A comparison of response characteristics from web and telephone surveys

Catherine A. Roster, Robert D. Rogers and **Gerald Albaum**
University of New Mexico
Darin Klein
Microsoft Corporation

Increasingly, web surveys are being used to supplement telephone survey data and some predict internet methods will one day replace telephone interviews as the primary method for surveying general populations. Despite these trends, few studies have systematically compared response differences between the two methods. This article describes a study in which both telephone and web surveys were used to collect data on the corporate reputation of an international firm. Findings reveal significant differences in sample characteristics, response effects and overall costs. In addition to demographic differences, the web garnered a lower response rate, more item omissions, and produced more negative or neutral evaluations than did the telephone survey. Factor structure for the corporate reputation construct was simpler in the web-based data. Predictability of behavioural measures was essentially equivalent between the two modes; however, cost-per-contact was significantly lower in the web survey.

Introduction

A growing number of researchers regard the web as a speedy, cheap and effective alternative to traditional data collection methods. Not only can web surveys deliver large samples within a short period of time, but also they can do so without the costs of interviewers, training, postage, data entry, and a myriad of other associated expenses. Furthermore, industry experts argue that web-based surveys can offer higher quality data due to elimination of interviewer error and built-in checks that prohibit respondent errors (McCullough 1998; Dillman 2000).

Concerns about internet access, technology unevenness, coverage error and sample representativeness, limited the early use of web surveys to

finite populations, such as 'internet users,' or 'technology innovators'. Although technology barriers still pose some legitimate concerns for web-based research, the potential for wider deployment of web surveys is ballooning as the incidence of household computer ownership and internet accessibility continues to rise. According to a recent Nielsen//NetRatings report (2001), over 459 million people now have home-based internet access worldwide. Although households in the US and Canada still account for 40% of the world's on-line population, internet usage in European markets, Japan, and other developed/developing nations continues to explode at a double-digit pace.

Substantial data collection efficiencies, cost advantages, and wider dissemination of internet access among diverse groups are fuelling trends toward usage of web-based surveys in studies designed to be representative of more general populations. Although the future of web-based research appears promising, the movement toward using the web to generate inferential samples has some experts concerned. In a study of mail, telephone and web surveys used in political polling for the 1999 New Zealand general election, Hoek *et al.* (2002) concluded that web-based samples were seriously biased and that the biases noted could not be corrected by traditional weighting variables such as age and gender. In a discussion of sampling issues for web-based surveys, Bradley (1999) emphasises that 'a sample of internet users is only representative of internet users'. Couper (2000) stresses that demographic differences between web samples and the general population are only 'part of the story' and proposes that 'the key question is whether the two populations are similar on the substantive variables of interest' (p. 471). This issue, notes Couper, has received far less attention than it deserves from the research community as it generally requires more cumbersome and expensive designs, such as conducting surveys with different methods.

One example of the trend toward 'mainstreaming' web-based research is a movement within corporate and commercial research firms, and even among academics, to begin supplementing or replacing telephone with web-based surveys (Willems & Oosterveld 2003). For instance, internet surveys now account for 47% of Harris Interactive's polling business and the firm credits the transition from traditional telephone surveys to web-based surveys for their profit superiority in the industry (Einhart 2003). Each method is comparable for gathering data from large samples with sufficient statistical power in a short amount of time even though each relies on a different mode of data collection (person-administered versus computer-administered). This trend continues in spite of the fact that very

little, if any, research has been reported concerning differences between these two methods that may significantly impact results. While there has been research done on the issue by commercial research companies and other research organisations, findings have been treated as proprietary and, thus, have not been published and made available to the broad research community. Techniques are available to help overcome differences in sample characteristics, but differences in substantive variables such as attitudes, intentions or behaviours, if such differences exist, may render subsamples incompatible and therefore make it inappropriate to combine results gathered via these two methods. Furthermore, substantive differences may have far-reaching implications for trend analysis of key indices monitored on a routine basis should web surveys eventually replace telephone surveys in commercial and corporate research studies.

