advancements taking place in the last half of 20th century, which have greatly changed ways and methods of collecting data. Especially parallel to the growth of computer and telecommunication technologies, a dramatic growth of survey research and techniques has been seen for particularly the last two decades. Attempting to explain the phenomenon, Gates and Jarboe (1987) reported that three factors have contributed to the change. These are developments in electronics technology, computer software, and environmental forces that oppose traditional data collection methods. With the evolution of data collection techniques, the computer has been adopted to improve data collection efficiency and minimize some of the inherent problems and biases in paper and pencil survey techniques (Malhotra, 1993, p. 404).

Indeed, one of the early applications of computer technology to research was in the area of data collection (Neal, 1989), and, along with making survey research more scientific and accurate, computer applications have introduced efficient and effective ways to gather and analyze data, while permitting researchers to study matters and questions that would not otherwise be possible (Anderson & Gansender, 1995; Saltzman, 1995). Computer-assisted personal interviewing, computer-assisted self-interviewing, fully-automated telephone interviewing, computer-disk-by mail (DISKQ), and computer-generated fax surveys are only a few new surveying methods that depend on computer technologies to mention. In fact, playing an ever increasing role in survey research, computers can be used for all phases of a survey, namely instrument design, sampling, field monitoring, coding and data editing, data capture, data cleaning, scale-index construction, data base organization, data base retrieval, data analysis, and documentation (Karweit & Meyers, 1983).

The utilization of e-mail in survey-based research is also a result of rapid developments in computer and communication technologies. E-mail’s speed and low cost coupled with the recent exponential increase in the usage of computers have led e-mail to be debated as a new mean of conducting sample surveys. Several researchers have conducted surveys by way of e-mail to understand what potentials of e-mail questionnaires are in collecting data. Nevertheless, the research for this study revealed that the literature on e-mail surveys mostly consists of empirical studies dealing with one aspect of the e-mail surveying and, thus, lacks a comparative perspective of the issue. As will be shortly discussed, some of the few studies only touch upon response rates of e-mail surveys while the rest deal exclusively with quality of data collected. Intending to fill the research gap, this study presents the four most important concepts of sample surveying as they apply to e-mail survey technique. These are sample representativeness, data quality, response rates, and advantages and disadvantages of e-mail surveying. However, before going further, a definition and background of e-mail within the context of this study shall firstly be presented because e-mail describes many new forms of communication that depend on computer and communication technologies.

A LOOK AT E-MAIL

The Electronic Mail Association (EMA), a Washington-based trade association, defines e-mail as “the generic term for the non-interactive communication of data, images or voice messages between a sender and designated recipient(s) by systems utilizing telecommunication links”. Following the definition, one evidently can combine e-mail with such technologies as facsimile, telex, communicating word processors, and similar technologies. Nonetheless, in this study the e-mail referred to is the form called text-e-mail, which can be described as transformation of
An e-mail survey can be defined as a self-administered questionnaire sent respondents through e-mail. Kiesler and Sproull (1986), two pioneers in the field, suggested that the utilization level of e-mail in survey research depends on its comparability to the other available methods of data collection. Based on Kiesler and Sproull’s suggestion, the researcher proposes that e-mail sample surveys be evaluated along four dimensions. These are the formation of a representative sample, the quality of the data, response rate, and the advantages and disadvantages of e-mail questionnaires. Indeed, for any particular research project, the relative importance attached to these factors will vary.

**FORMATION OF A REPRESENTATIVE SAMPLE**

It is well known that the random sample selected for any sample-survey research must typify or represent true population. In fact, for a survey that is free of respondents’ error (response effects), a survey and decisions based on the survey data usually can be as good as the sample selection error involved in the survey. Stated simply, a survey that does not embody the target sample will ultimately be a waste of time and resources. At this point, one may wonder about the connection between a survey method and representation of a sample. It is an obvious one. Other things being constant, the greater the coverage allowed by a survey technique, the more valid the results because more representative samples are possible. Moreover, responses to survey questions can be influenced by the mode of survey has, for a long time, been recognized by survey researchers (Ayidiya & Mckee, 1990).

