Exploring Free Questionnaire Data with Anchor Variables:
An Illustration Based on a Study of IT in Healthcare

Ned Kock, Texas A&M International University, USA

Jacques Verville, Auburn University Montgomery, USA

ABSTRACT

This paper makes an important methodological contribution regarding the use of free questionnaires, illustrated through a study that shows that a healthcare professional's propensity to use electronic communication technologies creates opportunities for interaction with other professionals, which would not otherwise be possible only via face-to-face interaction. This in turn appears to increase mutual trust, and eventually improve the quality of group outcomes. Free questionnaires are often used by healthcare information management researchers. They yield datasets without clear associations between constructs and related indicators. If such associations exist, they must first be uncovered so that indicators can be grouped within latent variables referring to constructs, and structural equation modeling analyses be conducted. A novel methodological contribution is made here through the proposal of an anchor variable approach to the analysis of free questionnaires. Unlike exploratory factor analyses, the approach relies on the researcher's semantic knowledge about the variables stemming from a free questionnaire. The use of the approach is demonstrated using the multivariate statistical analysis software WarpPLS 2.0. The study leads to a measurement model that passes comprehensive validity, reliability, and collinearity tests. It also appears to yield practically relevant and meaningful results.

Keywords: *Anchor Variable, Electronic Communication, Factor Analysis, Healthcare, Partial Least Squares, Structural Equation Modeling*

INTRODUCTION

Can the use of electronic communication technologies, such as email and social networking tools, improve the quality of the work conducted by groups of healthcare professionals? This paper provides an affirmative answer to this question, while making primarily a methodological contribution regarding the use of free questionnaires in survey research on healthcare information and communication technologies.

The methodological contribution is illustrated through a study showing that a healthcare professional's propensity to use electronic communication technologies creates opportunities for interaction with other professionals; opportunities that would not otherwise be available only via face-to-face interaction.

DOI: 10.4018/jhisi.2012010104

The end result appears to be an improvement in the quality of the work outcomes generated by groups of healthcare professionals. This improvement seems to be significantly mediated by an increase in mutual trust.

Survey research has been extensively used in the field of information systems and other fields that inform healthcare information management research (Galliers, 1992; Galliers et al., 2006; Ju et al., 2006; Orlikowski & Baroudi, 1991). It has also been extensively used in healthcare information management research itself (Erstad, 2003; Miller et al., 2004). Survey research enables healthcare information management researchers to study human-technology interaction and outcomes based on data that is both geographically distributed and builds on relatively large samples. Geographically distributed datasets are difficult to obtain through data collection approaches that rely on local samples such as field, case, and experimental research (Creswell, 2009). Large samples are typically difficult to obtain through intensive data collection approaches like field, case, and action research (Denzin & Lincoln, 2000; Kock, 2006). As such, survey research provides a good complement to other research approaches used in the field of healthcare information management.

In survey research typically questionnaires are used to collect data about a particular topic (Creswell, 2009; Drew & Hardman, 1985). Questionnaires can be designed with a general topic in mind or, more specifically, with certain constructs in mind (Creswell, 2009; Ehremberg & Goodhart, 1976). The former are referred to here as free questionnaires, where the component questions are not tied to a particular set of constructs. They include questions on a general topic, with the questions not necessarily expected to group around underlying constructs.

When questionnaires are designed with specific constructs in mind, the constructs are purported to be measured through multiple indicators. In this case, each indicator refers to a question-statement in the questionnaire, and is frequently measured on a Likert-type scale. The constructs and indicators are also expected to pass a confirmatory factor analyses (Ehremberg & Goodhart, 1976; Hair et al., 2009).

Even a questionnaire designed with certain constructs in mind may include questions that are not expected to be associated with specific constructs. This may happen as a researcher adds free questions to take advantage of a data collection opportunity. For example, a questionnaire may include 20 questions related to 5 key constructs, and another 15 free questions that are not specifically related to any underlying construct yet are anticipated to provide additional insights into the topic under study.

