Appalachian
STATE UNIVERSITY®
BOONE, NORTH CAROLINA

Assessing The Measurement Invariance Of Latent Job Satisfaction Ratings Across Survey Administration Modes For Respondent Subgroups: A MIMIC Modeling Approach

Authors:
Brian G. Whitaker & Jessica L. McKinney

Abstract

To date, many studies have assessed the measurement invariance of a wide variety of measures across Internet and paper-and-pencil conditions; however a relative dearth exists in the literature investigating measurement invariance across administration modes for differing subgroups of respondents. Using MIMIC modeling, this study assessed whether gender and age of the respondent systematically influenced responding according to administration mode above and beyond measurement invariance. Consistent with past research, this study demonstrated that job satisfaction ratings were indeed measurement invariant across Internet and paper-and-pencil conditions, however, older respondents tend to differentially rate job satisfaction according to administration mode. Implications are discussed .

Introduction

Measurement is often defined as the systematic assignment of values on constructs to represent characteristics of persons, entities, or events (Vandenberg & Lance, 2000). In the field of industrial/organizational psychology, researchers and practitioners employ the use of differing types of measures in order to characterize persons, groups, and/or organizations. Measurement has gained research attention over the last decade or so, most notably in the industrial/organizational psychology literature, specifically because it delineates, defines, and gauges the degree of congruence between theory and obtained data. Recent advances in theory in conjunction with the advent of more sophisticated analytical/ measurement tools has allowed many researchers to begin to address the issue of measurement invariance across populations, subgroups, or experimental conditions.

Measurement invariance is most often tested using multi-group confirmatory (MGCFA), which takes place over a number of successive steps involving a series of nested models. At each step, an additional constraint of invariance is added to subsets of parameters (i.e., factor loadings, error variances, factor variances) and the fit of the resulting model is assessed to determine if the forced invariance has worsened the fit of the model (Vandenberg & Lance, 2000).

Due to any number of interpersonal, social, cultural, or demographic differences, respondents may not approach, and thus respond to, a given measurement instrument in the same way. As such, to date, researchers and practitioners have examined measurement invariance over several types of grouping conditions including multisource performance appraisal ratings sources (Cheung, 1999; Woehr, Sheehan, & Bennett, 1999), gender (Whitaker, 2004a), and across cultures (Lam, Hui, & Law, 1999; Ployhart, Wiechmann, Schmitt, Sacco, & Rogg, 2003).

Recently, researchers have also begun investigating the extent to which differing employee attitude survey administration modes affect measurement invariance as more and more researchers and practitioners have turned to Internet-based administration as a method for collecting survey/questionnaire data as opposed to the traditional paper-and-pencil modes of administration (Church, 2001; Thompson, Surface, Martin, & Sanders, 2003). While disadvantages of Internet-based assessment do exist (i.e., multiple submissions, lack of experimental control, drop- out, questionable external validity; Kranzt & Dalal, 2000; Reips, 2002), the move to Internet-based survey administration is not surprising given the reported advantages of collecting data over the Internet. For example, Internet- based assessment grants one access to a relatively large number of participants, better generalizability, and cost savings over traditional paper-and-pencil surveys (for reviews, see Barak & English, 2002; Buchanan, 2001; Reips, 2002; Thompson et al., 2003; Yun & Trumbo, 2000). Despite the obvious widespread utilization of the

Internet as a data collection medium, in practice one frequently overlooked and untested assumption is that the measurement instrument in use functions the same across Internet and paper-and-pencil administration methods (i.e., measurement invariant; Cheung, 1999). However, if the measurement instrument functions differently across administration modes, those collecting survey data should feel compelled to analyze and report the data separately according to administration method. If separate analyses are not done and nonequivalent data sets are combined a true picture of employee attitudes may not emerge. Thus, an unnecessary intervention may be implemented on the basis of wrongful data compilation or conversely, a much-needed intervention may go unimplemented. Either alternative could prove costly for an organization. If, on the other hand measurement invariance does indeed exist for a given measurement instrument, those concerned may focus their attention on the interpretation of the combined results and apply interventions as necessary, resting assured that both data sets are statistically comparable.

Recent research has indicated that surveys may not function in the same way across administration mode. For example, Joinson (200 I) demonstrated that participants who perceive themselves to be visually anonymous and experience heightened feelings of private self-awareness tend to disclose more personal information than those who perceive themselves non-anonymous with reduced private self-awareness. Given that paper-and-pencil administration is usually carried out in the presence of an administrator while Internet surveying usually lacks this control (Barak & English, 2002), one may expect that survey respondents would experience feelings of anonymity and private self-awareness to differing degrees based on administration mode. If this were the case, survey respondents would likely self-disclose to differing degrees based on administration mode. Similarly, respondents may respond differentially to surveys based on various technical characteristics present during Internet administration. For example, survey respondents may use differing Web browsers, different monitor sizes, or Internet connections. If researchers ignore this technical variance, error may be introduced with an attendant disruption in data quality (Reips, 2002).