Can an e-mail questionnaire allow researcher reach a representative sample? Unfortunately, experiences have shown that the answer to the question is rather a complex one. Because relevant population for an e-mail survey is restricted to those having access to e-mail, a subject population’s e-mail accessibility level may be considered as an answer to the question. Nevertheless, it should be borne in mind that the availability or accessibility of e-mail does not guarantee acceptance or usage (Komsky, 1991). Moreover, although there are some mixed statistics on e-mail communication traffic, concentrating on such aggregate numbers is not too helpful if only the problems of incompatibility and low sophistication level...
among various e-mail packages and systems still persist. Incompatibility, for example, can cause discrepancies between the form of questionnaire sent and that received by respondents. Similarly, some older versions of e-mail do not feature a “response” function, which allows respondents to respond an e-mail questionnaire with one computer key (Oppermann, 1995).

Constituting a representative sample for an e-mail survey is further complicated by the fact that all computer users still share similar demographic characteristics.

However, with time, as the novelty of e-mail fades, reactions of computer users toward unsolicited e-mail messages change, and e-mail boxes receive more than a questionnaire daily, it is very likely that researchers employing e-mail will have to find some ways of increasing response rate.

Currently, most computer users appear to be young and well educated and to have above average incomes. Oppermann (1995) reported that college graduates were 2.5 times as likely to have a computer than were high school graduates, and households with an income of more than US$ 50,000 were about five times more likely to have PCs than were those earning below US$ 20,000. The implication of these numbers is that e-mail surveys of heterogeneous groups such as households with simple random sampling can be flawed because of the non-coverage error. The non-coverage error exists when some members of the population are not covered by the sampling frame and thus have no chance of being selected into the sample (Dillman, 1991). Some of the researchers have concluded that in the immediate future it is unlikely that most electronic users can be considered as representative of the general population. It has rightly been suggested that a stratified sampling scheme can be a better option than random sampling to lessen degree of potential non-coverage error associated with e-mail surveys (Oppermann, 1995).

In my judgment, nonetheless, the decision on whether surveying through e-mail is appropriate should be better left to individual surveyor, for in reality sample surveys vary in content and thus in sampling frame. For instance, a marketing researcher might find e-mail sample surveys quite beneficial for a task, say getting feedback related to a new software, or a product targeted to affluent and young populations. Moreover, a stratified sampling scheme, instead of random sampling, can always be employed to lessen degree of potential non-coverage error might associate with e-mail surveys (Oppermann, 1995).

QUALITY OF THE DATA

Obtaining complete and accurate responses from respondents is, of course, the main aim of any survey research (Frey, 1983, p. 44). Data quality that e-mail surveys yield has been investigated with the following empirical studies although the issue is still in its infancy.

In their empirical study, Kiesler and Sproull (1986) found less item incompleteness with the e-mail survey than with the paper mail and similar responses between paper and e-mail, but more socially undesirable responses in the e-mail survey than the paper questionnaire. In another study, Sproull (1986) found that the questionnaire data were slightly more comprehensive in the conventional study (mail) than in the e-mail study, and there were no differences in the nature of answers provided by the participants of the two methods. Difference in extremity of responses was, however, found as the e-mail survey elicited more extreme responses than the paper survey. Kiesler (1989) stated that because e-mail communication loosens social concerns and constraints on people, respondents are less concerned about making negative statements and revealing socially inappropriate truths. Ayidiya and Mckee (1990) also suggested that e-mail might lessen the so-called acquiescence of participants. E-mail communication might slow some respondents’ propensity to agree with survey statements more often than they would choose under normal circumstances. Like postal surveys, e-mail surveys can promote less acquiescence because of the absence of surveyors and total control of respondents over the timing of his/her response. Computer-mediated communication generally conveys little social information, so respondents experience less evaluation anxiety than when they respond in the other methods of surveying such as personal or telephone interviewing (Sproull & Kiesler, 1986; Kiesler, Zubrow & Moses, 1985).

