The Challenge of The World Wide Web:
A Comparative Analysis Model of Web Survey Application with Other Modes of Information Collection in Research Participation

Noppadon Kannika
Graduate School of Computer and Engineering Management,
Assumption University of Thailand,
Hua Mark, Bangkapi,
Bangkok, Thailand 10240
E-mail: noppadon@ksc.au.edu

Abstract

The research participation in Web survey design of students at Assumption University of Thailand is much more complex than other modes of information collection. The important findings in this research release were that Web survey produced high non-response rates because most of the students were not well informed about the Web survey project due to their invalid email addresses and they did not use their university email accounts. However, the finding released that Web survey was more likely to produce 'Yes' answers of the sensitive questions from the respondents than other modes of information collection. I, therefore, can conclude conditionally that Web survey is more likely to produce high quality of research data than other modes of information collection in the sensitive research projects. Nevertheless, Web survey, in comparison to the other modes of information collection, needs more future research to find strategies to get respondents to participate in the research on Web, to motivate them to complete the Web questionnaire, and to understand more about the factors that affect the respondents' research participation on the Web.

Keywords: comparative analysis model, information technology, research process on WWW, Web survey design, Web survey implementation, response rate, non-response, research participation.

1. INTRODUCTION

Today mankind has the accelerator pedal of technology pressed hard to the floor. (Martin, J., 1996) Survey research is almost completely a set of linked information processing tasks.   A world of rapidly changing information technology presents us with both the ways of capitalization on new technologies and the historical principles on which the survey research profession rest. (Reginald P. Baker, 1998) Especially, technological developments are providing new options—unthinkable a few years ago—for self-administered surveys. (Don A. Dillman, 1999) E-mail, fax, the Web, touchtone data entry—all of which are self-administered—have become feasible survey methods. (Clayton, 1998) Web survey research presents the survey mode which is based on computer-assisted, self-administered questionnaires answered without the presence of the interviewer. The questionnaires are based on HTML forms usually presented in standard Web browsers, and the responses are immediately transferred through electronic networks, usually the Internet. Web questionnaires have been placed on thousands of Web sites and Internet survey panels now include millions of Internet users. (Vasja Vehovar, 2002) While the proponents of Web surveys point to the astonishing growth rate of the Internet as a reason for their optimism regarding Web surveys of the general public, there remain major concerns regarding

the future penetration of this technology. There are many different estimates of the penetration of the Web. (Mick Couper, 2000) However, information technology continues to improve all aspects of doing Web surveys: standardized software, user-friendly interfaces, attractive multimedia, high speed of transmission, and low access costs. In the years ahead, increased Internet penetration, massive Web usage, and technological improvements will further expand the application of Web surveys. (Vasja, 2002) The rapid development of surveys on the World Wide Web (WWW) is leading some to argue that soon Internet (and, in particular, Web) surveys will replace traditional methods of survey data collection. (Mick Couper, 2000) The traditional collection of information in surveys has undergone a transformation in the means by which data are gathered. During the period 1940 – 1970, Mail surveys and face-to-face interviewing provided the primary mechanisms for collecting data. The increased availability of and access to telephones, improved sampling methods, and improved technology have resulted in a substantial increase in the use of the telephone as the means by which information is gathered in sample surveys. (Lars Lyberg, 1991) For reason of cost and ease of implementation, paper-and-pencil mail surveys are more frequently used for social research than are either telephone or face-to-face interviews. However, the last two decades of research aimed at improving mail survey methods are examined. Discussion of this research is organized around progress made in overcoming four important sources of error: sampling, non-coverage, measurement, and non-response. Little research in this area has been conducted. (Don A. Dillman, 1999) Face-to-face interviewing is the mode in which an interviewer administers a structured questionnaire to a respondent within a limited time period and in the presence of the respondent. Research during the last decade has challenged the once strongly held view that face-to-face interviewing obtains the best quality data. The interviewer, by virtue of his or her presence, may motivate and encourage respondents to answer the survey questions as well as probe for more complete and accurate responses. A risk exists, however, that the quality of data obtained in face-to-face interviews can be affected by the presence of members of the household who are not part of the study. Another risk, labeled *social desirability bias*, may exist in this data collection environment. (Lars Lyberg, 1991) Surveys which use telephone interviewing are less expensive and often proceed more rapidly than face-to-face interviewing. These surveys can often be implemented rapidly with a smaller staff than face-to-face interviewing. (Bradburn, 1983) However, one way to summarize the status of telephone survey methodology is to review its major known strengths and weaknesses, focusing on coverage error, non-response error, sampling error, and measurement error properties. (Groves, 2002) The utility of a data collection technique does not only depend on costs, speed, and response rate, but also on the quality of the data collected (Herman, 1977)