While the aforementioned literature suggests that elements may exist with the potential to impact cross-mode measurement equivalence, and indeed some researchers have demonstrated systematic differences according to administration mode (Dillman & Bowker, 200 I ; Reips, 2002b), fortunately, to date, research has generally demonstrated that survey respondents view the dimensionality of the constructs similarly and that items measure the constructs in the same way across administration mode. Indeed measurement invariance across Internet and paper-and- pencil administrations has been replicated for measures such as self-monitoring (Buchanan & Smith, 1999), personality (Buchanan, 200 I ; Pasveer & Ellard 1998), self-efficacy & technology course satisfaction (Burnkrant
& Taylor, 2001), organizational justice (Stanton, 1998), 360 feedback (Fenlason, 2000), selection (Ployhart, et al., 2003), performance appraisal (Maurer, Raju, & Collins,

1998), and employee attitudes (Etchegaray, Sexton, Helm-reich, & Thomas, 2003; Spera & Moye, 2001 ; Whitaker, 2004b). Furthermore, evidence exists suggesting that the two administration modes may be qualitatively equivalent as well. Preckel and Theimann (2003) found that item discrimination, reliability estimates, and criterion-related validity were remarkably similar for an online and paper-and-pencil measure of intellectual giftedness. Similarly, Meyerson and Tryon (2003), investigating the extent to which construct validity is affected using differing administration methods, demonstrated similar estimates of internal consistency and equivalent validity coefficients with other relevant scales. Importantly, these results suggest that the nomological network of a given measure may hold up well across Internet and paper-and-pencil administrations. Thus, it seems reasonable to assume that, as a whole, respondents use the same psychological metric when responding to Internet or paper-and-pencil surveys.

To begin with, the present study re-examines the invariance of ratings over Internet and paper-and-pencil administration modes. Given the preponderance of research findings indicating that ratings generally remain invariant, the following hypothesis is proposed:

HYPOTHESIS ONE
Job Satisfaction Ratings Will Demonstrate Measurement Invariance Over Internet and Paper-and-Pencil Administration Modes

While identifying any potential sources of invariance across administration methods is of great practical importance, this study also seeks to extend the findings of the aforementioned literature by investigating whether differing subgroups respond equivalently when providing ratings via paper-and-pencil and Internet administration method.

To date, very few researchers have investigated whether ratings made by members of differing subgroups are invariant with respect to data obtained from differing administration methods. I n the sole study carried out assessing subgroup cross-mode rating tendencies, Reips and Franek (2004) investigated whether anonymity, voluntariness, and age affects latent job satisfaction ratings. Their results suggest that across administration mode (Internet vs. paper-and-pencil), age remains positively correlated with latent job satisfaction ratings. This study notwithstanding, the dearth of research into cross-mode subgroup ratings invariance stands in stark contrast to the vast literature indicating that a) males and females differ with respect to computer anxiety, b) age is correlated with computer anxiety and c) technical variance impacts survey responding.

Factors Impacting Subgroup Cross-Mode Invariance
Many studies have demonstrated that females report higher levels of computer anxiety than males (Brosnan, 1998; Liu, Reed & Phillips, 1992; Rosen & Weil. 1994). Furthermore, these effects have been demonstrated in a variety of environments and across age groups (Wi lliams, Ogletree, Woodburn, & Raffeld 1993).

Computer anxiety has generally been conceptualized as a fear related to the use of computers, or feelings of intimidation and hostility toward this form of technology (Meier, 1985; Glass & Knight, 1988). Research has shown that computer anxiety is negatively related to performance on computer-based selection tests (Frericks, Ehrhart, & O'Connell, 2003) as well as performance on simple tasks that require the use of a computer (Brosnan, 1998; Mahar, Henderson, & Deane, 1997; Rosen & Weil, 1995). Furthermore, Mahar et al. (1997) demonstrated that those with higher levels of computer anxiety tended to complete simple computer tasks more slowly.

More recently, Durndell and Haag (2001), investigating the effects of gender on computer anxiety, demonstrated that males tend to report greater computer self-efficacy, lower computer anxiety, more positive attitudes toward the Internet and longer use of the Internet than females. The authors further argue that the literature on attitudes and anxiety toward computers is likely to extrapolate to the Internet. Thus, if females generally exhibit more computer anxiety than males, one may expect a disruption in the quality of data obtained from Internet administration, which in turn, may negatively impact ratings invariance across pencil-and-paper and Internet administration modes.

Similarly, another stream of research has shown that older employees consistently demonstrate higher levels of computer anxiety than younger workers (Honeyman & White, 1987; Johnson, 1987; Rosen & Weil, 1995). Researchers theorize that due to the wide-spread advent of computer technology, re-education becomes a very salient necessity for the older employee. In the employment setting, when progress in the individual career is threatened by lack of technological background the older employee facing the loss of a job or opportunity for advancement becomes a prime candidate for joining the ranks of the computer anxious.

Lastly, technical variance may affect latent responses as a function of demographic variables, particularly age. As previously outlined considerable technical variance in Internet-based studies exist as respondents may utilize computers with differing hardware/software characteristics (i.e., differing browsers, monitors, or network connections) which may ultimately contribute to a different appearance of, and responding to, online surveys. Working under the assumption that older respondents are more likely to use older or more outdated monitors or Web browsers than younger respondents, one may expect technical variance to influence the ratings of older respondents to a greater degree than those of younger respondents during online responding. Thus, a greater degree of technical variance for older employees may lead to ratings invariance across modes as a function of age.