RESPONSE RATE

A survey’s response rate can broadly be defined as the percentage of the total attempted questionnaires or interviews that are completed (Malhotra, 1993, p. 201). As with any survey method, an e-mail survey can be negatively affected by low response rates, especially when underlying population is not uniform in the characteristics. Low response rate surveys, in which those who do not respond are different from those who do, lead to the development of less valid inferences, not mentioning...
unrecompensed costs. In fact, what is known as non-response error, the statistical difference between a survey that include only those who responded and a survey that also includes those who failed to respond (Zikmund, 1994, p. 172), is one of the commonly perceived negatives associated with the mail questionnaire technique (Koff, 1992). The magnitude of non-response bias increases as the response rate decreases and thus the general assumption is that the higher the response rate, the lower the potential of non-response error and, therefore, the better a survey is.

As far as response rate is concerned, e-mail survey has been used as an effective means of gathering data in academic, scientific, and business contexts (Anderson & Gansender, 1995). Researchers using e-mail questionnaires have reported response rates ranging from 19.3% to 76%.

Kiesler and Sproull (1986), examining the response rate associated with e-mail survey vs. postal survey, found a higher response rate for the paper survey (75% vs. 67%). Sproull (1986) compared e-mail survey with face-to-face interviewing. The study found a participation rate of 73% for the e-mail survey and 87% for interviewing. Rafeli (1986) achieved a response rate of 48% for an on-line survey and 82% for a mailed survey sent to 172 users of a bulletin board system. Komsky’s 29-item on-line survey (1991) generated a response rate of 41%. Parker (1992) reported on a major corporation’s use of an e-mail survey to gather data from its employees who were working overseas. The response rate associated with the e-mail survey (68%) was significantly higher than those of the mail pouches (38%). Walsh et al. (1992) attained a response rate of 76% from a 93-item on-line survey of 300 oceanographers. Schultd and Totten (1994) had a response rate of 56.5% for a mailed survey and 19.3% for an e-mail survey. Kittelson (1995) obtained a response rate of 28.1% for an e-mail survey versus 76.5% for a postcard survey. Finally, Anderson and Gansender (1995) captured a response rate of 76% from a sample of 488.

Comparing to the typical low response rate of regular mail surveys, the response rates associated with e-mail method are, indeed, very promising. This is especially so when one takes into account that most of the responses were attained without using any response inducement technique. However, with time, as the novelty of e-mail fades, reactions of computer users toward unsolicited e-mail messages change, and e-mail boxes receive more than a questionnaire daily, it is very likely that researchers employing e-mail will have to find some ways of increasing response rate. The strategy to increase responses in e-mail questionnaires is similar to that of mail questionnaires, that is to encourage respondents to reply, but the issue of augmenting the response rate of e-mail surveys through some incentives and procedures is yet to be explored.

ADVANTAGES AND DISADVANTAGES

Each survey technique has its merits and shortcomings, and so has an e-mail survey. First, surveying through e-mail has the potential of radically changing the economics of conducting surveys. The cost of sending and receiving a questionnaire through e-mail is very low. According to an estimate, the marginal cost of storage, communications, and dissemination of a thirty page document can be less than a penny (Kambil, 1995). Moreover, while the cost of most surveying methods is prone to be linear, proportional to the size of the sample under the study, an e-mail survey usually has a fixed cost, enabling a surveyor to enlarge the sample if necessary. Further cost savings are also realized in terms of reproduction and printing costs of questionnaires. In fact, as a cost reduction tool in surveying, e-mail can be employed to send prior-notifications or follow-ups, the common motivational variables for increasing low response rate of postal surveys.