Therefore, in this research, the comparative analysis model is applied for exploring the effects of different modes on the survey participation and information quality. The principles of comparative analysis model is the same as those governing "ordinary" survey research. (De Leeuw, E.D. 1987) Typical steps taken in the comparative analysis model are: a precise definition of the research problem, data collection (i.e., collection of relevant articles or papers), coding of the variables of interest, and statistical analysis (Wolf, F.M., 1986) This research, in the comparative analysis model, will be empirically compared the design approaches of Web survey to other modes of data collection in survey research including mail survey, phone survey, and face-to-face interview.

2. RESEARCH OBJECTIVES

In this study, a comparative analysis model of Web survey application with other modes of information collection in research participation, researchers aim to achieve the research objectives including 1) To study the comparative analysis model of Web survey application with other modes of data collection method in survey research regarding the survey participation, data quality, and survey measurement, and 2) To develop the applications for the design approaches of Web surveys in both social science and information technology.

3. RESEARCH PROCESS ON WWW

Web surveys are proliferating at a rapid pace. Despite this increase, we have yet to see much empirical research aimed at exploring various features of Web survey design. (Mick P.Couper, 2001) In this paper, the researcher outlines the research design on world wide web (www) as shown in the figure 1.
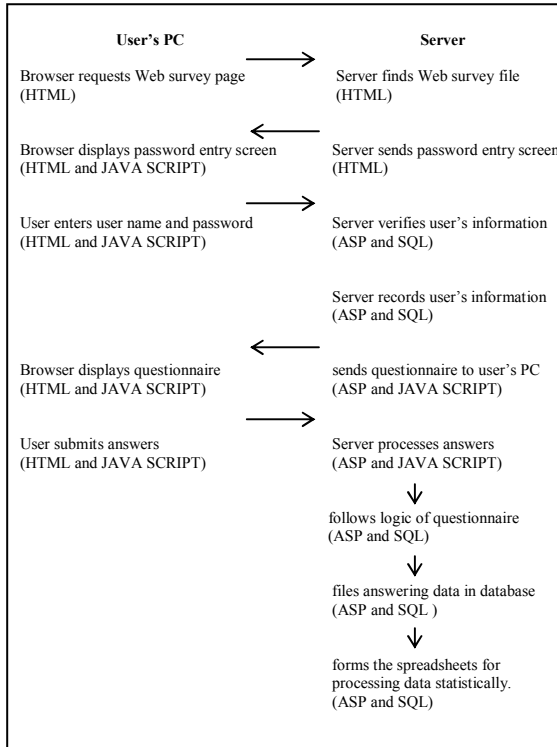
Figure 1. Research Process on World Wide Web
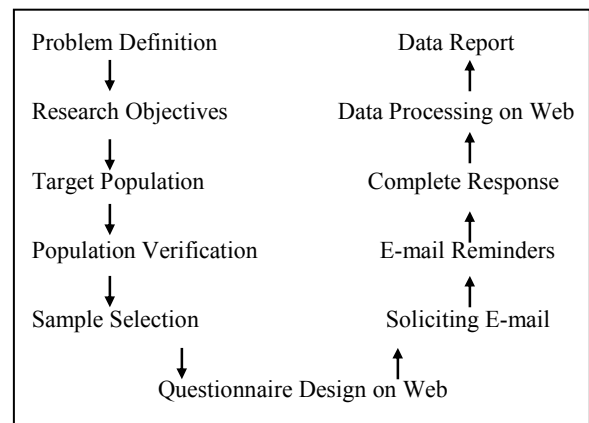
4. WEB SURVEY IMPLEMENTATION

The survey research was designed on the world wide web. There are different programming languages and styles that one can use for building a Web questionnaire,

some of which are quite sophisticated. (Don Dillman, 2000) In this study, every element of the target population has an email account and free access to the Web which makes the use of the Web for surveying attractive. In figure 2, researcher presents how Web survey works.