Demographics and Job Satisfaction

The present study suggests that a disruption in data quality may be made manifest in the degree to which the background variables of gender and age correlate with latent job satisfaction differentially across the modes of administration. For example, it has been shown over the course of many studies that females and older employ-ees generally report higher levels of job satisfaction than males or younger employees on paper-and-pencil surveys (Bedeian, Ferris, & Kacmar, 1992; Gibson & Klein, 1970; Glenn, Taylor, & Weaver, 1977; Reips & Franek, 2004; Varca, Shaffer, & McCauley, 1983). While several explanations have been proposed for women's higher levels of job satisfaction, the explanation most commonly cited is that women may be socialized to refrain from expressions of discontent to a greater extent than men (Hodson, 1989). In terms of age, the ubiquitous link between age and job satisfaction is thought to result for from younger workers adopting higher expectations than older workers. Younger workers therefore become more frequently and more easily disappointed while older workers have abandoned their expectations, and are thus less unhappy (more satisfied). An alternative explanation for this effect is that those dissatisfied may choose to leave the organization, while those satisfied remain and grow older and perhaps move into more satisfying jobs than those held by younger workers.

Given the previously outlined findings, awareness of differential attitudes toward the use of computers by differing subgroups may be an important factor to consider when developing and designing Internet-based survey initiatives or programs. As research has shown that even minimal involvement in computer-related activities can become an anxiety-provoking situation for those suffering from computer anxiety (Paxton & Turner, 1984), those prone to computer anxiety may read instructions improperly, ignore disclaimers more frequently, or respond in a half-hearted manner to a greater extent than those who do not suffer from computer anxiety, thus potentially disrupting data quality. In the end, survey responses may be influenced by demographic variables that may be ascribed to the rater. For example, the rater's sex or age may influence latent job satisfaction variability, even if these variables operationally have nothing to do with the construct for which the rating is being made (Guion, 1983).

Given the recent surge in online surveying, it is not hard to imagine a scenario in which a researcher/practitioner gathers attitudinal data at Time I using traditional paper-and-pencil surveying techniques, then later implements Internet surveying at Time 2. If female employees' or older employees' computer anxiety negatively impacts responding to the Time 2 survey, cross-mode measurement equivalence may break down as females or older workers may no longer respond similarly or view the dimensionality of the constructs in the same way across administration mode. Thus, it is possible that observed attitudinal differences in data sets between Times I and 2 may lead one to conclude that a shift in attitudes had occurred for females or older employees where, in reality, any observed differences in responding may have simply been a reflection of and reaction to computer anxiety.

In order to assess the effects of administration mode on subgroup responding, the present study seeks to identify and test latent job satisfaction variability across administration mode using a model known as a MIMIC model (multiple indicators, multiple causes, Joreskog &

Goldberger, 1975). MIMIC models extend MGCFA by analyzing the indirect effects of background variables, exogenous variables, or any other type of covariate on latent satisfaction variability in order to further analyze measurement invariance and population heterogeneity (Muthen & Muthen, 2001). In a MIMIC framework, an indirect effect for a given covariate would indicate that mean differences for latent satisfaction variability were found as a function of differing levels of that covariate. Logically, satisfaction ratings inherently necessitate that evaluative judgments be made regarding a target construct. However, the administration mode employed for data collection may impact the manner in which responding is carried out for a given level of a grouping variable (i.e., gender or age)

This study seeks to further contribute to the literature by employing a repeated measures design to investigate the impact of differing survey methodologies on a widely-used job satisfaction measure, the Job Descriptive Index (JDI: Smith, Kendall, & Hulin, 1969). To date, very few researchers have investigated cross-mode equivalence using a repeated measures framework. Barak and Cohen (2002), investigating measurement equivalence of a popular measure of career interests, found that while test-retest correlations were relatively high over a six-week period, three of the six subscale means were elevated for the Internet condition. As this study illustrates, the fundamental advantage of the repeated measures design over a between subjects design is the reduction in error variance associated with individual differences (Robin- son, 1981). That is, the use of a repeated measures de- sign removes error variance from the denominator of the F-ratio test thereby increasing statistical power to detect mean differences. For the purposes of this study, if gender or age leads to small systematic differences in cross-mode responding, a repeated measures design may be required to detect cross-mode differences as a between-subjects design may lack the power to discern any existing differences. Using a repeated measures de- sign to detect small differences in cross-mode data is a nontrivial endeavor as organizations typically analyze data from one time period to the next using a repeated measures framework as a means of tracking important organizational variables over time. If small systematic differences in employee responding emerge due a change in administration mode from one time period to the next, an organization may perceive that a statistically significant shift in employee attitudes had changed over time, when in reality, this disruption occurred as a result of differential responding according to survey technique. Thus, detecting whether small changes in the data take place as a function of administration mode becomes particularly salient. Past research assessing measurement equivalence with between-subjects analyses has failed to take this into account. Given that organizations typically utilize repeated measures analyses, this study hopes to more closely reflect analyses more typically seen in applied settings. Thus, the following hypothesis is proposed:

HYPOTHESIS TWO
Latent Job Satisfaction Variability Will Differ For Gender and Age Across Internet and Paper-and-Pencil Administrations

Method

Participants. Two hundred and sixty nine employed MBA and doctoral students from a private, mid-sized Midwestern university were recruited to take part in the study. Of this number, survey responses to both the online and paper-and pencil survey were obtained from 117 participants, reflecting a response rate of 43.4%. Analyses revealed no differences in gender or age between respondents and non-respondents. This sample size is adequate for the factor analytic purposes of this study. Guadagnoli and Velicer (1988) recommend a sample size of between 100 and 200 for stable factor solutions, while Nunnally (1978) suggests a participant to item ratio of 10:1. While lacking three participants to fulfill the Nunnally (1978) criteria, we argue that, for the most part, both of these guidelines are satisfied with the current sample size.

