

Pre-screening Job Applicants with Interactive Voice Response and Web-Based Technologies: Are the Methods Equivalent?

Keith Hattrup
San Diego State University

Matthew S. O'Connell
Select International, Inc.

Jenine R. Yager
San Diego State University

This study investigated the degree to which Interactive Voice Response (IVR) and Web-based technologies result in differences in responses to initial pre-screening items used in personnel selection. Over 50,000 applicants for entry level manufacturing jobs completed several basic employment eligibility items, and items designed to measure turnover risk, using either an IVR or Web-based job application process. As hypothesized, scores in the IVR condition were generally lower than in the Web condition, indicating overall lower qualifications among IVR respondents. Differences between conditions were very small, however, suggesting that the two approaches are largely comparable. On the other hand, the two approaches drew somewhat different samples of participants. African Americans and Hispanics were more likely to use the IVR method, whereas Whites and Asians tended to use the Web approach more frequently.

The use of computer-assisted technologies in testing has grown significantly in popularity in recent years, owing to improvements in the efficiency and accuracy of test administration and scoring (Donovan, Drasgow, & Probst, 2000; Stanton, 1998). Interactive voice response (IVR) technology, in particular, has been recently identified as an efficient method for collecting simple response data from large numbers of participants (Corkey & Parkinson, 2002; Thornburg, 1998; Tourangeau, Steiger, & Wilson, 2002). IVR has many advantages compared to traditional paper and pencil testing or interview methods, including automatic logical branching, self-pacing, access to hard-to-reach participants, elimination of interviewer bias, automatic and digital recording of information, standardization, greater confidentiality, and reduced cost (Corkey & Parkinson, 2002).

Yet, little is known about the equivalence of data gathered using IVR technology versus other approaches. Although several studies have recently compared the use of IVR with traditional and computer-assisted interviewing techniques in public opinion surveys (e.g., Corkrey & Parkinson, 2002a, 2002b), research has not examined the effectiveness of IVR when used as an initial screening tool in personnel selection contexts. This is a significant practical limitation given

that a key concern for organizational decision makers is whether the scores obtained from an IVR would lead to the same conclusions as scores obtained through other approaches. As Anderson (2003) noted in a review of research on the use of computerized technologies in assessment, “the most central question (in terms of equivalence) is whether use of new technology produces the same quantity and quality of applicants for an organization” (p. 126). Thus, the purpose of the present research is to compare an IVR approach for screening job applicants with an internet-web approach.

Interactive Voice Response Technology in Personnel Screening

Interactive voice response uses computer-assisted technology for administration and recording of interview responses via telephone. With IVR, a human interviewer is replaced by a high-quality recorded script to which the interviewee responds. Responses can be provided using the keys on a touch-tone telephone, or with the use of voice recognition technology, the respondent can speak his/her answer directly using the telephone receiver. Its use in public opinion surveying and customer service applications has grown substantially, owing to advantages related to standardization and efficiency (Thornburg, 1998). Several studies have recently sought to examine the extent to which IVR technologies can be effectively applied in the context of personnel selection.

Using a sample of students, Bauer, Truxillo, Paronto, Weekley, and Campion (2004) compared the perceived fairness of IVR with face-to-face and live telephone interviews. As expected, because of its non-personal characteristics, the IVR interview was rated lower in terms of several procedural justice factors, including interpersonal treatment, two-way communication, and openness. Because the IVR did not differ from the other techniques in other dimensions of fairness, the authors concluded that there “do not appear to be any major negatives in terms of structural fairness” (p. 135) associated with IVR versus other interviewing techniques. The use of a student sample may limit generalizability of these findings to some degree, however.

