

Paper-and-Pencil and Web Surveys: Different Methods, Same Results?

Kristofer J. Fenlason & Carrie Christianson DeMay
Data Recognition Corporation

This paper reviews how two commonly used, but different, survey administration methods may affect survey results. Four types of results are examined: Quantitative data, non-response/missing data, comment length, and measurement equivalence. Although more research remains to be conducted, indications from several studies are positive, suggesting that practitioners and researchers may well combine and compare paper-and-pencil and web-based survey results. Practice implications and suggestions for survey practitioners and researchers are offered.

Organizations conducting surveys today have a wide array of administration methods from which to choose. More “traditional” methods such as paper-and-pencil and telephone interviewing have been augmented in recent years by fax and automated telephone response systems (Kuhnert and McCauley, 1996) and most recently by Internet-based/web-based solutions.

Several forces, both general and specific, appear to be driving organizations to use web-based surveying. Kraut and Saari (1999) noted that as organizations connected their employees via e-mail, the Internet, and corporate intranets, electronic surveying increased. A recent report by the U.S. Department of Commerce (2002) indicates that both access to and use of the Internet has been steadily climbing for the last few years.

Based on the authors’ consulting experiences, an increase has been seen in corporations that want to move to web-based data collection for the purpose of increasing alignment with strategic initiatives to become more “e-business based.” Such initiatives are also being facilitated by an employee-base that is increasingly more comfortable with, and dependent on, the Internet and e-mail. In addition to these broader reasons driving organizations to web-based solutions, several authors have noted that electronic surveying brings with it several specific advantages over paper-and-pencil surveys including: increased speed (e.g., Bachmann, Elfrink, & Vazzana, 1996; Oppermann, 1995; Schaefer & Dillman, 1998), the potential for lower cost (e.g. Kraut and Saari, 1999; Sproull, 1986; Weible & Wallace, 1998), and a greater degree of control over questionnaire flow (e.g., Dillman, 2000).

Yet, despite these driving forces, moving an organization to a “paperless” option introduces a number of challenges and such a conversion is not likely to occur overnight. For instance, Kraut and Saari (1999) noted a number of potential concerns with electronic surveying including confidentiality,

access/coverage problems, data accuracy, technological problems, and initial investment costs. (Although it should be noted that a number of these concerns have been or are being addressed successfully in particular programs e.g., Christianson DeMay & Armstrong, 2000).

Nevertheless, in some cases one or more of these issues may be enough to eliminate the option of surveying using exclusively electronic methods. For instance, while office workers in a company may have unfettered access to PCs, it is unlikely that assembly workers have a similar opportunity. Even if access exists for certain groups (e.g., via shared PCs or kiosks), training or privacy concerns may still make it a practical impossibility to execute an exclusively web-based survey.

One solution to this dilemma is to use a dual-method administration process that provides respondents with a choice of either paper-and-pencil or web survey. There is good reason to focus on these two methods. As Macey (1996) commented, paper-and-pencil surveys have been the dominant method of organizational surveying and this is likely to continue for some time. Because paper-and-pencil surveys have been such a well-established and widespread data collection method for decades, it is safe to say that most workers have at least a moderate level of comfort in using this method. At the same time, electronic surveying shows great promise in terms of widespread use and acceptance. For instance, Christianson DeMay and Toquam Hatten (2001) reported that 97% of respondents chose to respond using a web-based versus paper rating form in a multi-source feedback program.

For some organizations, a dual-method approach may be a temporary bridge until access issues are resolved. For others, it may be a relatively permanent state of affairs. In either case, it is important to examine the possible implications of collecting data using these different modes, as researchers and practitioners go about comparing, combining, and interpreting data from these different administration methods.

What “Results” Are We Comparing?

Interest in, and research on, the possible effects of administration method has existed for more than 15 years (e.g., Kiesler & Sproull, 1986; Schuman & Presser, 1981). With the advent of the Internet, this interest has been renewed and research is being updated (e.g., Church, 2001; Fenlason 2000; Stanton, 1998; and Yost & Homer, 1998). A comparison across several recent studies finds four broad types of results being commonly compared across these two methods.

Items and Scale Scores When thinking about survey results, the first image that probably springs to mind is that of the quantitative (data) responses to individual questions (items) or combinations of items (scales or factors). These may be in the form of any metric but are most often in the form of frequencies of occurrence, percentages of response, or mean scores.