A study of electronic communication in healthcare organizations is discussed here. The study builds on survey data, and employs variance-based structural equation modeling (SEM) techniques (Chin, 1998; Chin et al., 2003; Maruyama, 1998; Schumacker & Lomax, 2004). It employs the multivariate statistical analysis software WarpPLS 2.0 (Kock, 2011). In the context of this study, a methodological contribution is made through the proposal of an anchor variable approach for the analysis of free questionnaires. An empirical contribution to the field of healthcare information management research is also made. As mentioned earlier, the study shows that trust in other professionals, in the context of distributed healthcare organizations, is an important mediating variable in the relationship between electronic communication technology orientation and quality of inter-professional outcomes.

FREE QUESTIONNAIRES

Whether a questionnaire can be seen as a free questionnaire or not depends on its relationship with a particular topic of investigation. Also, only a subset of the questions may make up a free questionnaire with respect to a particular topic. Frequently questionnaires are designed with one main topic T_1 for which a hypothesized model with theoretical constructs exist. These questionnaires may have additional exploratory questions that refer to a different but related topic T_2. For a researcher studying topic T_2 the questionnaire will be a free questionnaire.

The above example refers to the general situation where questions have been added to a questionnaire to explore topic T_2, but without a rigid organization of the questions around constructs. This may be done to explore difference facets of a general phenomenon, associated with topics T_1 and T_2, or to take advantage of the data collection opportunity to obtain additional research data. The latter is a common reason for the existence of free questionnaires (Creswell, 2009; Drew & Hardman, 1985).

Other reasons why free questionnaires may be produced are fully exploratory survey data analyses conducted to complement other research investigations, and exploratory data collection within the context of consulting projects. For example, a free questionnaire may be used to capture additional data in the context of a set of qualitative studies (Denzin & Lincoln, 2000). Also, a free questionnaire may be used to identify the reasons behind a business problem, as part of a consulting project.

In summary, even though they may appear to stem from poor research design, free questionnaires may become available to researchers for various reasons that are unrelated to poor research design. They may be a valuable source of data that could be difficult to replicate in the future, or provide the basis for the beginning of important research projects where future targeted data will be collected. It is important to devise strategies to deal with free questionnaires. This is one of the main goals of the discussion presented here.

ANCHOR VARIABLE VERSUS EXPLORATORY FACTOR ANALYSES

One problem posed by free questionnaires is that sets comprising different variables are likely to exist in which the variables are redundant (Creswell, 2009; Maruyama, 1998; Rosenthal & Rosnow, 1991). That is, the different variables in a set will essentially measure the same underlying construct. The direct use of these variables in multivariate models is problematic because it introduces collinearity into the models, which in turn leads to distorted results and misleading conclusions (Echambadi & Hess, 2007; Maruyama, 1998; Miller & Wichern, 1977).

The typical solution to this redundancy problem is to conduct an exploratory factor analysis on the dataset generated based on a free questionnaire (Ehremberg & Goodhart, 1976; Thompson, 2004). The exploratory factor analysis identifies sets of variables that may be associated with underlying constructs, called factors. Exploratory factor analysis algorithms uncover factors based on strong inter-correlations that usually occur among variables that are associated with the same underlying factor (Hair et al., 2009), even though correlated variables are not always redundant (Hamilton, 1987).

The main problem with exploratory factor analysis is that the algorithms used do not incorporate semantic knowledge about the variables when grouping them (Hair et al., 2009; Thompson, 2004). As such, different variables may be found to "belong" to the same factor because of strong inter-correlations, and yet may actually refer to different underlying constructs (Hamilton, 1987; Thompson, 2004). The strong inter-correlations may simply be due to strong underlying associations among different constructs.

One can expect the above problem to occur in survey research aimed at testing SEM models, which are essentially models with latent variables (Kline, 1998; Schumacker & Lomax, 2004). In these models, sets of variables stemming from answers to redundant questions are used to measure latent variables; in other words, they are used as indicators of the latent variables (Maruyama, 1998). This allows for the minimization of error via the calculation of latent variable scores based on weighted averages of the component variables, which are the sets of redundant variables.