Van Iddekinge, Eidson, Kudisch, and Goldblatt (2003) recently explored the use of IVR technology in administering and scoring biodata items for screening cashiers and courtesy clerks at a large supermarket chain. In their study, 1,043 applicants completed a 42-item biographical interview/questionnaire via telephone, using IVR technology. Analysis demonstrated significant relationships with a job performance criterion, and very little adverse impact against women and African Americans. The latter finding is important because it suggests that the use of this newer technology does not appear to result in substantially different hiring rates among participants from varying ethnic backgrounds, who may differ in terms of exposure to computerized technologies (cf. Van Iddekinge et al., 2003). The criterion-related validity of the IVR-biodata was lower than expected, leading the authors to call for additional research designed specifically to compare IVR and alternative test administration technologies.

In a more recent study, Shepard and Robie (2005) observed equivalence between personality measures administered via the Web versus IVR. Although there

was greater evidence of non-responding and missing data with the IVR, the two methods yielded equivalent measurement of the latent constructs that were measured by the two instruments. The authors called for additional research designed to examine a variety of outstanding issues, including differences in response styles, acquiescence, and measurement of various constructs. Indeed, given the small number of studies that have been conducted, additional research on the application of IVR as a prescreening tool in personnel selection is clearly warranted.

Focus of the Present Study

The present study examines whether responses to initial job applicant screening items are equivalent across IVR and Web-based approaches when used during the first stage of a multi-method employment hiring process. Of particular interest is whether the two methods yield equivalent mean scores on items designed to assess employment eligibility and overall turnover risk. Differences between the two methods would imply the existence of potentially important method effects related to the use of different pre-employment screening approaches. Such differences may suggest that the latent constructs measured by the instruments, including sources of trait-relevant and irrelevant variance differ between the methods. From a practical point of view, such differences would complicate the interpretation of scores obtained from different pre-screening methods, and would likely require the choice of a single method to avoid confounding selection decisions with pre-screening methods. In this context, an additional concern is the degree to which the two methods result in differences in adverse impact of the items when used to make selection decisions.

As part of a large scale recruitment campaign for a major U.S. automobile manufacturer, advertisements were placed in the local media and an application process was opened for candidates to apply for entry-level manufacturing positions. Candidates could complete the first stage of a multi-stage hiring process either by calling a toll-free line and completing the initial parts of the application through an IVR line, or by logging onto a web-site and completing the same pieces of the application online. Thus, of interest in the present study is the degree to which the administration modality used for the initial screening resulted in differential responses to the test items. Based on previous research, we expected to see some differences in the mean level response to each of the key items.

Specifically, a number of previous studies have demonstrated main effects of survey or test administration modality on responses to items that address sensitive or personal issues. Overall, those test administration methods that involve greater anonymity and confidentiality are often associated with higher levels of agreement to items measuring negative or socially undesirable behaviors (e.g. Kobak, Taylor, Dotti, Greist, Jefferson, Burroughs, Mantle, Katzelnick, Norton, Henk, & Serlin, 1997; Tourangeau & Smith, 1996). Compared with paper-and-pencil administration, computerized test administration has been linked with greater self-disclosure and candor (Davis, 1999; Turner, Ku, Rogers, Lindberg, Pleck, & Sonenstein, 1998); although several studies have observed that differences across test modalities are not significant (Knapp & Kirk, 2003; Millstein, 1987). In a study comparing IVR and

telephone interviewing, Corkrey and Parkinson (2002a) observed higher levels of admissions of marijuana and alcohol abuse among participants completing the IVR. In virtually all of these previous studies, authors noted that the greater confidentiality and anonymity afforded by IVR technology may lead to higher levels of admission of negative behavior or wrong-doing.