Comments A second type of result is qualitative in nature. Opportunities to offer comments or respond to “open-ended” questions are fairly common in surveys. Although comments are qualitative responses, they can still be compared in terms of the quantity of what is written (i.e., length by word or character count) as well as content themes and valences (positive or negative tone of the comment).

Completion and Response Rates A third type of result that may vary by method, is the extent to which surveys are completed and returned for inclusion in the data set. This can involve the extent to which the individual items within a given survey are answered (completion) and the percentage of surveys returned (overall response rate).

Measurement Equivalence A fourth type of result to be compared is the extent to which the same questions, delivered via different administration methods, are measuring the same things (i.e., measurement equivalence).

What We Know So Far...

The following four points, based on the existing literature, present a brief review of the “state of the art” of our knowledge about the comparability of results from paper-and-pencil and Internet/web methods.

Differences in Item or Scale Scores Several studies have investigated differences between item or scale-level responses that may be attributed to administration method. The obvious concern here is that if one method produces more or less favorable responses that appropriate interpretation and use of the results may be difficult at best and misleading at worst. A recent

example is a study by Kraut, Oltrogge and Block (1998) who found telephone responses to be significantly more favorable than paper-and-pencil responses. Fortunately, there is a growing body of studies that address the similarities and differences of on-line/web-based and paper-and-pencil methods.

Before the current focus on Internet/web surveying, there were a few studies published on electronic, but not strictly web-based surveying. For instance, Rosenfeld, Giacalone, Knouse, Doherty, Vicino, Kantor, and Greaves (1991) found higher job satisfaction scores under computer assessment conditions; however, this was true only for subjects who were low self monitors. Lautenschlager and Flaherty (1990) found greater social desirability with a computerized data collection method. However, Booth-Kewley, Edwards, and Rosenfeld’s (1992) replication of this study did not reveal such administration effects. In sum, these early studies do not indicate a compelling, broad-based pattern of differences on items or scales.

More recent studies have focused specifically on a direct comparison of web and paper-and-pencil administration methods. Stanton (1998) found that responses to paper-and-pencil surveys showed higher levels of perceived distributive and procedural fairness than web responses, but found no significant mean differences on a consistency scale. In their comparison of web surveys to paper-and-pencil surveys, Yost and Homer (1998) found web survey mean responses to be higher (more favorable) than their paper-and-pencil counterparts. However, these differences disappeared when they controlled for job type. Similarly, Magnan, Lundby, and Fenlason (2000) found higher (more positive) scores for web versus paper-and-pencil but these differences dropped out when demographic differences were taken into account. In this case, production employees completed only the paper-and-pencil version, while headquarters employees had access to the web or paper version. When production workers were filtered out of the comparison, only one scale differed significantly and the direction of the difference switched from web to paper-and-pencil being higher. Fenlason’s (2000) study of web versus paper-and-pencil within a multi-source feedback application found, a “...pervasive pattern of non-significance...” of differences that indicated no compelling mean differences between web and paper-and-pencil administration methods on an item level.

Finally, Church’s (2001) comparison of two instances of on-line and paper-and-pencil surveys found only minor differences in mean scores and these differences were inconsistent. Perhaps more significantly, he found

that survey method only accounted for a fairly trivial amount of unique variance (2.5%). His conclusion was that some method effects did appear to exist but that they were inconsistent and fairly limited in their effect.

Based on the studies available so far, there does not appear to be a consistent, compelling effect of method on quantitative results. Nevertheless, the results of three of these studies point out an important caveat for the interpretation of results. That is, practitioners and researchers need to be cautious about interpreting differences without first determining if significant demographic differences exist that could have influenced the responses.

Completion and Response Rates Another potential concern about different administration methods is the effect of method on issues of survey completion (the extent to which all or only some of the individual items within a given survey are answered) and response rate (the percentage of number of surveys returned).

Both incomplete surveys and lower survey response rates can affect the quality and validity of the data obtained. For instance, in the case of multi-source feedback surveys, any significant lowering of item completion rates could have important implications given that these instruments typically have fewer respondents than an employee opinion survey. Missing one or two responses could cause the data set to fall below a minimum N threshold for reporting, resulting in an incomplete report, or no report at all being generated. Another example would be if the members of a particular sub-group within a larger organizational unit were less inclined to complete all the items or return the survey as a whole. In this case, the results could be less representative of the sub-group and could also cloud interpretation of the responses of the larger organizational unit.