Several companies recently have introduced survey-software packages that work with e-mail systems to create, collect, and tabulate survey results.

Next, sending questionnaires out and receiving returns is definitely very fast with an e-mail questionnaire. If desired, a forward-chain with e-mail can be used to send all questionnaires out simultaneously. An e-mail questionnaire will be available to potential respondents in most cases almost immediately, no matter time or location of potential respondents. None of the existing survey techniques, including the faxing method, can provide researchers with such a speed in reaching most of the respondents. The delivery of an e-mail questionnaire is almost certain; when the delivery is not complete, in a moment the researcher is able to see it. Therefore, identification, elimination, and replacement of unreachable respondents are quite easy with e-mail surveys. This is especially important when a survey is targeted to a mobile population or fast data are needed.

The asynchronous nature of e-mail communication is another positive feature of e-mail surveying. Unlike phone or personal interviewing, e-mail does not require the respondents to be instantly accessible; consequently, the survey can be read and responded to at the respondents’ convenience. This, in turn, might increase the quality of responses by giving respondents the freedom to think about their propensities and beliefs.
Also, e-mail surveys do not require usage of paper at any stage, except for coding the respondents’ answers for analytical purpose unless there is a pre-loaded software for this specific task. Less contact with paper is not only cost efficient but also, to some extent, ecologically sound. By allowing surveyors to encode the data without transcribing from paper, e-mail questionnaires can eliminate numerous hours of tedious work (Dacko, 1995) and reduce potential respondents mistakes such as deciphering respondents’ handwriting.

On the other hand, survey research through e-mail has its unique drawbacks as well. First, at the moment there is no comprehensive listing of people having e-mail boxes, it is difficult to conduct large scale and not within-organization surveys. Although there are several utilities, such as Finger, Whois, Netfind, to collect the Internet e-mail addresses, it has been estimated (Parker, 1995, p. 286) that only one or two per cent of all the Internet users can be located with one of these methods, because the majority of users are not on systems that allow these utilities to identify them. A good way to obtain e-mail addresses of the general population could be getting them from such on-line companies as CompuServe, America Online, and Prodigy or commercial e-mail carriers such as MCI or AT&T. However, because these companies have to be concerned with the privacy of their customers, and some of the companies have announced that they can perform specific surveys, it then follows that this is not a viable option.

Another major hindrance of employing e-mail surveys is that e-mail is still not a standardized tool, despite growing demand of users for standardization. For instance, LAN e-mail packages are different not only from one another but also from the Internet e-mail. Before the change was made, a company-wide survey revealed that a major company was using 22 different e-mail installations (Mayor, 1995). As mentioned earlier, non-standardization of e-mail systems can cause discrepancies between the form of questionnaire sent and that received by respondents. For example, if one sends out an e-mail questionnaire in binary, unless conversion is possible, an Internet e-mail user respondent may not receive the survey because the Internet e-mail packages accept only ASCII format. Similar technicalities further affect the questionnaire design and the usage of visual aids with e-mail surveys. For example, the Internet e-mail is normally text-based: what makes difficult, for instance, to send a multimedia file as an attachment. Finally, because e-mail responses can not be anonymous, the execution of e-mail surveys dealing with such sensitive issues as politics and organizations may be very difficult.

Table 1 compares e-mail-surveying technique with traditional data collection methods across a variety of variables. It should be noted that the emphasis is on the typical, and for any particular research project these assessments may not hold true. Also, because e-mail surveying is still at an early developmental stage, this study’s assessments of e-mail survey technique may