The mean age of the participants was 29.37 (SD = 7.25) with an average organizational tenure of approximately 28.09 months, working an average of 27.6 hours per week. Participants held jobs in multiple categories including office and administrative support, education, and sales. The final sample was split between gender; male 53%, females 47%.

Measures. Participants were recruited by the primary experimenter (a male) and asked to anonymously complete all six items from each of two subscales of the JD! (*Supervision* and *Wt1rk Itself*) on both a paper-and-pencil and Internet version of the survey. Participants were asked to supply satisfaction ratings with their primary jobs in mind. That is, if a participant was employed in more than one position, he/she was instructed to make satisfaction ratings based on the position he/she considered primary. The JDI was chosen because it is considered one of the most carefully constructed measures of job satisfaction in use today (Rosznowski, 1989). These scales and items were chosen on the basis of Gregson's (1987) findings that using the six items that loaded highest on each dimension, even when converted into a five-point scale, loaded the same as they did for Smith. et al. (1969) when the test was developed. As such, researchers can abbreviate the JDI and use a Liker! scale (*strongly agree* to *strongly disagree*) without interfering with dimensionality. Furthermore, these scales demonstrated high internal consistency (.86 and .84. respectively), and do not inter-correlate very highly (*Suspension-Work Itself; .37*).

In order to minimize memory effects across administrations, the 12 items making up the three subscales were first embedded within 52 filler items for the paper-and-pencil administration, then embedded then embedded within 52 different filler items for the Internet administration. The 12 JD! items maintained their order and position between administrations, with the filler items serving to "mask"' their repeated presence. For this task, all filler items for both versions were taken from the Devine Inventory (Conoley & lmpara, 1995) and set to the same five-point scale as the JD! items. The Devine Inventory is a behavioral profiling instrument that taps 33 behavioral and personality descriptors including aggressiveness, sociability, initiative, etc. In order to n:ducc the possibility of confounding administration mode with filler item use, the JCI sub-facets were embedded within the same Devine Inventory construct for both administration modes. For example, if the JD! Supervisor factor was embedded within the Sociability facet of the Devine Inventory for paper-and-pencil ad- ministration, the Supervisor factor was embedded within the same Devine inventory scale for Internet administration, however, different Sociability items were used. This way, a balance was struck between minimizing memory effects, by using differing items across modes, while decreasing the confounding effects of this practice. As analysis was only carried out on the 12 JDI items, data from filler items were omitted from analysis. In order to model both traditional paper-and-

pencil and Internet surveying conditions while counterbalancing for order effects, the following procedure was used: During class time, sixty-two of the 1 17 participants (53%) completed the paper-and-pencil survey first. To each survey was attached an identification number which they were instructed to take with them as they would be asked to supply this number during Internet surveying. Once paper-and-pencil surveying had been completed participants were informed that they would receive an email (the address was supplied to the experimenter by the participants) one week from the then-current date that contained a link to the Internet survey. They were further informed that the Internet survey would remain accessible 24 hours a day, for a seven-day period and that the survey nay be taken at any point during this period. The email containing the embedded link and access password to the survey was sent via email one week later. While the embedded link was not individualized each participant was given a unique password to the survey site. This password was a number which also served as a code for identification purposes, allowing the researcher to pair the paper-and-pencil and Internet responses. In order to ensure that the differing surveys appeared as similar as possible to those with differing browsers and/or monitors, visual comparisons of the paper-and-pencil and Internet surveys were carried out over varying browsers and resolutions. In the end the paper-and-pencil and Internet versions of the survey were identical in format and length, however, minor variations in appearance were evident. For example, at higher resolutions, more items were displayed on-screen than at lower resolutions. Despite this, the only differences between the surveys were the masking items.

The other 55 participants (47%), received essentially the same treatment in the reverse order. Participants were informed that they would receive an email with an embedded link to the survey, which they had one week to complete. Participants were instructed that Internet surveying would be done first and paper-and-pencil administration would be done two weeks from the then-current date. The email sent to participants followed the same procedure as above. The emailed link took them to the instruction page, and the instruction page was linked to the survey page. Two weeks later, the experimenter returned to the class and paper-and-pencil administration was carried out following the same procedure outlined above.

Results

Following the steps outlined by Vandenberg and Lance (2000), we first investigated whether the covariance matrices differed significantly across Internet and paper-and-pencil administrations in order to test Hypothesis One. This test is an omnibus test of overall measurement invariance. In this model everything is constrained to be equal (factor form, factor loadings, item indicator variances, and factor variances), so obtaining acceptable fit here indicates that no further model testing is required. Upon comparing the covariance matrices across the two administration modes, the chi-square significance test indicated that the covariance matrices are non-invariant across the