The study reported in this paper examines differences in telephone versus web-based survey samples and responses. The context of the study was a single survey of corporate reputation conducted for a multinational corporation in a south-western metropolitan area of the United States. The study used web-based and telephone-based data collection procedures to obtain data from a general population of local residents. Specifically, the study objectives were to (1) examine sample equivalence, (2) ascertain differences in data quality, including response rates and item omissions, (3) uncover any response effects and (4) determine cost differences between the two methods. Differences in any or all of the above factors could have a significant impact on the way decision-makers use their data, especially when the assumption is made that the two methods produce comparable data.

Background

The impetus of much of the published research comparing different survey methods has been to establish which of the methods studied produces superior results. As many of these studies were conducted before web-based data collection methods had come fully into their own, few include electronic survey data collection methods. Given the relatively recent advent of web-based research, all-method comparisons incorporating web surveys are rare and those that have included web-based research have concentrated on comparing web with other electronic methods, such as e-mail (e.g. Backman *et al.* 1996; Tse 1998), or to traditional mail surveys as both are self-administered (e.g. Mehta & Sivadas 1995; Stanton 1998;

Cobanoglu *et al.* 2001). Broader review-type studies have dealt with more general methodological issues surrounding web survey issues (Zhang 1999; Couper 2000). Furthermore, a wide variety of indices have been employed to evaluate 'superiority'. For instance, studies have compared *response rates* (Yu & Cooper 1983; Hox & de Leeuw 1994; Backman *et al.* 1996; Gjestland 1996), *data quality* (de Leeuw & van der Zouwen 1988; Mehta & Sivadas 1995; Schaefer & Dillman 1998; Tse 1998), *item omissions* (Stewart 1982; Durand *et al.* 1983; Omura 1983), and *response effects* (Bishop *et al.* 1988; de Leeuw *et al.* 1996). The ultimate definition of 'superiority' is made somewhat contentious by the fact that different parties value and assess these response characteristics differently, depending on the researcher's paradigm and the study's objectives.

Perhaps more importantly, although many studies have sought to compare methods in terms of sample equivalence or data superiority in one respect or another, few have addressed the issue of *comparability* of response content in data collected via different methods. This issue is worthy of closer scrutiny than it has received thus far, especially given the trend toward mixed-mode data collection as a means of improving response rates (Dillman & Tarnai 1988; Dillman 2000). The few published studies that have attempted to compare response effects and content have produced mixed results. For example, Srinivasan and Hanway (1999) found that telephone IVR responses to a series of customer satisfaction questions elicited higher ratings than those given in a mail survey. On the other hand, Willems and Oosterveld (2003) show that using a so-called hybrid approach (internet and telephone combined) in a crossover research design produced comparable results – factor structures for 10 scaled-items were equivalent. Similarly, Taylor (2000) finds comparable results on several items when comparing internet versus telephone surveys, although responses to some items in online surveys were 'substantially different' from telephone survey results and could not be corrected using traditional correctional weighting procedures.

Non-equivalence may be caused by *modality* or *population* effects (Willems & Oosterveld 2003). Modality effects are response differences at the individual level due to the mode of data collection, or the way respondents interact with the questionnaire. Population effects are response differences at the total sample level resulting from population biases at an aggregated level because some subpopulations are over or under-represented in the sample. With telephone-based surveys it is possible to obtain a representative national sample, whereas internet surveys are by definition restricted to respondents with internet access.

Although internet access is relatively high in the US, nearing 70% according to a recent Arbitron study (McKillen 2002), and in business-to-business (B2B) situations would even be higher, there is still substantial variation in technology access among key demographic groups. For instance, while on the rise, household computer penetration rates among US Hispanic consumers lag behind their Anglo counterparts (Lach 2000). Moreover, computer penetration and internet access are lower in most parts of Asia, Latin America, and Central and Eastern Europe (Parmar 2003) than in other parts of the world.