In variance-based SEM, partial least squares or closely related algorithms are employed (Chin, 1998; Chin et al., 2003). The weights are calculated as multiple regression coefficients, and the latent variable scores are calculated as exact linear combinations of the component variables, where the indicator scores are combined based on the weights. This leads to a solution in which the error terms of the regression equations linking latent variable scores and their indicator scores are reduced to zero.

The anchor variable approach proposed here differs from exploratory factor analysis in that it relies on the researcher's semantic knowledge about the variables in a free questionnaire; knowledge which may precede or build upon the inspection of a free questionnaire. The anchor variable approach brings together two main elements, namely: the analysis of coefficients of correlation among variables, ordered by strength; and the analysis of the meaning of the questions from which the variables were obtained.

Another difference between the anchor variable approach proposed here and exploratory factor analysis is in the basic criterion used for uncovering candidate redundant variables. The criterion proposed here is that of a confirmatory factor analysis (Maruyama, 1998; Miller & Wichern, 1977; Thompson, 2004), and refers to acceptable convergent validity. More specifically, bivariate correlations among variables must be equal to or greater than 0.5 for them to be considered as possibly "belonging" to the same latent variable (Hair et al., 2009; Kline, 1998). Still, whether the variables are selected as "belonging" or not to the same latent variable, or as indicators of the latent variable, depends on whether they pass the semantic analysis done by the researcher.

Some degree of subjectivity is associated with the semantic analysis, as it essentially relies on the researcher identifying similarities in meaning among the questions that refer to correlated indicators. Because of that, it is recommended that a full validity and reliability analysis, as well as an additional collinearity analysis, be conducted on the dataset (Fornell & Larcker, 1981; Hair et al., 2009; Kline, 1998; Nunnaly, 1978). These are part of a confirmatory, as opposed to exploratory, factor analysis conducted after latent variables are identified. Arguably these will serve as solid validations of the semantic choices made by the researcher; if the dataset passes all of the validity, reliability and collinearity tests.

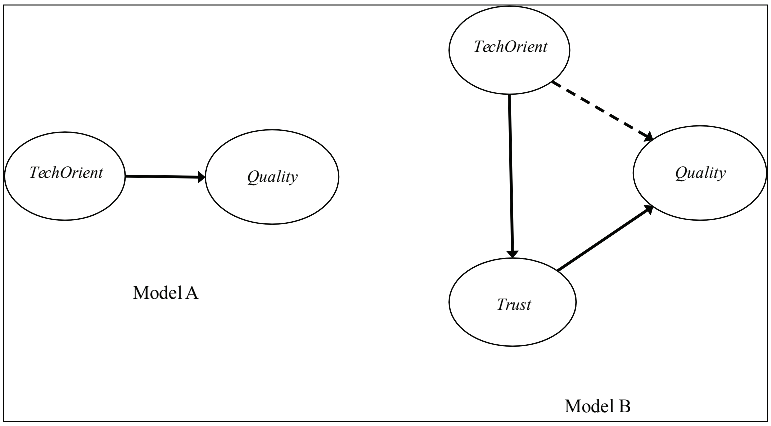
A STUDY OF ELECTRONIC COMMUNICATION IN HEALTHCARE ORGANIZATIONS

A fundamental theoretical idea that has been finding increasing support from empirical research is that electronic communication technologies affect performance via mediating variables, where the technologies act as facilitators of individual or group processes that are somewhat independent of the technologies (Kock et al., 2006; Meijer, 2008; Thatcher & Brown, 2010). That is, the individual or group processes exist prior to the use of the technologies; technology facilitation leads to adaptations of those processes. This seems to be particularly true for group tasks that involve distributed knowledge (Kock et al., 2006; Meijer, 2008), such as healthcare-related tasks.