administrations $[x^2 (I 10) = 216.47, p < .05]$ see Table I. However, given this statistic's extreme sensitivity to model complexity, sample size, and model misspecification, many researchers have recommended assessing model fit based on various absolute and incremental fit indices (Bentler & Bonett, 1980; Cheung & Rensvold, 2001; Ding, Yelicer, & Harlow, 1995; Hu & Bentler, 1999). The obtained fit indices indicated acceptable fit for the model
(CFI = .92, RMSEA = .08, SRMR = .07). While the obtained values for these fit indices are slightly lower than contemporary values required (CFI = 0.95, RMSEA = 0.06), given that smaller sample sizes are widely known to affect the estimates of fit measures (Boomsma, 1983; Hu, Bentler, & Kano, 1992), we chose to use the more liberal cutoff values associated with more established criteria. Furthermore, more traditional guidelines also corroborated model fit. For example, Carmines and Mciver
(1981) suggested that a X^2/df ratio of no more than three serves as an adequate indicator of good fit. Based on this indicator, the obtained cross-mode covariance matrices are equivalent (x^2/df ratio = 1.97). Given these results, it is argued by the authors that consistent with Hypothesis One, job satisfaction ratings are indeed invariant across paper-and-pencil and Internet administrations.

In order to test for the effects of the covariates (gender and age) on latent job satisfaction variability across administration modes, a MIMIC model was assessed. This test examines the significance of the relationships between the covariates and latent job satisfaction across administration modes by comparing a model in which these relationships are constrained to be equal across administration modes to one in which the relationships are free to vary across administration mode. The constrained model produced a significantly worse fit to the data in terms of the chi-square significance test $[i\uparrow x2(4) = 60.92, p < .05]$ as well as a change in CFI of .0I. The change in CFI of .0I indicates a decrement in model fit, according to guidelines developed by Cheung and Rensvold (2002).

Next, we let each covariate vary across administration modes one at a time and compared the fit each of these models to the model in which all relationships were constrained to be equal across administration mode. Only the model assessing the improvement in model fit when the effects for age were free to vary produced a significant improvement in terms of the chi-square and practical indexes $[i\uparrow x^2(4) = 15.97, p < .05]$. Table 2 shows that in the paper-and-pencil condition, age was positively and signif-

Table I
Results of Tests of Measurement Invariance Across Paper-and-Pencil
and Internet Administrations

Model	df	$x2$	Mt	$t.x2$	CFI	RMSEA	SRMR
Equal covariance matrices	1 10	216.47'		.92	.08	.07	
Baseline MIMIC model	124	293.76'		.91	.09	.08	
Constrained MIMIC model	136	354.68'	12	60.92'	.90	.08	.08
Age	132	338.71 '	4	15.97'	.91	.08	.08
Gender	132	350.03'	4	4.65	.91	.09	.08

Note-CF!, comparative fit index; RMSEA, root mean square error of approximation; SRMR. standardized root mean square residual. $p < .05$.

icantly correlated with both satisfaction dimensions. That is, in the paper-and-pencil condition, over both satisfaction variables, satisfaction increases with age. However. Age was not related to either satisfaction construct for those responding via the Internet administration mode.

The model for gender failed to produce significant increases in the chi-square difference test, nor any significant changes in any of the practical indices [$X2(4) = 4.65, p < n.s.,$ CF!= .91, RMSEA = .09, SRMR = .08]. Inspection of Table 2 indicates that despite administration mode, females report higher levels of satisfaction than males. Thus, Hypothesis Two was partially supported. While the relationship between respondent gender and latent job satisfaction variability remained invariant across administration mode, respondent age related to latent job satisfaction variability differently across paper-and-pencil and Internet administrations.

The results of an effect size analysis suggest that age interacts with administration mode to produce a value decrement of approximately 0.2 on a five-point Likert scale as one moves from paper-and-pencil administration to Internet surveying. While this value may seem minimal, on a practical level, this decrease in latent satisfaction ratings as a function of administration mode could (and in this data set actually did) lead to apparent significant differences in satisfaction for older employees.

Discussion

Importantly, these results reaffirm past research that has shown measurement invariance between differing survey administration modes (e.g., Ployhart et al., 2003; Whitaker, 2004), however, the MIMIC method used in this study suggests that respondent characteristics may systematically influence the responses made to the measurement device differentially across administration mode.

Consistent with past research indicating that females generally report higher levels of job satisfaction than males (e.g., Kacmar & Ferris, 1989), the results of the present study indicate that females tend to be more satisfied with both the supervision and the work itself than their male counterparts. Furthermore. the results of the MIMIC modeling suggest that this relationship is consistent across survey administration modes.

Conversely, a significant finding emerged for the respondent age covariate. These results suggest that while age was positively related to job satisfaction for ratings obtained from paper-and-pencil survey administration,

this relationship was not evident for ratings provided via Internet administration. An important point to bear in mind regarding this finding is the repeated measures design of the study. That is, the same employees responded to the same survey items over differing administration modes.

Given these results, it may serve organizations well to first identify if this issue exists within its workforce, and then assist those employees who feel uncomfortable using newer forms of technology for data collection purposes. For example, Reips (2002) advocate the adoption of a multiple site entry technique wherein an experimenter (or organization) may strategically segment the its sample (or workforce) on any number of variables (for example by age). send each segment a unique URL to the same on-line survey, and upon data collection, analyze the data for evidence of self-selection measured via drop-out. Upon diagnosing that older workers within the organization differentially refrain from online surveys, the organization may intervene by instituting programs meant to alleviate distress surrounding this data collection medium.