Several of the previously reviewed studies assessed some aspect of item completion or response rates. Yost and Homer (1998) examined response time/latency and response rates and found significantly shorter response times in the web survey conditions; however, they found no significant differences in final response rates. Stanton's (1998) comparison of web and paper-and-pencil methods included an analysis of the "quality" of the data, which examined, in part, the proportion of missing data (i.e., items that were not completed). Stanton found that web-based data tended to have a lower percentage of missing values.

Fenlason (2000) found that the percentage of missing values was not significantly different across survey

items (comparing web and paper-and-pencil) and the differences that did exist were small (the largest difference was less than 1%). The direction of these differences indicated that paper-and-pencil had slightly fewer missing values than Internet (paper-and-pencil mean missing values = 1.35%, web mean missing values = 1.58%). However, in this study, respondents had the option of skipping a question or choosing a "Not Applicable" option, both of which were coded as "missing" when captured (making it impossible to disentangle actual omitted items from NA responses).

Church (2001) examined item completion rates in two studies of web vs. paper-and-pencil surveys. The first showed no significant difference in item completion rates. Yet, when skipped items and Don't Know (DK) responses were examined separately, it was found that the on-line method yielded seven times as many skipped items as paper. The opposite pattern was found when the use of a DK option is examined, although the differences in the rates of using DK were not nearly as pronounced as with the skipped responses.

The second study showed similar results for the skipped items (which were about five times more numerous in the on-line versus the paper-based survey). In contrast to the first study, there were no significant differences in the number of DK responses across the two methods, nor significant differences in item completion rates.

Although not addressing overall response rates, per se, Church's (2001) paper notes what appears to be a preference for choosing the on-line option. Across two studies, the preference was for on-line (58% chose on-line in Study One and 55% in Study Two, with the balance choosing paper-and-pencil). Other notable demographic differences in response include that those who were younger tended to choose the on-line method more than the paper-and-pencil (in Study One -- about 64% of those 36 years or younger and about 62% of those 36 - 50 chose on-line). The choice was more mixed for older employees. Of those 51 or older, about 53% preferred paper-and-pencil versus on-line of about 47%). In the second study age was even more clearly related to choice of on-line, with about 70% of those 35 or younger choosing this method, followed by about 57% of those 35 - 50 and 42% 51 years or older. Church found no differences for gender in either study and found a difference on tenure in only one study (with longer tenured employees being less likely to use on-line).

It appears that the question of whether there are overall differential response rates between web and paper-and-pencil has not yet been fully answered for at least two reasons:

First, early studies citing differences did not specifically address web-based surveys, but rather surveys delivered by other electronic means. While studies have found lower response rates with on-line surveys than with paper-and-pencil surveys, readers of the literature would do well to carefully differentiate between more recent web-based and other electronic formats, typically e-mail. For instance, both Schaefer & Dillman (1998) and Mehta & Sivadas (1995) found lower response rates for electronic surveys than paper-and-pencil but these studies used an e-mail format.

Second, in several of the more recent studies, it is not possible to ascertain meaningful overall response rates for the two different methods. For instance, Spera and Moye (2001) found an overall response rate of 77% (within a multinational company) and report that about 50% completed the survey on the web. However, access to the survey via the company intranet was non-random and in some areas local management was allowed to choose the method of survey. Likewise, an overall response rate was unavailable for Stanton's (1998) study because surveys were returned from a web page without identifying information. Similarly, Young et. al. (2000) could not compute a meaningful web vs. paper response rate because while participants were offered a choice (of paper or web), very few actually had access to the company intranet and response rate within the group with access was apparently not tracked.

Open-Ended Comments Comments are commonly used across a wide variety of surveys, including customer, employee/organizational and multi-source feedback instruments. To date, the published literature on Internet and electronic surveying has focused on the length of the comments themselves. Kiesler and Sproull (1986), for example, found that comments from electronic surveys were twice as long as from paper-and-pencil versions. Similarly, in Mehta and Sivadas (1995) e-mail survey, respondents wrote more than those who completed a paper-and-pencil version. Similarly, Yost and Homer (1998) found longer comments among respondents who completed a web versus a paper-and-pencil survey. However, there was no difference between methods in terms of the percentage of respondents who provided comments.