For example, in healthcare settings, individuals with different professional backgrounds and expertise must collaborate to achieve common goals (Robinson & Casalino, 1996; Liang et al., 2010). Those individuals are often based in different and geographically distributed organizations, such as hospitals, doctors' offices, and laboratories (Erstad, 2003; Liang et al., 2010). In order to successfully collaborate, a certain degree of mutual trust must be established. One's propensity to use electronic communication technologies may create opportunities for interaction with other professionals, which would not otherwise be possible only via face-to-face interaction (Kock & DeLuca, 2007). This, in turn, could increase mutual trust, and eventually improve the quality of group outcomes.

Since the discussion presented here is mostly methodological, a certain degree of theoretical succinctness is called for, as is some

Figure 1. Models with constructs (Quality: quality of inter-professional outcomes; TechOrient: electronic communication technology orientation; Trust: trust in other professionals)



restraint in the amount of space dedicated to supporting the resulting model based on a review of the literature. Otherwise this paper would far exceed what would be reasonable in terms of length. The theoretical ideas presented above are fairly straightforward, and are depicted on the two models, A and B, shown in Figure 1. The models assume that the unit of analysis is the individual healthcare professional. They contain three theoretical constructs: *Quality*, or the quality of inter-professional outcomes achieved in healthcare organizations, as perceived by an individual healthcare professional; *TechOrient*, the electronic communication technology orientation of an individual healthcare professional; and *Trust*, the trust of an individual healthcare professional in other professionals.

Model A makes explicit the hypothesis that *TechOrient* is positively associated with *Quality*. Model B comprises three hypotheses, which can be summarized as follows: *TechOrient* is positively associated with *Trust*; *Trust* is positively associated with *Quality*; and the association between *TechOrient* and *Quality* is non-significant (noted as a dashed arrow) when one controls for the effect of *Trust* on *Quality*. Models A and B together express the meta-hypothesis that *Trust* is a "perfect" mediator of the relationship between *TechOrient* and *Quality* (Baron & Kenny, 1986).

DATA COLLECTION AND PRELIMINARY ANALYSES

The data for this study was collected through a mail survey sent to healthcare professionals in North America. The survey instrument was developed based on a previous qualitative study and literature on collaboration in healthcare settings involving the use of information and communication technologies. Respondents were asked to indicate the degree to which they agreed with the question-statements in the survey questionnaire (Appendix A). Question-statement items were measured using a five-point Likert-type scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

A small pilot study involving 30 healthcare professionals was used to pre-test the instrument and to identify any ambiguities and distracting stylistic problems with the survey questionnaire. In addition, a small number of individuals were interviewed regarding the questionnaire for additional refinement. The majority of changes pertained to rewording, sorting, and elimination of some of the questions.

After finalizing the survey instrument, the questionnaire survey was mailed to a random group of 1,820 healthcare institutions in North America. Among the participants, 9 (4.2%) were medical doctors, 136 (63%) were either registered nurses or nurse practitioners, and 72 (32.8%) identified themselves as other professional healthcare workers. The effective return rate was 11.9 percent, for a total of 217 usable questionnaires returned.

In keeping with good practice, we examined the differences between early and late respondents (Lambert & Harrington, 1990). Our analysis included the late responses as non-responses and compared them with the early responses (Abraham et al., 2008). Thus, for our analysis we used two sets of responses, where 60 were marked as early responses and 60 as late responses. We performed a non-parametric comparison of means test for the differences between the means of early and late responses, using a randomly selected set of 16 survey items. We concluded that there was no evidence to suggest that the respondents were not a representative sample of the population, and we proceeded with further analyses.

Anchor variables were initially selected from the questionnaire, which is a free questionnaire in the context of this study. They were selected to match the variables in models A and B above. Indicators were then identified based on correlations between the anchor variables and other variables in the dataset. This was followed by a variance-based SEM analysis, which included a confirmatory factor analysis and a collinearity analysis. The software used to conduct the analyses was WarpPLS 2.0 (Kock, 2011).