| Table 1 - A comparison of e-mail surveying with traditional surveying methods |
|---------------------------------|-----------------|-----------------|-------------------|------------------|
| Survey Type                     | Factor          | Personal interviews | Telephone interviews | Mail questionnaire | E-mail questionnaire |
| Cost                            | Very high       | Moderate to low    | High               | Moderate          | Lowest             |
| Data collection speed           | Moderate to low | High              | Moderate           | Moderate          | Fast               |
| Response rate                   | Very high       | Moderate           | Moderate           | Low               | Moderate to high   |
| Geographic flexibility          | Limited         | High              | Moderate           | Very high         | Highest            |
| Time flexibility                | Limited         | Moderate           | High               | High              | Highest            |
| Diversity of questions          | High            | Moderate to low    | Moderate to long   | Moderate          | Moderate           |
| Questionnaire length            | Long            | Moderate to short  | Moderate to long   | Moderate to short | Moderate to short  |
| Use of visual aids              | Highest         | None              | Usually high       | Low               | Low                |
| Item non-response               | Low             | Moderate           | Moderate to high   | Moderate to low   | Low                |
| Ease of prior-contact           | Difficult       | Easy              | Easy               | Easiest           | Easiest            |
| Ease of follow-up               | Difficult       | Easy              | Easy               | Easiest           | Easiest            |
| Respondent anonymity            | Lowest          | Moderate           | High               | High              | High               |
| Potential interviewer bias      | High            | Moderate           | None               | None              | None               |
| Respondent convenience          | Low             | Moderate           | High               | High              | High               |
| Eliciting sensitive data        | Low             | Moderate to high   | High               | High              | High               |
| Control of field force          | Low             | Moderate           | High               | High              | High               |
| Sample control                  | High            | Moderate to high   | Low                | Moderate          |                     |
| Difficulty of carrying out      | Very difficult  | Difficult          | Easy               | Easy to medium    |                     |
contradict with those of research when more substantial data are available, or when e-mail systems become more versatile bringing about changes at the very structure of e-mail communication.

CONCLUSIONS AND RECOMMENDATIONS

E-mail can provide a good opportunity for those researchers who have a limited research budget or who are interested in fast data gathering. E-mail is cheap, fast, and its asynchronous nature can lead to better response quality. Because e-mail obliterates time and zone constraints, surveying with e-mail can prove to be very beneficial when the sample is scattered and mobile. At present, lack of information on e-mail addresses and non-standardization of e-mail packages inhibit e-mail’s full potential. Especially when the research project involves sample surveys of heterogeneous populations such as households, usage of e-mail may not be appropriate because of the high non-coverage risk. Accordingly, it is too early to declare that the e-mail technique has become a rival or a better technique than major non-computerized data collection techniques.

As it is mentioned throughout the study, the usage of e-mail in sample surveys and research on e-mail surveys alike are still in their infancy, and, thus, there is a great need for more research. Additional studies must be undertaken to reveal the possibilities e-mail has to offer to survey-based research. As increasingly decreasing postal questionnaire response rates attest, the popularity of e-mail as a survey medium will almost certainly have an impact on the willingness of respondents to participate in e-mail questionnaires. Thus, some studies need to be carried out in order to find some ways of augmenting e-mail survey response rates. E-mail surveys have so far mostly been used within organizations and educational institutions, so it will be interesting to see studies which experiment with e-mail surveys upon different populations. Another avenue for future research can be the assessment of the extent of non-coverage risk associated with e-mail survey technique across various populations.

As the unit cost of computers falls, the number of e-mail service providers increases, and e-mail packages become more compatible, e-mail questionnaires will very likely be more and more used in sample surveys. Several companies recently have introduced survey-software packages that work with e-mail systems to create, collect, and tabulate survey results. Software packages plus improvements in e-mail usage can perhaps in the not too distant future make “virtual surveyor” a reality if legal regulations taking place between now and that time do not prohibit usage of e-mail for surveying purposes.

REFERENCES

KIESLER, Sara. E-mail brings out the best and worst in group interaction. Infoworld, p.4-9, 1989.
MAYOR, Tracy. Universal e-mail. CIO, 88, 1995.
PARKER, Tim. How the Internet e-mail works. Using the Internet, 2nd ed. 1995.