Thus, in general, the IVR approach might be expected to yield different mean level responses on sensitive items as compared to a Web-based approach, given possible perceived differences in confidentiality and anonymity. Furthermore, when compared with normal Web-based surveying, IVR might be expected to yield different responses as a result of differences in the basic human-technology interface that the two methods use. In particular, the Web-based approach, like normal paper-and-pencil methods relies on visual-textual information; respondents read survey items and then provide a response using a keyboard or mouse. In contrast, IVR uses auditory information and respondents usually speak their responses. Thus, compared to the Web-based approach, the IVR approach allows less time for respondents to read items and provide a response. In other words, the IVR interface has a greater resemblance to a normal telephone conversation; questions are asked and answers are expected within the context of a normal conversational flow. As a consequence, IVR respondents likely have less time to consider the question and their answer, and because normatively expected or desirable responses are likely less available in memory than truthful ones, responses in IVR assessments should be lower in social desirability.

Thus, we hypothesize the following:

H1: Compared to a Web-based pre-screening method, the IVR approach will lead to higher levels of disclosure of undesirable or negative behaviors.

Method

Sample

Participants included a total of 54,218 applicants who either completed the initial job pre-screening using the IVR system or the Web. Participants responded to advertisements placed in a variety of media, including newspaper, television, and radio, for applications for entry-level production team member positions at a major automobile manufacturer. The minimum requirements for application included an age of 18 years and a legal right to work in the U.S. There were no other requirements for applying for these positions. Of the sample of applicants, 23,092 completed the IVR version, whereas 31,126 completed the application using the Web-based system.

Gender information was available for 14,153 participants. Of this number, 76.7% of the participants were male. There was some difference in the gender distribution across the two methods. Specifically, 73.9% of the respondents who completed the Web version were male, whereas 69.8% of the IVR respondents were

male. Race information was available for 14,049 participants. Of this number, 22.7% were White, 7.0% were African American, 68.2% were Hispanic/Latino, 1.6% were Asian, and 0.5% were Native American. As might be expected given the distribution of wealth in the U.S. and access to internet resources, some differences in the use of the IVR versus Web-based system was observed across race groups. Specifically, of the African American applicants, 55.1% completed the Web-based application, whereas 81.7% of the Asian participants, 60.9% of the Hispanic/Latino sample, 64.4% of Native Americans, and 70.8% of the White/Caucasian participants completed the Web-based application. Of the participants who completed the IVR, 18.5% and 18.7% provided race and sex information, respectively. In contrast, of those who completed the Web system, 23.2% and 23.4% provided race and sex information, respectively.

Procedures

As noted, applicants were informed of the opportunity to apply for positions with this organization, and were invited to complete the first stage of a multi-stage hiring process either by calling a toll-free line and completing the initial parts of the application through an IVR line, or by logging onto a web-site and completing the same application online. Because of the high demand for these jobs, the window for initial applications was limited to 2 weeks. This was proven to be sufficient in that over 14,000 applications were taken in the first 24-hour period and, as noted above, over 50,000 valid candidates were logged by both systems by the end of the 2-week period. Both application mediums were available continuously and without interruption for the 14-day period.

Because an extremely high applicant turnout was expected, and because this was only to be used as the first step in the application process, the script used in both modalities was directed at a small number of variables. In the IVR condition, applicants were asked to speak and spell their names and addresses, which were then transcribed within 24 hours into the database. All applicants were asked to provide information on the following: (1) Whether or not they were 18 years of age or older, (2) whether or not they had a legal right to work in the U.S., (3) the number of full-time jobs they had in the past 5 years, (4) the number of times they had been terminated (not laid off) in the past 5 years, and (5) the number of unexcused and unscheduled absences they felt were acceptable during a 12-month period. The latter three items were chosen to provide separate pieces of information relevant to an applicant's turnover risk, based on the premise that past behavior is an effective predictor of future behavior. In the present sample, intercorrelations among the three items ranged from .05 to .08, indicating very little commonality among the items. Hence, we treated the items separately in the analyses reported below, rather than forming a composite scale. This is consistent with the practical decision to treat each item as a separate screening assessment, rather than forming an overall composite measure from the three items. Additional information regarding work history, etc. was gathered as part of the second step in the selection process, which is not discussed in this paper.