Fenlason (2000) found that comments tended to be longer in web responses compared to the paper-and-

pencil version. This web/paper multi-source survey provided the opportunity for comments about the manager's strengths and opportunities for improvement. The same pattern of responses was present across both the strengths and opportunities for improvement comments, with paper-and-pencil comments tending to be shorter (average length of 24.66 words for strengths and 23.49 for opportunities) compared to web comments (an average of 35.41 and 33.97 words for the strength and opportunities comments, respectively).

Although based on a handful of studies, the existing evidence indicates that, compared to paper-and-pencil surveys, open-ended comments on web-based surveys tend to be longer.

There could be several reasons for longer comments, but one explanation is that many web-based surveys effectively place no limit on how much respondents can write, while paper-based surveys have a finite amount of available space. Although, it can be argued that some respondents to paper surveys effectively "stretch" this limit by either writing smaller or attaching additional sheets. However, it is likely that providing longer comments in web surveys is a lower effort task than that of adding to paper-based surveys (e.g., more effort is required to hand write or type and separately attach additional comments).

It is also possible that web-survey respondents submit a greater volume of comments because of the relatively lower level of feedback about the volume they are writing. One of the current conventions for comment entry on web surveys is to provide a text box that allows only limited viewing of the comments as they are entered. For example, the comment entry box may only show a few lines at a time, and writers must scroll within this box to review and edit their comments. Thus, much less feedback on the volume being written is displayed than would be typically available when using a word processing program.

It is notable that none of the studies to date have tackled the issue of comparing comments beyond the metric of length. So content and valence of comments remain unexplored territories for comparison across methods.

Measurement Equivalence An area of key concern to researchers and practitioners alike is measurement equivalence. That is, whether responses collected via one method (e.g., the web) are psychometrically equivalent to responses collected via another method (e.g., paper-and-pencil). This equivalence issue is roughly analogous to students from two different sections of a class who take the same test, but under different conditions (e.g.,

one section takes a written test and the other section has the questions read to them). In order to compare and combine scores from both tests, the assumption is the tests measure the same things and are not affected by some other factor outside of the test content. Even if the questions are the same, it is possible for other conditions to affect how students might answer the questions (e.g., some students' hearing may not be as good as others, affecting how well those in the non-written condition comprehended the questions). So, it is possible that differences in scores can be due to factors other than a student's knowledge (i.e., may be due to the method in which the test was delivered).

Several researchers have investigated measurement equivalence in surveys across a variety of contexts, including: examining cross-national/cross-cultural equivalence in organizational surveys (e.g., Ryan, Chan, Ployhart, & Slade, 1999); 360 feedback to examine the equivalence of performance constructs across rating sources (e.g., Woehr, Sheehan & Bennett, 1999); and probing the equivalence of web and paper-and-pencil methods (e.g., Stanton, 1998).

The purpose of this paper is not to provide a detailed treatment of how to test for measurement equivalence. The interested reader is referred to several articles that provide a technically rich treatment of the methods and analyses involved. See Reise, Widaman, & Pugh (1993) for a discussion of determining equivalence using either IRT or SEM. Joreskog and Sorbom (1993) provide a detailed treatment of establishing equivalence using multiple groups confirmatory factor analysis (MG-CFA) within the LISREL 8 statistical package. Young, et al., (2000) provide an applied example of IRT with survey data while Ryan, et al., (1999) and Stanton (1998) provide applied examples of SEM with survey data.

Researchers and practitioners alike should be aware of the issue of measurement equivalence in the use of dual-method surveys. When equivalence can be established, then data comparisons (e.g., means, standard deviations, completion rates) can be interpreted with the confidence that differences are likely not due to the different methods measuring different things. It should be noted that not all the studies reviewed in this paper that compare and contrast data from different administration methods have directly examined the issue of measurement equivalence (e.g., as assessed by either IRT or SEM). Consequently, readers of the literature should be aware of whether equivalence is assessed before such comparisons are made.

This review found few published studies empirically assessing the measurement equivalence of on-line/web and paper methods. Though not specifically directed at a web application, a 1995 study by King and Miles did empirically assess equivalence and found no evidence to suggest that responses collected from paper-and-pencil and computerized administrations were not equivalent. Stanton's (1998) study was one of the first published investigations that specifically examined measurement equivalence of web versus paper and pencil. In short, Stanton found that the two methods were equivalent.