One issue that has plagued comparative methodological studies of this nature has been the lack of a consistent process for the analysis of alternative survey data collection methods. Albaum and Peterson (1985) have offered a comprehensive paradigm that can be used to guide comparative evaluations on survey methods. The process suggested by these authors separates the issues confronting researchers with data collected via different methods into three distinct stages. The first stage entails an evaluation of sub-sample characteristics in order to determine sample equivalence between or among methods and each to the general population from which findings are to be inferred. The second stage entails an evaluation of dependent variables to assess comparability in terms of response effects, including consideration of factors that may detract from data quality, such as item omissions or response rate. The third stage involves overall cost/benefit comparisons of survey methods including factors such as response speed, overall costs and predictive differences. This framework was used to guide our investigation of differences in telephone and web-based samples for a single study on corporate reputation of an international firm among the 'general population'.

Methodology

Two methodological issues guided the structure of this research project. One was the need to have the data collected by a common source. In order to overcome problems that might confound an assessment of survey results if multiple sources were to be used, both the web and the telephone surveys were conducted by a single commercial vendor. Second was the need to use a common measurement instrument. An existing proprietary questionnaire measuring corporate/company reputation was adapted for use by the vendor for both surveys. This instrument included 22 Likert-type scales (five-category response format) assessing level of agreement or disagreement with statements associated with the corporation's reputation

in the community, a 10-point overall measure of corporate reputation on a scale of 'very positive' to 'very negative' that was to be used to evaluate 10 organisations in the area, and seven 4-point statements exploring general community concerns on a rating scale of 'a very serious problem' to 'not a problem at all'.

The population of interest was adults in a major metropolitan area in the south-west. The same vendor was used to acquire both samples. The vendor purchased a sample of telephone numbers within the metropolitan area from a list broker and the ultimate sample was determined using a conventional probabilistic method in which the last digit of each number on the list was replaced with a random number (Tucker *et al.* 2002). Data were collected by the vendor's call centre. The web sample was comprised of randomly selected addresses purchased from a company that maintains a panel from the same metropolitan area. In effect, the study was a field experiment using an 'after-only' design employing standard procedures utilised by commercial marketing research firms.

Findings

Our first stage of evaluation concerned issues regarding subsample equivalence between general population characteristics of telephone and web samples. Table 1 shows distributions of six demographic characteristics for each of the sample groups. The samples differed from each other on all characteristics except household income, and both samples were different from the population based on the 2000 United States Census of Population. Overall, major differences between the two sample groups are summarised as follows:

- Age – a greater proportion of respondents aged 25–50 responded to the web survey than the telephone survey, whereas the opposite held for respondents aged 65 and older.
- Ethnicity – Anglo respondents were the majority group in both survey groups, whereas Hispanic respondents were relatively more prominent in the telephone survey.
- Education – a greater proportion of respondents to the telephone survey than to the web survey were college graduates or had postgraduate degrees or work.
- Time in Community – respondents to the web survey (54.7%) were relatively more likely to have lived in the community for no more than 20 years than respondents to the telephone survey (36.9%).

Table 1 Demographic characteristics of respondents (percentage distributions)