Table 1
Proportions of respondents who were older than 18 years and eligible to work in the U.S.,
and means and standard deviations of turnover risk items for participants who provided
and those who did not provide gender and ethnicity information

Group	<u>Proportions</u>		<u>Prior Jobs</u>		<u>Terminations</u>		<u>Absences</u>	
	>18 years	Eligible	Mean	SD	Mean	SD	Mean	SD
IVR								
Information provided	97.8	99.9	1.69	0.93	0.02	0.18	1.23	1.38
Information withheld	96.8	99.8	2.07	2.34	0.18	1.09	1.86	2.77
Web								
Information provided	99.8	100	1.53	0.89	0.03	0.46	0.91	1.25
Information withheld	99.5	100	1.70	1.11	0.16	0.59	1.32	2.10

Results

Table 1 presents the summary statistics for each of the pre-screening items among participants who provided sex and race information and among those who withheld information about demographic status. Tables 2 and 3 provide summary statistics for the pre-screening items in each testing condition among participants in each of the different sex and race groups. Results of statistical tests of differences in item responses across test modes, race, and gender are presented in Table 4. As can be seen in the tables, there were generally very small differences across the two test administration modalities in responses to the items. As can be seen in Table 1, scores for participants who provided sex and race information were generally better than for participants who withheld sex and race information on each of the items, except the item that addressed eligibility for U.S. employment.

Although significantly fewer respondents who used the IVR than the Web met the minimum eligibility criteria (i.e. being older than 18 and eligible for employment in the U.S.), the overall magnitude of these differences was very small, particularly in the case of the item assessing eligibility for U.S. employment. The main effects of race and of gender, and the interactions involving these variables were not statistically significant, despite the large sample sizes. The values shown in Table 2 indicate that adverse impact would not occur in any comparison involving the groups for either test modality. For example, in the case of the IVR, the adverse impact ratio for Asian participants would only be 96.9 / 98.1, or 98.8. Thus, in terms of basic employment eligibility criteria, the two methods yield nearly identical pools of candidates.

Means and standard deviations of the three turnover risk items are presented in Table 3. Although the main effect of test modality on mean number of reported previous jobs was non-significant in the tests involving race groups, the same main effect was significant in analyses involving gender groups, owing to differences across the analyses in missing data. As can be seen in Table 3, participants who used the IVR reported having significantly more previous jobs than Web participants. However, this difference was very small, accounting for only 0.7% of the overall variance in responses to the item. A main effect of gender in the same analysis

revealed that women held fewer previous jobs than men; however, in this case, gender accounted for only 0.4% of the overall variance in this item. A main effect of test modality was also observed for the number of absences that respondents considered acceptable; as expected, more absences were reported in the IVR condition. However, again, this difference was practically trivial, accounting for only 1.2% of the variance at most. Main effects of race and of gender were also observed in the number of absences considered acceptable; these effects accounted for 0.4% and 1.2% of the variance, respectively. Interactions involving sex or race were non-significant for the absence item. No significant differences were observed across the conditions in the number of previous terminations.

Adverse impact analyses were done for these three items using an overall sample selection ratio of .50. This was based on the idea that the present pre-screening items would be used to reject the bottom 50% of the sample of applicants, allowing the remaining participants to proceed to the next step of the hiring process. In the present case, despite some differences in the distributions in the different testing conditions, neither the IVR nor the Web method resulted in a violation of the 4/5ths rule when decisions were based on the number of previous jobs held or the number of reported terminations. However, differences in adverse impact ratios were observed across the conditions for the number of absences that respondents thought would be acceptable. Using the White group as comparison, adverse impact ratios were .83, .95, .78, and .85 for African Americans, Asians, Latinos, and Native Americans in the IVR condition, but were .87, .86, .90, .92 for the same groups in the Web condition, respectively. Thus, use of the IVR method appears to result in some differences in selection decisions, particularly adverse impact, when participants are asked about the number of absences they think would be acceptable. Adverse impact ratios satisfied the 4/5ths rule in every other comparison, however, suggesting that overall, the results show considerable similarity in the kinds of practical decisions that might be made using the two methods.