Figure 2. Web Survey Work Process

| User's PC	Server
Browser requests Web survey page (HTML)	Server finds Web survey file (HTML)
Browser displays password entry screen (HTML and JAVA SCRIPT)	Server sends password entry screen (HTML)
User enters user name and password (HTML and JAVA SCRIPT)	Server verifies user's information (ASP and SQL)
	Server records user's information (ASP and SQL)
Browser displays questionnaire (HTML and JAVA SCRIPT)	sends questionnaire to user's PC (ASP and JAVA SCRIPT)
User submits answers (HTML and JAVA SCRIPT)	Server processes answers (ASP and JAVA SCRIPT)
	follows logic of questionnaire (ASP and SQL)
	files answering data in database (ASP and SQL)
	forms the spreadsheets for processing data statistically. (ASP and SQL)

5. RESEARCH METHODS

This research is based on the specific target population of Thai students at Assumption University. Cooperation from the Registrar made it possible to apply a simple random sampling (SRS) to students to be drawn by samplers at the survey research center at Assumption University. The total designed sample size was 2,500 students. The equal size of the sample allocation was determined by 500 students per each group including Web survey group 1, Web survey group 2, Traditional Mail survey, Paper-Pencil (PP) Phone survey, and Paper-Pencil (PP) Face-to-face interview. Data collection was conducted from June 25 to July 10, 2002. The researcher also conducted the post research by phone for interviewing the non-respondents in the Web survey projects. Some statistics were applied to the data analysis including frequency, percentage, mean, standard deviation, variance, correlation, and other appropriate statistics. SPSS/FW was used for data processing.

6. RESULTS

Figure 3 shows the research participation behavior of respondents which was classified by the timing of the study. According to the rule of thumb for response expectations (Scott Crawford, 2002, Course Notes), the figure illustrates a sharp increase at 30.2% out of the total web survey participants who responded during the first 24 hours of data collection by Web survey before declining to 14.2%. However, by using the reminder email, the result shows the leveling off in research participation behavior.

Figure 3. Response Behavior on Web Classified by Timing of the Study (percentage of total respondents who participated with Web survey
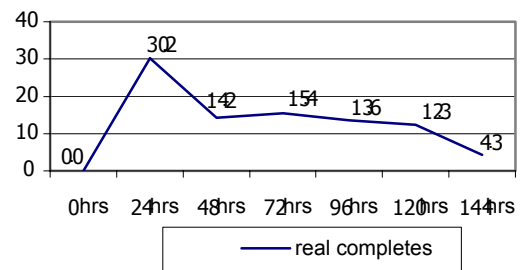
When the comparative analysis model of Web survey and other modes of data collection in research participation is used, table 1 shows that the response rate of the phone survey (81.2%) is higher than those other modes of data collection.

Table 1. Response Rate Classified by Modes of Information Collection

Modes	Complete		Partial		None	
	n	%	n	%	n	%
Mail	198	39.6	22	4.4	280	56.0
Phone	406	81.2	15	3.0	79	15.8
Face-to-face	203	40.6	10	2.0	287	57.4
Web 1	88	17.6	16	3.2	396	79.2
Web 2	47	9.4	13	2.6	440	88.0

However, the Web and Mail surveys are more likely than other modes to produce "Yes" answers to sensitive questions as shown in figure 4. A socially desirable answer is one influenced by the respondent's concern over how he or she will be viewed by the interviewer, according to social norms. In mail surveys, the respondent does not have to be concerned about how his or her answer is being viewed by someone with whom he is interacting. (Don A. Dillman and others, 1996)

Figure 4. Effect of Information Collection Modes on the "Yes" answers of Sensitive Questions (base on total respondents of each mode)
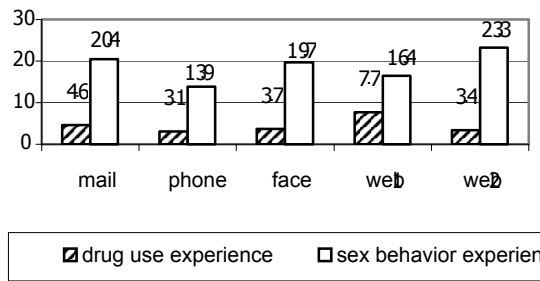

Table 2. "Yes" answer on Sensitive Questions Classified by Self-Administered and Interviewer-respondent Modes (base on total respondents of each mode)

	Drug Use Experience	Sexual Behavior Experience
Self-Administered Modes (Web and Mail)	5.2	19.8
Interviewer-Respondent Modes (Phone and Face-to-face)	3.4	15.8