Characteristic	Sample group		*p*
	Telephone	Web	
Age	(*n* = 251)	(*n* = 307)	<0.001
18–21	9.2	5.9	
22–24	4.4	4.6	
25–34	13.5	21.2	
35–49	29.9	37.8	
50–64	26.7	24.4	
65 and older	16.3	6.2	
Gender	(*n* = 251)	(*n* = 284)	<0.001
Male	50.2	33.1	
Female	49.8	66.9	
Ethnicity	(*n* = 242)	(*n* = 258)	<0.001
Anglo	53.7	72.1	
Hispanic	38.4	17.8	
Black/African American	0.8	1.9	
Native American	0.8	1.9	
Other	6.2	6.2	
Household annual income	(*n* = 202)	(*n* = 241)	<0.50
Less than $10,000	5.4	6.6	
$10,000–$19,999	9.9	8.7	
$20,000–$29,999	11.4	15.8	
$30,000–$39,999	8.3	21.2	
$40,000–$49,999	16.3	10.8	
$50,000–$59,999	10.9	12.0	
$60,000 and greater	27.7	24.9	
Education	(*n* = 245)	(*n* = 271)	<0.001
Some high school	3.7	0.0	
High school graduate	18.8	13.3	
Some college/associate degree/vocation school	34.3	50.9	
College graduate (4 years)	25.3	20.3	
Postgraduate work or degree	18.0	15.5	
Time in community	(*n* = 249)	(*n* = 271)	<0.002
Less than 6 six years	15.3	21.4	
6 to 10 years	8.4	13.7	
11 to 20 years	13.3	19.6	
More than 20 years/native	63.1	45.4	

- Gender – although there were significant differences in respondents' gender, this does not reflect population differences as the sample for the telephone survey aimed at a quota of 50% female. Still, two-thirds of the web survey respondents were female.

Table 2 Obtained sample size and response rates

	Survey method	
	Telephone	Web
Number of original attempts	2173	974
Number contacted	620	974
Number responded	251	318
Number of unusable responses	0	46
Number of usable responses	251	272
Response rate (number of usable/number contacted) (%)	40.5	32.6
Response rate (number of usable/number of attempts) (%)	11.5	32.6

When significant differences among subsamples exist, two options are available for researchers (Albaum & Peterson 1985). First, responses can be adjusted on the basis of one or more key sample or population characteristics and then the subsamples combined. Adjustment can be by weighting or by using covariance analysis to remove the influence of the sample characteristic(s) from subsequent analyses. The second option is that of accepting subsample differences and treating the data accordingly. If this is done, a relatively large sample size is needed from both groups.

This brings us to the second stage of evaluation, which examines response effects and content. Turning to measures of effect, the first concern was whether response rates were equivalent for the two methods. Table 2 shows results from this assessment. For the telephone survey, 620 households were contacted resulting in a sample of 251 usable responses, a response rate of 40.5%. This rate seems very high for a telephone survey

Table 3 Item omissions, by survey mode

	Mean and (std. dev.)		Range
Variable	*p(t)*	Item omission %	Item omission %
Demographic (6 questions)	(0.04)		
Telephone		4.4 (7.6)	0–19.5
Web		14.5 (7.1)	3.5–24.2
Likert scales (22 statements)	(0.01)		
Telephone		25.2 (10.8)	11.2–56.2
Web		32.6 (6.5)	22.6–43.7
10-point scales (10 scales)	(0.48)		
Telephone		12.6 (13.2)	1.6–37.1
Web		16.1 (8.0)	8.5–29.6
4-point scales (7 scales)	(0.00)		
Telephone		2.2 (2.3)	0–6.0
Web		10.2 (3.5)	6.6–17.3

and is misleading. When the total number of attempts (i.e. calls made) is used in determining response rate, the rate drops to 11.5%, which is a more widely reported rate for telephone surveys. Similarly, the response rate for the web survey may be overstated from what one would obtain if a sample were drawn from a frame that is not a panel. For the web survey, 974 invitations were sent out to members of a panel resulting in 318 respondents who went to the website. Of those, 272 responses were usable for a response rate of 27.9%.