Table 2
Proportions in each group who were older than 18 years of age and eligible to work in the United States

Group	Older than 18 years of age		Eligible to work in the U.S	
	IVR	WEB	IVR	WEB
Race				
African American	98.7	100.0	100.0	100.0
Asian	96.9	100.0	100.0	100.0
Hispanic/Latino	97.6	99.8	99.9	100.0
Native American	100.0	100.0	100.0	100.0
White	98.1	99.9	99.7	100.0
Sex				
Women	97.8	99.8	99.9	100.0
Men	97.8	99.9	99.9	100.0

Table 3
Means and standard deviations of turnover-risk items in each group

Item/Group	IVR			WEB		
	Mean	SD	d ^a	Mean	SD	d ^a
Number of Previous Jobs						
African American	1.77	0.96	0.10	1.54	0.86	-0.06
Asian	1.47	0.88	-0.25	1.63	0.86	0.05
Hispanic/Latino	1.69	0.95	0.01	1.51	0.89	-0.09
Native American	1.74	0.73	0.07	1.61	0.68	0.02
White	1.68	0.34		1.59	0.90	
Women	1.62	1.04	-0.10	1.42	0.95	-0.16
Men	1.72	0.88		1.57	0.86	
Number of Terminations						
African American	0.02	0.18	0.07	0.01	0.12	0.00
Asian	0.00	0.00	-0.08	0.03	0.21	0.18
Hispanic/Latino	0.03	0.20	0.11	0.02	0.15	0.07
Native American	0.00	0.00	-0.08	0.06	0.24	0.48
White	0.01	0.13		0.01	0.10	
Women	0.03	0.21	0.05	0.02	0.14	-0.03
Men	0.02	0.17		0.03	0.53	
Number of Absences						
African American	1.22	1.30	0.18	0.96	1.26	0.15
Asian	1.06	1.22	1.08	1.08	1.48	0.25
Hispanic/Latino	1.28	1.42	0.94	0.94	1.27	0.13
Native American	1.32	1.38	1.08	1.08	1.51	0.26
White	1.00	1.23	0.78	0.78	1.17	
Women	1.33	1.33	0.10	0.98	1.24	0.07
Men	1.19	1.40		0.89	1.26	

^ad is the standardized mean difference between groups. d values represent comparisons of each race group with the White group, and women with men

Discussion

This study was designed to address the practical question of whether IVR technology results in different mean level responses to items designed as an initial pre-screening for job applicants as compared to a Web-based approach, and the degree to which such differences may translate into differences in adverse impact. Although considerable interest has recently been expressed in the use of IVR technology for assessing job applicants, very little research of a comparative nature has been conducted to evaluate the extent to which IVR systems elicit different responses from potential job applicants compared to more traditional methods. Based on a review of previous research on the use of IVR in public opinion surveying, we hypothesized that the IVR approach would yield a higher proportion of participants claiming negative or socially undesirable behaviors than a Web-based approach.

Table 4
Results of tests of differences in response by test mode, race, and sex

Effect	Chi-square values			F-ratios	
	18 or older	Eligible	Jobs	Terminations	Absences
Race					
Test mode	126.44**	8.73*	2.22	0.33	5.53*
Race	4.51	3.05	1.58	0.72	11.32*
Test mode*Race	1.43	0.00	2.10	0.17	1.42
Sex					
Test mode	129.92**	10.07*	77.90**	0.12	134.49*
Sex	0.03	0.27	43.20**	0.02	16.55*
Test mode*Sex	0.35	0.00	1.20	1.22	0.58

Note: Results of main effects of test mode differ between the two sets of analyses due to missing data