Thus, one would expect that respondents in interviewer-respondent modes or telephone and face-to-face surveys are more likely to provide socially approved answers and less willing to admit to behaviors that are viewed negatively by society, and presumably by the interviewers as well. The experimental research literature has repeatedly found that many research participants tend to portray themselves in the most positive light possible (Dovidio and Fazio, 1992). An early study by Hochstim(1967), in which women in a face-to-face interview were more likely to say their health was "excellent" (44%) compared to women in a telephone interview (37%) and a mail survey (30%), illustrates the tendency to overstate a socially approved answer. Researchers have continued to find evidence to support the notion that the mail survey respondent is least likely to be influenced to give a socially desirable response (compare DeMaio, 1984; de Leeuw and van der Zouwen, 1988). Compared to this latest research, those samples who participate in the self-administered surveys, including Web and mail modes, are relatively higher than the interviewer-respondent surveys in producing "Yes" answers to the sensitive questions.

7. POST RESEARCH OF WEB SURVEY

As shown in table 1, the research findings show that Web surveys encounter the problem of unit non-response at 88.0% in Web survey group 2 and 79.2% in Web survey group 1. The following research findings in this section are released by the post research of Web survey unit non-respondents.

Table 3. Factors affecting on the unit non-respondent to Web survey (Multi-responses)

Factors	%
1. Have not received invitation email yet	52.7
2. Have not checked email yet	27.4
3. Cannot approach the question web page	8.1
4. Ambiguous passwords and identification process	6.2
5. Limitation of user's Internet access	5.6
6. Low attractive incentive	1.9
7. Low attractive web page	0.6

From table 3, the research findings show that 52.7% of non-respondents have not received the invitation email yet due to the wrong email address and that most of the target users do not use the university email account. This is followed by 27.4% out of the total non-respondents who have not checked their email yet, 8.1% who cannot approach the question web page, 6.2% became the non-respondents due to the ambiguous passwords and identification process, and 5.6% who have some limitations to Internet access, respectively. The user's technical equipment affects Web survey participation. For certain Web surveys, only the technologically advanced users with the latest version of browser, higher speed of Internet access, and better PC platform and monitor can participate. Inadequate equipment also makes the Web survey longer, unpleasant, difficult or even impossible (Vasja Vehovar, 2002). However, there are some important factors which affect the research participation among those respondents. The research results illustrate that the response rate to web surveys usually varies across some characteristics of Internet users as shown in Table 4.

Table 4. Internet Heavy User and Research Participation on Web

Internet Heavy User (several times per week)	Research Participation on Web	
	Yes %	No %
Yes	91.0	9.0
No	66.4	33.6

p=0.0000

From table 4, the research findings show the significant difference between the Internet – heavy users and non-heavy Internet users. 91.0% of the total sample of Internet heavy users participate the research on the Web, while 66.4% of the total sample of non – heavy Internet users participate the project. In general, research participation depends strongly on the survey topic, its salience, and the respondent's involvement. For Web survey in particular, intensive users with heavy use and strong attitudes toward Internet issues are more likely to

participate in Internet use surveys (Findlater and Kottler, 1998). Similarly, satisfied customers are more likely to participate in customer satisfaction Web surveys (Enander and Sajti, 1999), and more involved e-commerce users are more likely to participate in Web surveys on e-commerce (Elder, 1999). Moreover, the findings in this research also show the relationship between the duration of Internet use and the research participation on the Web as shown in table 5.

Table 5. Relationship between Duration of Internet Use and Research Participation on Web

Duration of Internet Use	Research Participation on Web	
	Yes %	No %
Not over than 2 years	13.0	23.9
3 – 4 years	48.0	47.1
5 years and over	39.0	29.0

p=0.028

The duration of Internet use is another important factor which affects the research participation on the Web. From table 5, the first column shows that 39.0% of the respondents who have used the Internet for 5 years or more tended to participate in the research on the Web, compared to 29% who did not tend to participate in the project. Furthermore, the intensity of computer and Internet usage is the most important predictor of cooperation in a Web survey. (Vasja Vehovar, 2002) Some previous research papers state that the usual characteristics of Internet usage also determine the participation in Web surveys: respondents are young, educated, rich, and male (Batagelj, 1998; Flemming and Sonner, 1999). However, in this study, there is no any significant relationship between the demographical data of respondents and research participation. Nevertheless, the user's experience in the previous research participation on the Web is the most important predictor of cooperation in Web survey at this time. If the users have ever experienced previous research participation on Web, they tended to participate in Web survey research of this study as shown in figure 5.