Looking at item omissions, the web survey produced more non-responses (i.e. a greater average) to the demographic questions and all scaled questions, although the difference for the 10-point scales was not significant. As shown in Table 3, the largest difference in average percent

Table 4 Mean values of extent of agreement with positively-worded statements about the target corporation*

	Survey mode			
Scale Item**	Telephone[a]	Web[b]	*t*	*p*
Education support	4.14	3.98	1.66	< 0.10
Technology access	3.69	3.42	2.27	< 0.03
Technology training	3.51	3.49	0.16	< 0.88
Environmentally responsible	3.46	3.29	1.40	< 0.17
Local environmental support	3.08	2.99	0.56	< 0.58
Environmental programme support	3.45	3.56	0.90	< 0.38
Water conservation	3.45	3.31	1.01	< 0.32
Community health concern	3.44	3.18	2.05	< 0.05
Local charitable organisation support	4.12	3.93	2.09	< 0.04
Employee voluntarism for community support	3.94	3.79	1.49	< 0.14
Minority advancement	3.79	3.42	3.11	< 0.003
Advancement of women	3.81	3.38	3.43	< 0.002
Neighbourliness	3.74	3.46	2.54	< 0.020
Community commitment	3.91	3.72	1.69	< 0.100
Safe employee workplace	4.26	3.90	3.96	< 0.001
Employee job security	3.83	3.48	3.16	< 0.003
Has a future in the industry	4.31	4.03	3.41	< 0.002
Marketplace value	4.16	3.99	1.89	< 0.060
Technology leadership	4.24	4.15	1.01	< 0.320
Corporate citizenship	3.99	3.67	2.98	< 0.004
Local community health concern	3.60	3.37	2.15	< 0.04
Community health risk	3.32	3.12	1.66	< 0.10

*Scaled as 5-category Likert scales where 1 = 'strongly disagree' and 5 = 'strongly agree'.

** Scale items were positive statements about the target company's impact on these items.

[a]Number of responses varied from 110 to 223.

[b]Number of responses varied from 179 to 246.

of omission between the two survey modes of data collection occurred for the demographic questions and 4-point rating scales. The range of item omission percentages was greater for the telephone survey mode with the exception of the demographic questions and the 4-point rating scales. This might reflect the presence of interviewers causing differential responses to verbal and non-verbal cues.

The next set of concerns addressed differences in response content and how data substance might vary due to survey data collection method. As shown in Table 4, there were significant differences at $p < 0.05$ or less in the mean agreement scores for 11 of the 22 Likert statements on corporate/company reputation. This number exceeds the binomial chance probability. For all of the statistically significant scales (and for 21 of the 22 scales as well), the mean value was greatest for the telephone survey mode. Since all the statements were worded positively, the lower mean scores indicate that the web survey generated more negative and neutral evaluations than did the telephone survey.

For the 4-point rating scale measuring level of concern about various community issues, web respondents indicated more extreme concerns than did respondents to the telephone survey. Table 5 shows that five of the scales were significant at $p < 0.05$. Web respondents reported greater mean values for all scales.

A positive result emerges from the evaluation of the reputation of the target organisation and eight (of nine) other organisations in the local community. Although only two of the 10 organisations differed significantly ($p < 0.05$) based on survey mode, the telephone survey mode generated greater mean values for seven of the 10 items, which can be

Table 5 Mean values of attitudes concerning community problems*

	Survey mode			
Community concern	Telephone[a]	Web[b]	*t*	*p*
Education support	4.14	3.98	1.66	< 0.100
General health of local residents	2.65	2.86	2.56	< 0.020
Traffic problems	2.81	3.05	3.05	< 0.003
Air pollution	2.67	2.80	1.74	< 0.090
The quality of education	3.35	3.65	4.04	< 0.001
The condition of the local economy	3.10	3.42	4.55	< 0.001
The current water supply	3.65	3.55	1.52	< 0.130
The supply of water available 5 to 10 years from now	3.69	3.70	0.30	< 0.770

*Scaled from 1 to 4, where 1= 'not a problem at all' and 4 = 'a very serious problem'.
[a]Number of responses varied from 236 to 251.
[b]Number of responses varied from 263 to 297.