While much research has examined the difficulties faced by older workers in the workforce (e.g., Sterns & Doverspike, 1989), little research has examined the influence of age on the acceptance of technology. On the practical side, if the differential effects observed are a result of computer anxiety, then this suggests that organizations would do well to consider interventions aimed at reducing this anxiety through computer training, as computer anxiety is a temporary condition that can be reduced through a comfortable learning environment (Ayersman & Reed 1996). Indeed research has demonstrated the success with which cooperative learning alleviates the negative effects of computer anxiety (Dupin-Bryant, 2002; Keeler & Anson, 1995). However, if these findings do not directly result from an inherent tendency toward computer anxiety for the older worker, rater training in organizations should take into account the influence of these demographic variables on ratings. For example, specific training programs could be developed and targeted to older workers in order to influence latent satisfaction rating variability through a reduction in rater biases (e.g., halo, severity, etc.).

Lastly, the issue of technical variance should be considered in light of these findings. The participants in this study had the option of filling out the online survey from any computer to which they had access. Assuming that many chose to complete this survey at home, and the out- comes of this study indeed result from greater variation in browsers or monitors for older employees, organizations may wish to institute a policy of completing online satisfaction surveys at work only. This would allow the organization to control the browsers employed and the screen resolution settings in order to ensure standardized survey visualization across all employees.

On the theoretical side, much effort has been directed toward delineating the domain space that constitutes the job satisfaction criterion (e.g., Judge, Locke, Durham, & Kluger, 1998; Judge & Watanabe, 1998). The results of this study demonstrate that administration mode may interact with age to influence latent job satisfaction ratings. Our findings suggest, while seemingly innocuous,

Table 2
Relationships Between Covariates and
Latent Satisfaction Constructs Across Administration

Administration Mode	Variable	Satisfaction Dimension	
		Supervision	Work Itself
Paper and pencil	Age	.32'	.33'
	Gender	.25'	.29'
Internet	Age	.16	.19
	Gender	.22'	.26'

Note- For the gender variable. males were coded as 1 and females as 2. 'p < .05.

respondent demographic variables should be considered when studying the influences that weigh on job satisfaction ratings in future research.

AUTHOR NOTE

Correspondence concerning this article should be addressed to B. G. Whitaker, Department of Psychology, University of Akron, 302 Buchtel Mall, Akron, OH 44325 (e-mail: bgw3@uakron.edu).

REFERENCES

AYERSMAN, D., & REED, W. (1996) . Effects of learning styles. programming and gender on computer anxiety. *Journal of Educational Computing Research,* 28, 148-161.

BARAK, A., & COHEN, L. (2002). Empirical examination of an online version of the Self Directed Search. *Journal of Career Assessment,* 10, 387-400.

BARAK, A., & ENGLISH, N. (2002). Prospects and limitations of psychological testing on the Internet. *Journal of Technology in Human Services,* 19, 65-89.

BEDEIAN. A. G., FERRIS, G. R.. & KACMAR, K. M. (1992). Age. tenure, and job satisfaction: A tale of two perspectives. *Journal of Vocational Behavior,* 40, 33-48.

BENTLER, P. M. & BONETT, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin,* 88, 588-606.

BooMSMA, A. (1983). *On the rohustness of l/SREL (maximum likelihood estimation) against small sample si=e and non-normality.* Amsterdam: Sociometric Research Foundation.

BROSNAN, M. J. (1998). The impact of psychological gender, gender-related perceptions, significant others and the introducer of technology upon computer anxiety in students. *Journal of Educational Computing Research,* 18, 63-78.

BUCHANAN. T. (2001). Online personality assessment: Equivalence of traditional and WWW personality measures. I n U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of /nternet Science* (pp. 57-74), Lengerich, Germany: Pabst.

BUCHANAN, T.• & SMITH, J. L. (1999). Using the Internet for psychological research: Personality testing on the World-Wide Web. *British Journal of Psychology,* 90, 125-144.

BURNKRANT, S. R.. & TAYLOR, C. D. (2001). *Equivalence of traditional and Internet-based data collection: Three multigroup analvses.* Paper presented at the 16th Annual Conference of the Society of Industrial and Organizational Psychology, San Diego, April, 200 I .

CARMINES, E., & MclvER, J. (1981). *AnaZv::ing models with unobsen>-able variables.* In G.W. Bohrnstedt & E. F. Borgotta (Eds.), Social measurement: Current issues (pp. 65-1 15). Beverly Hills, CA: Sage.

CHEUNG, G. W. (1999). Multifaceted conceptions of self-other ratings disagreement. *Personnel Psychology,* 52, 1-36.

CHEUNG, G. W.. & RENSVOLD, R. B. (200I). The effects of model parsimony and sampling error on the fit of structural equation models. *Organizational Research Methods,* 4, 236-264. 962.

CHURCH, A. H. (2001). Is there a method to our madness? The impact of data collection methodology on organizational survey results. *Personnel Psychology,* 54, 937-962.

CONOLEY, J. C.. & IMPARA, J. C. (Eds.), (1995). The *twelfih mental measurements yearhook.* Lincoln: University of Nebraska, Burns Institute of Mental Measurements.

DILLMAN, D. A., & BoWKER, D. K. (2001). The Web questionnaire challenge to survey methodologists. In U.-D. Reips & M. Bosnjak (Eds.). *Dimensions of Internet Science* (pp. 159-178), Lengerich, Germany: Pabst.