Following the same analytical protocol as Stanton (1998), Ryan, et al., (1999) and King and Miles (1995) both Fenlason (2000) and Magnan, Lundby and Fenlason (2000) found measurement equivalence between web and paper-and-pencil survey versions.

Young, et al., (2000) examined equivalence using analyses based on IRT. They found that "The DTF and DIF tests resulted in no consistent relationship between mode of response and response patterns" (p7).

In sum, in the cases where direct comparisons of paper and web-based surveys have been conducted, it appears reasonable to expect that web and paper-and-pencil forms will result in equivalent measures.

Practice and Research Implications

The studies reviewed here provide researchers, practitioners, and clients with a good foundation upon which to build a richer understanding of dual-method (web and paper) survey efforts. The studies to date, are informative and tend to be encouraging in terms of the comparability of web and paper results. Yet, this area of study is still relatively new, and as such, relevant studies should be replicated in order to confirm these initial findings and test additional hypotheses.

Of course, it is unrealistic to think that organizations will cease using dual-method surveys until all the research questions are answered. The challenge is to use the initial research to guide current survey practice, as well as to integrate pertinent research questions into ongoing survey efforts. In this spirit, the following paragraphs provide some suggestions for both the practice of dual-method (web and paper) surveying and the continued research that must be conducted.

Continue to Examine Quantitative Results Most studies to date have focused on interpreting differences in terms of item and scale means. While means are certainly a good indicator of the central tendency of the underlying response distribution, they only tell part of the interpretive story. Few investigators have examined possible differences in commonly used response

percentages, These include “top-box” scores, as used in customer satisfaction surveys and percent favorable and unfavorable responses as often found in organizational surveys. As Young, et al., (2000) note, the same mean scores can result from different combinations of favorable, neutral and unfavorable percentages. Given that many survey programs make heavy use of collapsed percentages such as percent favorable and unfavorable, it is important to examine the possible effects of method on this form of the data. Finally, other forms of survey metrics are also likely to be informative and diagnostic (e.g., Church’s, 2001 examination of variance under different methods).

Explore Comment Issues Further An emerging finding about web-based surveys is the tendency for respondents to write more when completing open-ended questions, than with paper-and-pencil surveys. Consequently, end users of web survey data may well have a higher volume of comments to review than if the same survey had been executed on paper. Anecdotally, it does appear that end users notice and react to this difference, as the authors have already had several conversations with clients who have inquired about the possibility of limiting the length of comments submitted by web respondents.

One solution to these end user concerns about length would be to simply limit the space available much as it is limited on the paper forms (e.g., by placing a character or word count limit on the input submitted). However, such changes ought to consider the possibility of negative respondent reactions in terms of their expectations of how much they can or should be allowed to write.

Aside from a simple numerical limit there may be other more subtle ways of limiting the volume of comments. For instance, employees could write as much or as little as they wished, but a word count could be shown as the comment is typed as a form of feedback on length. Or, the comment box into which respondents type could be expanded, providing more visual feedback about how much they are writing without having to scroll up or down within the box.

The length of comments is not simply an issue of personal preference or heavier workload placed on the end user. The authors’ practice experience has shown that comments can have a sometimes profound emotional impact on the reader or object of the comments. An area in which this implication is probably most clear is in the use of multi-source feedback surveys. In this survey application the comments are, by definition, highly personalized. Van Velsor (1998) noted that comments in multi-source

reports may even “override the impact of the actual [quantitative] scores” (p.192). Given the potential for increased negative reactions (due to greater length) it is important to establish if the valence (positive or negative tone) of comments is also affected by method. Similar questions can be asked of content of the comments. Are they more wide-ranging or do web-based comments tend to go into more depth on a few select issues?

Take a Closer Look at Completion and Response Rates In terms of item-level completion rates, researchers and practitioners need to remain aware of the possible differential use of Don’t Know/Not Applicable responses versus respondents simply skipping questions. Not distinguishing between these categories of non-response is likely to cloud the interpretation of item-level non-response conclusions. This is an especially critical distinction when the surveys being studied are multi-source feedback instruments because raters are often given instructions to use a DK/NA option if they have insufficient opportunity to observe the behavior in question.

Church’s (2001) work in dissecting completion rates by examining NA/DK responses separately from actual skipped items is an important example to follow. Future investigations should continue to treat these responses separately where possible because indiscriminately combining them could mask important differences. Certainly, until a larger body of empirical work has been developed, generalizing about completion rates should be done carefully and should explicitly spell out what constituted a non-response to an item.