* $p < .05$ ** $p < .01$

As expected, respondents who completed the IVR system reported having significantly more previous jobs and tended to think that more unexcused absences were acceptable, as compared to participants who completed the Web-based system. Furthermore, of the IVR participants, a significantly greater proportion were under the age of 18 or were not legally eligible to work in the U.S. However, it is also important to point out that the differences between conditions, although statistically significant, were very small. In the case of the two eligibility items, differences across conditions in the proportions of eligible applicants varied from about 3% to only 1/10th of a percent, clearly not a large enough differences to cause practically meaningful differences across test modalities in selection outcomes, including adverse impact. Although statistically significant, the proportion of variance in the numbers of previous jobs reported and the numbers of absences considered acceptable that was explained by test modality was also trivial, ranging from less than 0.1% to 1.2%. In the case of the number of absences that respondents reported, differences in test modality in the shapes of the distributions was associated with more noticeable differences in adverse impact ratios, however. In fact, using an overall selection ratio of .50 for this item would result in a violation of the 4/5ths rule in the comparison involving Whites and Latinos who use the IVR version, but the same comparison would not lead to a violation of the 4/5ths rule for Web-based participants. This result needs to be interpreted with caution, however, given that the two screening methods drew somewhat different samples of applicants in terms of demographics and other variables, as discussed below. Further, differences in adverse impact ratios between the methods were small, and in only one case did the methods result a different determination about whether the simple 4/5ths rule might be violated.

Thus, overall, the results of the present research are somewhat consistent with the notion that IVR technology, as compared to Web-based technology, is perceived as offering greater anonymity and confidentiality, thereby resulting in greater

disclosure of negative and undesirable behaviors. However, differences observed across conditions in the present study were very small overall. Despite very similar means, however, the distributions of responses to the items varied somewhat across conditions, and in some cases resulted in slightly different adverse impact ratios. This suggests the need for additional research to better ascertain how and why responses to Web-based and IVR tests might differ. Taken together, results of the present study suggest that the two approaches yield very comparable information about mean levels of responses to items designed to measure basic employment eligibility and turnover risk.

Further, it is clear from the present findings that the two approaches draw somewhat different samples of applicants. Perhaps because of differences in access to internet resources, and experience with the World Wide Web, applicants who choose to apply using a Web-based system rather than an IVR, may tend to represent a greater proportion of individuals from Asian and White/Caucasian groups. In contrast, African American and Hispanic/Latino applicants may be somewhat more likely to utilize an IVR system, rather than the Internet, when applying for entry level production jobs. Yet, importantly, race and gender did not interact with test condition in the present study, meaning that differences between the groups in item responses do not depend on test modality and vice versa. Moreover, a number of the statistical tests of the main effects of test modality on item responses were significant after controlling for race and gender, indicating that test modality effects were partly independent of differences in respondent demographics across the different conditions. Overall, surprisingly few respondents (21%) provided information about sex and race, with more providing the information on the Web-based system. This has implications for tracking applicants in this organization. For example, adverse impact statistics may be calculated, but may have some undetermined generalizability to the entire population of applicants.

Like all empirical research, the present research has a few limitations that are important to note. Most importantly, as in other research designed to examine the equivalence of IVR and Internet-based selection techniques (Shepard & Robie, 2005), participants self-selected into the different conditions of the study. Specifically, applicants for the jobs advertised in the present context were allowed to choose whichever technological interface, Web or IVR, that they preferred. This, of course, means that there may be uncontrolled differences between the samples in access to a computer or telephone, or comfort with the techniques, and so on. Yet, despite these potential sample differences, the two methods yielded remarkably consistent information about the overall level of responses to the pre-screening items. Clearly, the present study is a small-scale investigation of the effects of test modality, specifically IVR versus Web-based screening, on responses to items used for initial assessment of employment eligibility and overall turnover risk. Future research is needed that builds upon the results of the present investigation to explore how test modality, applicant characteristics, and test construct interact to effect test responses and selection decisions.