Figure 5. Relationship between User's Experience in Research Participation on Web and the Cooperation in Web survey at this time
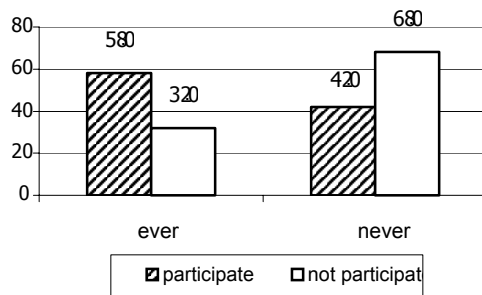
From figure 5, the research results show that 58.0% of the respondents who have experienced cooperation with research on the Web tended to participate in the research on the Web for this study. However, 68.0% of the respondents who have not experienced cooperation with research on the Web did not tend to participate in this Web survey.

8. CONCLUDING REMARKS AND DISCUSSION

In this research, the challenge of the world wide web: a comparative analysis model of web survey application with other modes of information collection in research participation, my target population is the university students at Assumption University in Bangkok, Thailand. I decided to select the Assumption University students because I have the complete population frame to which I can apply the probability sampling for selecting the representative sample from the population data base. I, therefore, do not indicate, either positively or negatively, if the research findings can be extrapolated to general internet users as a whole because I have not listed all general internet users and I need to apply the probability sampling in this project. However, I do indicate in this research project that the focus on the university students in the research participation of Web survey design is much more complex than for other modes of information collection. First, technology emerges as an additional factor that interacts with other features at the respondent, at the society, and at the survey organization levels. Second, technological changes are occurring extremely fast, permanently complicating a whole array of issues. Third, the very nature of the Web survey mode often introduces components of mix-mode surveys at the solicitation as well as at the responding stage. (Vasja Vehovar, 2002) The important findings in this research release are that non-response rate to Web surveys is to produce the high non-response rate of research participation on the Web which may affect the quality of survey data. However, Web survey is more likely to produce 'Yes' answers to the sensitive questions than other modes. Moreover, the findings of this study illustrate that most of the target users were not informed about the Web survey project. They have not received the invitation email due to a wrong given email address and they do not use the university's email account. The assumption of future research is that if they are well informed about the Web survey project, the response rate should increase. In the meantime, Web surveys provide many more options for the designer, far exceeding the relatively limited design features of traditional mail surveys. They may be used as powerful tools to guide the respondents through the survey, to motivate them to complete the task, and to provide a rich variety of visual stimuli to enhance the survey question on Web. (Mick Couper, 2001) Research findings in this paper show that the self-administered surveys, whether a paper-pencil mail survey or on the Web, can decrease the potential problems of socially

desirable answers in survey research. The research findings in this study also state the factors which produce the high non-response rates are due to ambiguous passwords and identification process, limitation of Internet access, low attractive incentive, and low attractive Web page design. It may be true, therefore, that a Web survey designer should consider these factors in future Web survey. For example, in the future research on the Web, Web survey designers should embed a unique identifier (ID) into the Uniform Resource Locator (URL) provided to each respondent. (Mick Couper, 2001) Web surveys now provide an enormous opportunity for methodological research and for understanding the role of the instrument in self-administered surveys. As the survey industry embraces the Web for data collection, many important questions remain to be addressed. One of the most important – and interesting – is whether Web surveys are a technological advance in the mail survey format or whether they are an entirely new format with multimedia capabilities that challenge the survey designer and present exciting new opportunities for questionnaire design. (Mick Couper, 2001) In this research paper, The Challenge of World Wide Web: A Comparative Analysis Model of Web Survey Application with other modes of information collection in research participation," I can conclude conditionally that Web survey is more likely to produce a higher quality of research data than other modes of information collection. However, the Web survey needs more future research to find strategies to get respondents to participate in the research on the Web, to motivate them to complete the Web questionnaire, and to understand more about the factors that affect the respondents' participation behavior. Moreover, I can say that the multimedia capabilities of the Web offer many opportunities and challenges in achieving the goals of maximizing research participation and data quality over than other modes of information collection in survey research.

9. BIBLIOGRAPHY

Batagelj, Z., Lozar, K., and Vehovar, V., 1998, Respondent's Satisfaction in WWW Surveys, paper presented at the International Conference on Methodology and Statistics, Preddvor, Slovenia.