It appears from these initial investigations that more work can be done in assessing overall response rates via web versus paper-based surveys. As noted earlier, overall response rates have been infrequently reported, often it seems due to limitations in the administration process (e.g., survey respondents have a choice of which method to complete, method is not tracked). It is possible that more specific results can be reported given that several current survey programs involve the use of tracking or “ticket” numbers that are unique identifiers for each survey respondent, whether taken by web or paper.

In addition, there is a rich survey literature concerning methods for increasing response rates (e.g., pre-notifications, follow-up reminders). Many of these techniques are well known and researched in other survey methods and this research should be replicated with web surveys (see Summers and Groehler, 2000 for one such account).

Examine Results Closely for “Sub-Group” Differences Dual-method survey users should be cautious when

interpreting differences as being caused by method before ruling out the influence of demographic variables. As Yost and Homer (1998) and Magnan, Lundby, and Fenlason (2000) demonstrated, meaningful differences in scores can evaporate when membership in a particular sub-group is taken into account. Based on these results, practitioners and researchers should strive to include, and evaluate, the effect of relevant demographic variables such as job level, production/non-production sites, management/non-management, etc. Church's (2001) method of evaluating the contribution of method over and above the contribution of demographic variables, offers one such approach to evaluating this influence.

Continue to Research Measurement Equivalence In the cases where measurement equivalence of web and paper-based surveys has been explicitly assessed (e.g., Stanton, 1998; Fenlason, 2000; Magnan, Lundby and Fenlason, 2000) equivalence has indeed been indicated. This is encouraging news for those who are interested or compelled to use both web and paper methods of data collection. While the evidence so far implies that web and paper-based data may be combined and compared without regard to administration method, these few studies should not be considered a "green light" to proceed without looking back. Practitioners and researchers should continue to conduct well-reasoned studies that add to the growing body of research directly addressing the measurement equivalence of dual-method surveys.

References

Bachman, D., Elfrink, J., & Vazzana, G. (1996). Tracking the progress of e-mail vs. snail-mail. *Marketing Research, Summer*, 30-36.

Booth-Kewley, S. Edwards, J. E. & Rosenfeld, P. (1992). Impression management social desirability, and computer administration of attitude questionnaires: Does the computer make a difference? *Journal of Applied Psychology*, 77, 4, 542-566.

Christianson DeMay, C., & Armstrong, S. (2000). Dual-process surveying: A technical perspective. In W. Macey (Chair). *Making the move to internet-based surveys: Lessons from experience*. Practitioner forum conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2000. New Orleans, LA.

Christianson DeMay, C. & Toquam Hatten, J. (2001). 360-degree assessment at the Boeing company: A dual-process (web/paper) approach. In M. Harris (Chair). *The internet and I/O psychology: Applications and issues*. Symposium conducted at the 16th Annual

Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2001. San Diego, CA.

Church, A. H. (2001). Is there a method to our madness? The impact of data collection methodology on organizational survey results. *Personnel Psychology*, 54, 937-969.

Dillman, D. A. (2000). Internet and interactive voice surveys. Chapter 11 in *Mail and internet surveys: The tailored design method*. Second Edition. New York: John Wiley & Sons. 352-412.

Fenlason, K. J. (2000). Multiple data collection methods in 360 feedback programs: Implications for use and interpretation. In T. Anselmi (Chair). *Current issues and challenges in the use of survey-based data*. Symposium conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2000. New Orleans, LA.

Joreskog, K. & Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the Simplis command language*. Chicago: Scientific Software.

Kiesler, S. & Sproull, L. (1986). Response effects in the electronic survey. *Public Opinion Quarterly*, 50, 402-413.

King, W.C., & Miles, E.W. (1995). A quasi-experimental assessment of the effect of computerizing noncognitive paper-and-pencil measurements: A test of measurement equivalence. *Journal of Applied Psychology*, 80, 6, 643-651.

Kraut, A. I., Oltrogge, C. G., and Block, C. J. (1998). Written vs. telephone surveys of employees: Are the data really comparable? A poster paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas, TX, April, 1998.

Kraut, A. I., & Saari, L. M. (1999). Organization surveys: Coming of age for a new era. In A. I. Kraut and A. K. Korman (Eds.). *Evolving practices in human resource management* (pp. 302-327). San Francisco: Jossey-Bass.