Table 6 Mean values of attitudes towards organisations in the local area*

Organisation	Survey mode		*t*	*p*
	Telephone[a]	Web[b]		
Utility company	6.12	5.83	1.33	< 0.19
Telephone company	4.88	4.35	2.46	< 0.02
Local university	7.59	7.20	2.14	< 0.04
Local food manufacturer	6.95	6.89	0.28	< 0.78
Chain retailer	6.96	6.92	0.20	< 0.85
'Target' organisation	7.18	7.35	0.79	< 0.44
Components manufacturer	6.14	6.25	0.48	< 0.63
Bank X	5.99	5.65	1.43	< 0.16
National laboratory	8.00	7.70	1.66	< 0.10
Bank Y	5.80	5.95	0.69	< 0.50

*Scaled from 1 to 10, where 1 = 'very negative' and 10 = 'very positive'.
[a]Number of responses varied from 173 to 247.
[b]Number of responses varied from 224 to 291.

interpreted as a directionally more positive evaluation than that from the web survey (see Table 6). These are relative interpretations as the absolute values were neutral or positive.

So far the analysis has focused on differences in magnitude, direction, and non-response at the level of individual items. But do these differences in modes of survey data collection also have an impact on techniques used to analyse the data, such as multivariate analysis? To examine this question the Likert-scaled data on company reputation for the target company were factor analysed using an eigenvalue of 1 as the criterion for extraction of factors. Based on Varimax rotation, the number of factors, the percentage of total variance these factors explain, and the reliability of the analysis appears in Table 7 for each survey mode. It appears that web survey respondents had a much simpler underlying structure of reputation in mind than did the telephone survey respondents. Although more of the

Table 7 Factor analysis characteristics

Analysis characteristic	Survey mode	
	Telephone	Web
Number of factors	6	2
Percentage variance explained	78.2*	68.5**
Reliability (θ)	0.84	0.92

*With 2 factors, percentage of variance explained is 50.8.
**With 6 factors, percentage of variance explained is 82.3.

total variance was explained by telephone responses, it took four additional factors to generate only 10 percentage points more explained variance.

Reliability was assessed by the coefficient theta (θ), which is based on the number of items factor analysed and the first (i.e., the largest) eigenvalue. In effect, theta is a special case of coefficient alpha (Carmines & Zeller 1979, pp. 60–61). Responses given by web participants had a greater reliability, but the reliability that emerged from both survey modes is acceptable. In sum, reputation data obtained by telephone respondents generated a more complex structure in the minds of respondents, and were slightly less reliable than data generated by the web. This may very well reflect influence of the interviewer and the interaction between interviewer and respondent.

The third stage of evaluation involved an overall cost/benefit analysis of the different survey modes. A key question here was whether superiority in predictive validity differed between telephone and web survey responses. For this analysis, factor scores were calculated and regression analyses were conducted using the appropriate factor scores (6 for the telephone mode; 2 for the web mode) as independent variables and the separate reputation evaluation (assessed by a 10-point scale of negative/positive) of the target organisation as the dependent variable. Both regression models were significant at $p < 0.001$, and the percentage of variance explained, R^2, did not differ much between the two modes: $R^2_{\text{tel}} = 0.48$; $R^2_{\text{web}} = 0.50$. Similar significant results were obtained when the same independent variables were regressed on a question that asked for extent of support toward the target company expanding its operations in the local area: $R^2_{\text{tel}} = 0.46$; $R^2_{\text{web}} = 0.51$.

Last, a benefit of using a common vendor was the availability of cost data on both data collection methods. The cost per interview for the telephone survey was $30 and for the web $14. While both methods are capable of generating large samples in a relatively short amount of time, there are distinct cost advantages to using web-based surveys as opposed to telephone interviews.

Summary and conclusions

The study sought to determine if there were differences in data collected by a telephone survey and by a web survey. Although differences are to be expected as each relies upon different modes of administration, the extent and nature of these differences is of great practical concern as more and

more researchers supplement or replace traditional telephone surveys with web-based surveys. Our data were derived from a real world application that relied upon conventional industry standards to obtain data from multiple survey methods in a single study of a general population's attitudes. Our findings suggest there can be important differences in terms of sample demographics, response effects, substantive content and overall cost/benefits.