DING, L. YELICER, W., & HARLOW, L. (1995). Effects of estimation methods, number of indicators per factor, and improper solution on structural equation modeling fit indices. *Structural Equation Modeling,* 2, 1 19-144.

DUPIN-BRYANT, P. (2002). Reducing computer anxiety in adults learning to use microcomputers. *Journal of Extension,* 40.

DURNDELL, A., & HAAG, Z. (2001). Computer self efficacy, computer anxiety, and attitudes towards the Internet and reported experience with the Internet, by gender, in an East European sample. *Computers in Human Behavior,* 18, 521-535.

ETCHEGARAY, J. M.. SEXTON, J. B., HELMREICH, R. L., & THOMAS, E. J. (2003). *Joh satisfaction ratings: Measurement equivalence across physicians and nurses.* Paper presented at the 18th Annual Conference of the Society of Industrial and Organizational Psychology, Orlando, April, 2003.

FEN LASON, K. J. (2000). *Multiple data-collection methods in a 360 .feedhack program: lmplications fiJr use and interpretation.* Paper presented at the 15th Annual Conference of the Society of Industrial and Organizational Psychology, New Orleans, April, 2000.

FRERICKS, L.. EHRHART, K.. AND O'CONNELL, M. (2003). *Computer anxiety and test per formance: Comparing selection test formats.* Paper presented at the 18th Annual Conference of the Society of Industrial and Organizational Psychology, Orlando, April, 2003.

GIBSON, J. L.. & KLEIN, S. M. (1970). Employee attitudes as a function of age and length of service: A re-conceptualization. *Academy of Management Journal.* 13, 41 1-425.

GLASS, C. AND KNIGHT, L. (1988). Cognitive factors in computer anxiety. *Cognilil'e Therapy & Research,* 12, 351-366.

Gt.ENN, W. D.• TAYLOR, P. A., & WEAVER, C. N. (1977). Age and job satisfaction among males and females: A multi-variate, multi-survey study. *Journal of Applied Psychology,* 62, 189-193.

GREGSON, T. (1987). Factor analysis of a multiple-choice format for job satisfaction. *Psychological Reports,* 61, 747-750.

GUADAGNOLI, E.. & VELICER, W. (1988). Relation of sample size to the stability of component patterns. *P.,ychological Bulletin,* 103, 265-275.

Gu10N, R. M. (1983). Comments on Hunter. Jn F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Perfiirmance measurement and theory* (pp. 267-275). Hillsdale, NJ: Erlbaum.

HoDsoN. R. (1989). Gender differences in job satisfaction: Why aren't women more dissatisfied. *Sociological Quarterly,* 30, 385-399.

HONEYMAN. D. S.. & WHITE, W. J. (1987). Computer anxiety in educators learning to use The computer: A preliminary report. *Journal of Research on Computing in Education,* 129-138.

Hu. L., & BENTLER, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional versus new alternatives. *Structural Equation Modeling.* 6, 1-55.

Hu, L., BENTLER, P. M.. & KANO, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin,* 112, 351-362

JoHNSON, N.D. (1987). Effects of inservice training on writing apprehension and computer anxiety in elementary school teachers. *Dissertation Abstracts International,* 48, 904-909.

JmNSON, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology,* 31, 177-192.

JORESKOG, K. G.. & GOLDBERGER, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable, *Journal of the American Statistical Association,* 10, 631-639.

JUDGE, T. A., LOCKE, E. A.. DURHAM, C. C., & KLUGER A. N. (1998). Dispositional effects on job and life satisfaction: The role of core evaluations. *Journal of Applied Psychology.* 83, 17-34.

JUDGE, T. A.. & WATANABE, S. (1998). Another look at the job satisfaction-life satisfaction relationship. *Journal of Applied Psychology,* 18, 939-948.

KACMAR, E.M.• & FERRIS, G.R. (1989). Theoretical and methodological considerations in the age-job satisfaction relationship. *Journal of Applied Psychology.* 14, 201-207.

KEELER, C.M., & ANSON, R. (1995). An assessment of cooperative learning used for basic computer skills instruction in the college classroom. *Journal of Educational Computing Research,* 12, 379-393.

KRANTZ, J. H.. & DALAL, R. (2000). Validity of Web-Based Psychological Research. In M. H. Birnbaum (Ed.), *Psychological experiments on the Internet* (pp. 35-60). New York: Academic Press.

LAM, S. S. K.. Hui, C, LAW, K. S. (1999). Organizational citizenship behavior: Comparing perspectives of supervisors and subordinates across four international samples. *Journal of Applied Psychology,* 84, 594-601.

Liu. M., REED, W.M.• & PHILLIPS, P.D. (1992). Teacher education students and computers: major, prior computer experience, occurrence, and anxiety. *Journal of Research on Computing in Education,* 24, 457-467.

MAHAR, D.. HENDERSON, R.. AND DEANE, F. (1997). The effects of computer anxiety. state anxiety, and computer experience on user's performance on computer-based tasks. *Persona/in· & lndil'idual Di!: .fi!rences,* 22, 683-692.

MAURER, T. J., RAJU, N. S.. & Cot.I.INS. W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. *Journal o/ Appli<'li Psychology,* 83, 693-702. Meier, S. (1985). Computer aversion. *Computers in Human Behlll•ior,* 1, 171-179.