Kuhnert, K. & McCauley, D. P. (1996). Applying alternative survey methods. In Kraut (Ed.), *Organizational surveys: Tools for assessment and change* (pp. 1-17). San Francisco: Jossey-Bass.

Lautenschlager, G. J. & Flaherty, V. L. (1990). Computer administration of questions: More desirable or more social desirability? *Journal of Applied Psychology*, 75, 3, 310-314.

Macey, W. B. (1996). Dealing with the data: Collection, processing and analysis. In Kraut (Ed.), *Organizational surveys: Tools for Assessment and Change* (pp. 204-232). San Francisco: Jossey-Bass.

Magnan S. M., Lundby K. M., & Fenlason, K. J. (2000). Dual media: The art and science of paper and Internet employee survey implementation. In T. Anselmi (Chair). *Current issues and challenges in the use of survey-based data*. Symposium conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2000. New Orleans, LA.

Mehta, R., & Sivadas, E. (1995). Comparing response rates and response content in mail versus electronic mail surveys. *Journal of the Market Research Society*, 37, 4, 429-439.

Oppermann, M. (1995). E-mail surveys: Potentials and Pitfalls. *Marketing Research, Summer*, 29-33.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Journal of Applied Psychology*, 114, 3, 552-566.

Rosenfeld, P., Giacalone, R. A., Knouse, S. B., Doherty, L. M., Vicino, S. M., Kantor, J. & Greaves, J. (1991). Impression management, candor, and microcomputer-based organizational surveys: An individual differences approach. *Computers in Human Behavior*, 7, 22-32.

Ryan, A. M., Chan, D., Ployhart, R. E. & Slade, L. A. (1999). Employee attitude surveys in a multinational organization: Considering language and culture in assessing measurement equivalence. *Personnel Psychology*, 52, 37 - 58.

Schaefer, D. & Dillman, D. (1998). Development of a standard e-mail methodology: Results of an experiment. *Public Opinion Quarterly*, 62, 378-397.

Schuman, H. & Presser, S. (1981). *Questions and Answers in Attitude Surveys*. San Diego: Jossey Bass.

Spera, S. D. & Moye, N. A. (2001). Measurement equivalence between paper and web survey methods in a multinational company. A poster paper presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2001. San Diego, CA.

Sproull, L. S. (1986). Using electronic mail for data collection in organizational research. *Academy of Management Journal*, 29, 159-169.

Stanton, J. M. (1998). An empirical assessment of data collection using the internet. *Personnel Psychology*, 51, 709-724.

Summers, L. & Groehler, L. (2000). High-tech adventures in boosting response rates. In M. Sederberg and S. Rogelberg (Chairs). *Improving the survey effort: Methodological questions and answers*. Symposium conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2000. New Orleans, LA.

U.S. Department of Commerce -- Economics and Statistics Administration (2002). A nation online: How Americans are expanding their use of the internet. www.esa.doc.gov/508/esa/nationonline.htm

Van Velsor, E. (1998). Designing 360-degree feedback to enhance involvement, self-determination, and commitment. Chapter 7 in *Maximizing the value of 360-degree feedback*. Walter Tornow, Manuel London and CCL Associates. 149 - 195.

Weible, R. & Wallace, J. (1998). Cyber Research: The impact of the Internet on data collection. *Marketing Research, Fall*, 19-24.

Woehr, D. J., Sheehan, M. K., & Bennett, W. Jr. (1999). Understanding disagreement across rating sources: An assessment of the measurement equivalence of raters in 360 feedback systems. Paper presented at the 14th Annual Conference of the Society of Industrial and Organizational Psychology, Atlanta, GA, April, 1999.

Young, S. A., Daum, D. L., Robie, C. and Macey, W. H. (2000). Paper versus web survey administration: Do different methods yield different results? In M. Sederberg and S. Rogelberg (Chairs). *Improving the survey effort: Methodological questions and answers*. Symposium conducted at the 15th Annual Conference of the Society for Industrial and Organizational Psychology, Inc., April, 2000. New Orleans, LA.

Yost, P. R., & Homer, L. E. (1998). Electronic versus paper surveys: Does the medium affect the response? Paper presented at the 13th Annual Conference of the Society of Industrial and Organizational Psychology, Dallas TX, April, 1998.