The rapid dissemination of internet technology ushers in the temptation to assume that internet samples are a viable alternative to telephone samples as a suitable frame for representing the general population. While distinctions may be dissipating, our findings support the argument that internet samples continue to over-represent some groups, like younger consumers, and under-represent important ethnic groups, such as Hispanics. Such differences, however, are likely to be less prevalent in the future and in the meantime can be dealt with by employing weighting schemes or sampling adjustments.

Perhaps a greater concern lies in the area of response effects and substantive differences in response content. It is here that differences in mode of administration are most likely to impact respondents' behaviours and exert a bottom-line difference in survey results, some of which cannot be as easily rectified as differences in respondent characteristics. Here, response rate as a percentage of total attempts was higher in the web survey, but it should be noted that our web frame was drawn from a panel while the telephone frame was not. The removal of interviewer apprehension also appears to have increased item omissions to demographic questions and led to more neutral or negative attitudinal evaluations in the web survey as opposed to the telephone survey. Respondents to the web survey also seemed to adopt a more streamlined cognitive response style to measures of corporate reputation. This is an interesting finding that warrants further research, especially in light of the fact that this simpler factor structure proved to be just as predictive of overall behaviours and attitudes as the more complex structure elicited by responses to the telephone survey.

Nevertheless, the issues most likely to continue to propel the transition from telephone to internet surveys reside in the overall cost/benefit differences between the two methods. The primary advantage for switching is lower costs for data obtainable at equitable speed. Commercial research firms know this already. Harris Interactive estimates that internet surveys are 15% to 20% cheaper to administer (Einhart 2003) and suggests results may be more accurate than those from

telephone surveys. Our findings substantiate the claim that there is a substantial cost advantage to collecting data using web-based surveys. In this particular study, cost for the web survey was 53% lower than for the telephone survey. Our findings also lend support to the notion that web surveys may be equally, if not more, accurate than telephone surveys in predicting behaviours.

Researchers will no doubt continue to gravitate toward web-based surveys because they are fast, cheap, and can produce large samples. However, caution should be exercised before assuming that results obtained from web-based surveys produce data equivalent to telephone surveys. The dawn of a new era is visible, but perhaps not yet upon us.

References

Albaum, G. & Peterson, R.A. (1985) A paradigm for methodological research on mail survey response. *Proceedings of the 1985 AMA Summer Educator's Conference*.

Backman, D., Elfrink, J. & Vazzana, G. (1996) Tracking the progress of email vs. snail mail. *Marketing Research*, **8**(2), pp. 31–35.

Bishop, G. *et al.* (1988) A comparison of response effects in self-administered and telephone surveys. In R.M. Groves *et al.* (eds) *Telephone Survey Methodology*. New York: Wiley, pp. 321–340.

Bradley, N. (1999) Sampling for internet surveys: an examination of respondent selection for internet research. *Journal of the Market Research Society*, **41**(4), pp. 387–394.

Carmines, E.G. & Zeller, R.A. (1979) *Reliability and Validity Assessment*. Beverly Hills, CA: Sage Publications.

Cobanoglu, C., Warde, B. & Moreo, P.J. (2001) A comparison of mail, fax, and web-based survey methods. *International Journal of Market Research*, **43**(4), pp. 441–452.

Couper, M.P. (2000) Web surveys: a review of issues and approaches. *Public Opinion Quarterly*, **64**(4), pp. 464–494.

de Leeuw, E.D., Mellenbergh, G.J. & Hox, J.J. (1996) The influence of data collection method on structural models: a comparison of a mail, a telephone, and a face-to-face survey. *Sociological Methods & Research*, **24**, pp. 443–472.

de Leeuw, E.D & van der Zouwen, J. (1988) Data quality in telephone and face-to-face surveys: a comparative analysis. In R.M. Groves *et al.* (eds) *Telephone Survey Methodology*. New York: Wiley, pp. 283–299.