MUTHEN, L., & MUTHEN 8. (2001). *Mp/us Us<'I"± Guide.* Los Angeles: Author.

MEYERSON, P. AND TRYON, W.W. (2003). Validating Internet research: A test of the psychometric equivalence of I nternet and i n-person samples. *Behlll•ior Research Methods, lnstrt1ments. & Comput<'l"s,* 35, 614-620.

NUNNALLY, J. C. (1978). *Ps1·chometric theo1T* (2nd ed.). New York: McGraw-Hill.

PASVEER. K. A.. & ELLARD. J. H. (1998). The making of a personality inventory: Help from the WWW. *Behlll·ior Res<:arch Method.1·. lnstrt1ment.1· and Computers,* 30, 309-313.

PAXTON, A.. & TURNER, E. (1984). The application of human factors to the needs of the novice computer user. *l11tematio11al .Journal o(Man-Machine Studies,* 20. 137-156.

PLOYHART, R. E., WFEKl.EY, J.. Hou z. B.. & KEMP. C. (2003). Wcb-based vs. pencil testing: Arc test scores comparable across applicants and incumbents'? *Personnel Psychology.* 56. 733-752.

PRECKEL. F.. & THIEMAN, H. (2003). Online- versus paper-pencil-version of a high potential intelligence test. *Swiss Journal of Ps1·clwlogr.* 62.131 -138

REIPS. U.-D. (2002). Standards for Internet-based experimenting. *Experimental Psychology,* 49, 243-256.

REIPS, U.-D.. & FRANEK. L. (2004). Mitarbeiterbefragungen per Internet oder Papier'? Der Einfluss van Anonymitiit. Freiwilligkeit und Alter auf das Antwortvcrhalten [Employee surveys via Internet or paper" The influence of anonymity, voluntariness and age on answering behavior]. *fl'irtscha fispsyclw/ogie,* 6, 67-83.

ROBINSON, P. W. (Ed.) (1981). *Fundamentals o(exp<'l"imental psrchology.* Englewood Cliffs, MJ: Prentice-Hall.

Ros EN. L.D.. & Wrn_, M.M. (1994). Computer availability, computer experience and technophobia among public school teachers. *Computers in Human Behavior,* 11. 9-31.

RosEN. L. D.. & WEIL. M. (1995). Adult and teenager consumer users of technology: Potholes on the information superhighway" *Journal o(Consumer Affairs,* 29. 55-84.

RosZNOWSKI. M. (1989). Examination of the measurement properties of the Job Descriptive Index with experimental items. *Journal o/ Applicd Ps1·c/10logy.* 74, 805-814.

SmH, P.C.. KENDALL. L. M.. & HUI.IN, C. L. (1969). *The mea.1·urement of satisfi1ctio11 in 1mrk and retirement.* Chicago: Rand McNally.

SPERA. T. J.. & MoYE. B. G. (2001). *Measurement equil'lllence hetween paper and lfrh s111Te\' methods in a multinational compam'.* Paper presented at the 16th Annual Conference of the Society of Industrial and Organizational Psychology. San Diego. April. 2001.

SIANTON. J. M. (1998). An empirical assessment of data collection using the Internet. *Personnel Psrclwlogr,* 51. 709-724.

STERNS. H. L.. & DoVERSPIKE, D. (1989). Aging and the training and learning process. In L. Goldstein. *Training and de1·elopme11t in organi=ation.* San Francisco: Jossey-Bass.

THOMPSON. I. f.. SURFACE. E. A.. MARTIN, D. I., & SANDERS, M. G. (2003). From paper to pixels: Moving personnel surveys to the Web. *Personnel Psychology,* 56, 197-227.

VANDENRERG, R. J.. & LANCE, C. E. (2000). A review and synthesis of the measurement i nvariance literature: Suggestions, practices. and recommendations for organizational research. *Organi=ationa / Research Methods,* 4, 4-70.

VARCA, P. E.. SHAFFER. G. S.. & McCAULEY. C. D. (1983). Gender dit: ferences in job satisfaction revisited. *Academ1· o/ Management .lournal.* 26. 384-353.

WHITAKER. 8. (April, 2004). *The Impact o/ Data Collection Methodology on Popular Data-Reporting Indices.* Poster presented at the 19th Annual Society for Industrial and Organizational Psychology Conference. Chicago.

WHITAKER. B. (July, 2004J. *Internet-ha.ml allitude asses.1·ment: Does gender af/{>ct mea.rnrement equil'lllence :'* Poster presented at the I 12th Annual American Psychological Association Convention. Honolulu.

WII.l.IAMS, S.. OGLETREE. S.. WOODBURN. W., AND RAFFELll. P. (1993). Gender roles, computer attitudes. and dyadic interaction performance in college students. *Sex Roles,* 29. 515-525.

WOEHR. D. J.. SHEEHAN. M. K.. & BENNET!, W. JR. (1999). *Understanding disagreement across rating sources: An assess11u•111 <(the mea.rnrt!ment equi\'(/ lence o(raters in 3M! degrel! fi'l!dhack srstems.* Poster presented at the Fourteenth Annual Conference of the Society for Industrial and Organizational Psychology. Atlanta.

YuN. G. W.. & TRUMBO. C. W. (2000). Comparative response to a survey executed by post, email. & Web form. *Journal o/ Computer-Mediatl!d Communication.* 6, 1-26.