All in all, the present study provides encouraging evidence on the usefulness and comparability of IVR technology as an initial pre-screening tool, as compared with a traditional Web-based approach. Although some differences in distributions

of responses to sensitive items are likely to be observed, the overall magnitude of these differences does not appear to be great. From a practical point of view, the use of IVR appears to yield useful and comparable information relative to other methods in the present context, at no greater financial cost. Organizations like the one that participated in the present study may benefit from broader application of IVR in personnel selection, accompanied by ongoing research designed to evaluate the equivalence of IVR with other more traditional approaches.

References

- Anderson, N. (2003). Applicant and recruiter reactions to new technology in selection: A critical review and agenda for future research. *International Journal of Selection and Assessment*, *11*, 121-136.
- Bauer, T. N., Truxillo, D. M., Paronto, M. E., Weekley, J. A., & Campion, M. A. (2004). Applicant reactions to different selection technology: Face-to-face, interactive voice response, and computer-assisted telephone screening interviews. *International Journal of Selection and Assessment*, *12*, 135-148.
- Corkrey, R., & Parkinson, L. (2002a). A comparison of four computer-based telephone interviewing methods: Getting answers to sensitive questions. *Behavior Research Methods, Instruments, & Computers*, *34*, 354-363.
- Corkrey, R., & Parkinson, L. (2002b). Interactive voice response: Review of studies 1989-2000. *Behavior Research Methods, Instruments, & Computers*, *34*, 342-353.
- Davis, R. N. (1999). Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments & Computers*, *31*(4), 572-577.
- Donovan, M. A., Drasgow, F., & Probst, T. M. (2000). Does computerizing paper-and-pencil job attitude scales make a difference? New IRT Analyses offer insight. *Journal of Applied Psychology*, *85*(2), 305-313.
- Kobak, K. A., Taylor, L. V., Dotti, S. L., Greist, J. H., Jefferson, J. W., Burroughs, D., Mantle, J. M., Katzelnick, D. J., Norton, R., Henk, H. J., & Serlin, R. C. (1997). A computer-administered telephone interview to identify mental disorders. *Journal of the American Medical Association*, *278*, 905-910.
- Knapp, H., & Kirk, S. A. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered surveys: Does methodology matter? *Computers in Human Behavior*, *19*(1), 117-134.
- Millstein, S. G. (1987). Acceptability and reliability of sensitive information collected via computer interview. *Educational and Psychological Measurement*, *47*(2), 523-533.
- Shepard, W., & Robie, C. (2005). Equivalence of tests administered on computer versus interactive voice response (IVR). In F. Drasgow (Chair), *Innovations in Computerized Assessment: Research on Practical Issues*. Symposium presented at the 20th Annual Conference of the Society for Industrial and Organizational Psychology, April, Los Angeles.

- Stanton, J. M. (1998). An empirical assessment of data collection using the Internet. *Personnel Psychology, 51*(3), 709-725.
- Thornburg, L. (February, 1998). Computer-assisted interviewing shortens hiring cycle. *HR Magazine, 73-79*.
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question content. *Public Opinion Quarterly, 60*, 275-304.
- Tourangeau, R., Steiger, D. M., & Wilson, D. (2002). Self-administered questions by telephone: Evaluating interactive voice response. *Public Opinion Quarterly, 66*, 265-278.
- Turner, C. F., Ku, L., Rogers, S. M., Lindberg, L. D., Pleck, J. H., & Sonenstein, F. L. (1998). Adolescent sexual behavior, drug use, and violence: Increased reporting with computer survey technology. *Science, 280*(5365), 867-873.
- Van Iddekinge, C. H., Eidson, C. E., Kudisch, J. D., & Goldblatt, A. M. (2003). A biodata inventory administered via interactive voice response (IVR) technology: Predictive validity, utility, and subgroup differences. *Journal of Business and Psychology, 18*, 145-156.