Dillman, D.A. (2000) *Mail and Internet Surveys: The Tailored Design Method*. New York: John Wiley & Sons.

Dillman, D.A. & Tarnai, J. (1988) Administrative issues in mixed mode surveys. In R.M. Groves *et al.* (eds) *Telephone Survey Methodology*. New York: Wiley & Sons, pp. 509–528.

Durand, R.M., Guffey Jr., H.J. & Planchon, J.M. (1983) An examination of the random versus nonrandom nature of item omissions. *Journal of Marketing Research*, **20**, pp. 305–313.
Einhart, N. (2003) The opinion catcher. *Business 2.0*, **4**(4), p. 87.
Gjestland, L. (1996) Net? Not yet. *Marketing Research*, **8**(1), pp. 26–29.
Hoek, J., Gendall, P. & Healey, B. (2002) Web-based polling: an evaluation of survey modes. *Australasian Journal of Market Research*, **10**(2), pp. 23–34.
Hox, J.J. & de Leeuw, E.D. (1994) A comparison of nonresponse in mail, telephone, and face-to face surveys. *Quality and Quantity*, **28**, pp. 329–344.
Lach, J. (2000) Crossing the digital divide. *American Demographics*, **22**(6), pp. 9–11.
McCullough, D. (1998) Web-based market research ushers in new age. *Marketing News*, **32**(19), pp. 27–28.
McKillen, D. (2002) Web growth levels off. *Medical Marketing and Media*, **37**(10), pp. 10–12.
Mehta, R. & Sivadas, E. (1995) Comparing response rates and response content in mail vs. electronic mail surveys. *Journal of the Market Research Society*, **37**(4), pp. 429–439.
Nielsen//NetRatings reports that 459 million people have internet access worldwide (2001, September) *Cableoptics Newsletter*, **12**(9), p. 11.
Omura, G.S. (1983) Correlates of item nonresponse. *Journal of the Market Research Society*, **25**, pp. 321–330.
Parmar, A. (2003) Net research is not quite global. *Marketing News*, 3 March, pp. 51–52.
Schaefer, D. & Dillman, D.A. (1998) Development of a standard e-mail methodology: results of an experiment. *Public Opinion Quarterly*, **62**, pp. 378–397.
Srinivasan, R. & Hanway, S. (1999) A new kind of survey mode difference: experimental results from a test of inbound voice recognition and mail surveys. Paper presented at the meeting of the American Association for Public Opinion Research, St. Pete Beach, FL.
Stanton, J.M. (1998) An empirical assessment of data collection using the internet. *Personnel Psychology*, **51**(3), pp. 709–725.
Stewart, D.W. (1982) Filling the gap: a review of the missing data problem. In B.J. Walker *et al.* (eds) *An Assessment of Marketing Thought and Practice*. Chicago: American Marketing Association, pp. 395–399.
Taylor, H. (2000) Does internet research work? Comparing online survey results with telephone survey. *International Journal of Market Research,* **42**(1), pp. 51–63.
Tucker, C., Lepkowski, J.M. & Piekarski, L. (2002) The current efficiency of list-assisted telephone sampling designs. *Public Opinion Quarterly,* **66**(3), pp. 321–338.
Tse, A.C.B. (1998) Comparing the response rate, response speed, and response quality of two methods of sending questionnaires: email vs. mail. *Journal of the Market Research Society*, **40**(4), pp. 353–361.
Willems, P. & Oosterveld, P. (2003) The best of both worlds. *Marketing Research,* Spring, pp. 23–26.
Yu, J. & Cooper, H. (1983) A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research*, **20**, pp. 36–44.
Zhang, Y. (1999) Using the internet for survey research: a case study. *Journal of the American Society for Information Science*, **51**(1), pp. 57–68.