Bradburn N.M., Sudman, S., and Associates, 1979, Improving Interviewing Methods and Questionnaire Design: Response Effects to Threatening Questions in Survey Research, San Francisco: Jossey-Bass.

Clausen, J.A., and Ford, R.N., 1947, Controlling bias in Mail questionnaires, Journal of the American Statistical Association, Vol. 42, 497 – 511.

Clayton, R., & Werking, G.S., 1998, Business surveys of the future: The World Wide Web as data collection methodology. In M.P.Couper,

R.P.Baker, J.Bethlehem, C.Z.F.Clark, J.,Martin, W.L.Nicholls,II,&J.M.O'Reilly (Eds.), Computer assisted survey information collection, New York: Wiley-Interscience

DeMaio, T.J., 1984, Social desirability and survey measurement: a review, In C.F.Turner and E.,Martin (eds.), Surveying subjective phenomena, Vol., 2, New York: Russell Sage.

Don A.Dillman, 1999, Coursepack for Self-Administered Mail & Web Survey Design, Michigan: University of Michigan.

Don A.Dillman, 2000, Mail and Internet Survey , New York: Wiley.

Dovidio,J.F., and Fazio, R.H., 1992, New Technologies for the direct and indirect assessment of attitudes. In J.M.Tanur(ed.), Question about Questions: Inquiries into the cognitive bases of surveys, New York: Russell Sage.

Edith D.de Leeuw and Johannes van der Zouwen, 1988, Data quality in telephone and face-to-face surveys: a comparative meta-analysis. In Robert M.Groves, Paul P.Biemer, and associates, Telephone survey methodology, New York: Wiley.

Elder, A., 1999, Power buyers or cautious shoppers? Segmenting the E-Commermce Market, Proceeding of the ESOMAR Worldwide Internet Conference Net Effects, Amsterdam: ESOMAR.

Enander, J. and Sajti, A., 1999, Online survey of online customers, value-added market research through data collection on the Internet, Proceeding of the ESOMAR Worldwide Internet Conference Net Effects, Amsterdam: ESOMAR.

Findlater, A. and Kottler, R.E., 1998, Web interviewing, Validating the application of Web interviewing using a comparative study on the telephone, The Worldwide Internet Seminar 1998 in Paris, France, Amsterdam: ESOMAR.

Herman, Mary B., 1977, Mixed-mode data collection: telephone and personal interviewing. Journal of Applied Psychology, 62, 399 – 404.

Hochstim, J.R., 1967, A critical comparison of three strategies of collecting data from households, Journal of the American Statistical Association, Vol.62, 976 – 989.

Kish L., 1965, Survey Sampling, New York: Wiley-Inter science.

Kojetin, B.A., Borgida, E., and Snyder, M., 1993, Survey topic involvement and nonresponse bias,

Proceedings of the survey research methods and cognition, Vol.2, 647 – 655.

Lars Lyberg, and Daniel Kasprzyk, 1991, Data collection methods and measurement error: an overview. In Paul P.Biemer, Robert M.Groves, Lars E.Lyberg, Nancy A.Mathiowetz, and Seymour Sudman, Measurement Errors in Surveys, New York: Wiley-Interscience.

Martin, J., 1996, Cybercorp: The New Business Revolution, New York: American Management Association.

Mick P.Couper, 2000, Web Surveys: A review of issues and approaches, American Association of Public Opinion Research, Vol.64: 464 – 494.

Mick P.Couper, 2001, Web Survey design and administered, American Association of Public Opinion Research, Vol. 65: 230 – 253.

Reginald P. Baker, 1998, Computer Assisted Survey Information Collection, The CASIC Future, New York: Wiley-Interscience.

Robert M.Groves, 1990, Theories and methods of telephone surveys, Annual Review of Sociology, Annual Review, Inc.

Scott D.Crawford, 2002, Coursepack for Web survey implementation, Michigan: University of Michigan.

Vasja Vehovar, Zenel Batagelj, Katja Lozar Manfreda, and Metka Zaletel, 2002, Nonreponse in Web Surveys. In Robert M.Groves, Don A.Dillman, John L. Eltinge, and Roderick J. A. Little, Survey Nonresponse, New York: Wiley-Interscience.

Wolf, F.M., 1986, Meta-analysis; Quantitative methods for research synthesis, Beverly Hills: Sage.