SELECTING ANCHOR VARIABLES AND IDENTIFYING RELATED INDICATORS

In order to test a theoretical model with data from a free questionnaire, a researcher should first identify anchor variables in the free questionnaire that match the variables in the model. This is done by inspecting each of the indicators (i.e., question-statements) in the questionnaire, and identifying the ones that appear to be the best semantic matches for the variables in the model. Those are essentially the indicators that are arguably the closest to those the researcher would have designed to measure the variables.

After the step above is completed, the researcher should create various sub-tables, one for each of the indicators in the dataset. Those sub-tables should contain the correlations between each indicator and all of the other indicators in the dataset. Figure 2 illustrates this, with sub-tables for 4 indicators, noted as *var13* … *var16*. These indicators refer to questions 13 to 16 of the free questionnaire used in the study of electronic communication in healthcare organizations, which is provided in Appendix A.

Next the researcher should reorganize the sub-tables that refer to the anchor variables, listing the correlated indicators in descending order of absolute strength (whether negative or positive). The researcher should then rename the indicators so as to clearly identify the latent variables and the indicators that refer to them.

The latent variable name should appear at the top of each sub-table, and the indicator names on the left should be the name of the latent variable followed by 1, 2, etc., depending on the number of indicators identified. These will be the indicators whose correlation to the anchor variable is equal to or greater than 0.5. This is illustrated on Figure 3, where the indicator previously called *var01* was renamed *Quality*, and two of the indicators were renamed *Quality1* and *Quality2*. Note that always the first indicator will be the one whose correlation is 1, as it refers to the anchor variable itself. The names of the original indicators, which in this case are *var01* and *var02*, are retained within parentheses for future reference. Those names remind the researcher that the indicators refer to question-statements 1 and 2 of the free questionnaire.

Once the indicators associated with the selected anchor variables have been identified, the researcher needs to create a modified version of the original dataset for a SEM analysis. This new dataset, extracted from the original dataset, will contain the columns that refer to the indicators that have been identified. Those columns should be renamed based on the results of the steps previously taken, with the new

Figure 2. Example of four variables with correlations

AK	AL	AM	AN	AO	AP	AQ	AR	AS	AT	AU
	var13			var14			var15			var16
var01	0.169		var01	-0.11		var01	0.326		var01	0.185
var02	0.124		var02	-0.092		var02	0.273		var02	0.291
var03	0.132		var03	0.011		var03	0.503		var03	0.385
var04	0.12		var04	0		var04	0.502		var04	0.315
var05	0.205		var05	0.145		var05	0.444		var05	0.286
var06	0.149		var06	0.114		var06	0.43		var06	0.433
var07	0.243		var07	0.036		var07	0.464		var07	0.408
var08	-0.059		var08	-0.044		var08	-0.369		var08	-0.221
var09	0.188		var09	0.066		var09	0.399		var09	0.381
var10	0.151		var10	0.174		var10	0.343		var10	0.314
var11	0.242		var11	0.193		var11	0.554		var11	0.394
var12	0.132		var12	-0.005		var12	0.175		var12	0.204
var13	1		var13	0.37		var13	0.193		var13	0.239
var14	0.37		var14	1		var14	0.188		var14	0.271
var15	0.193		var15	0.188		var15	1		var15	0.516
var16	0.239		var16	0.271		var16	0.516		var16	1
var17	0.201		var17	0.154		var17	0.448		var17	0.574

names referring to the indicators that have been identified. This is illustrated on Figure 4. As done before, the names of the original indicators are retained within parentheses for future reference.

The extracted dataset can then be used as the input for a SEM analysis; usually by being read by a SEM software tool. Latent variables should be created by aggregating the indicators that are expected to load on them. These are the indicators whose names start with the name of a latent variable and are followed by 1, 2, etc.

CONFIRMATORY FACTOR ANALYSIS AND DATA VALIDATION

The anchor variable approach to identifying latent variables and related indicators, proposed here, is an alternative to exploratory factor analysis. Still, it does not obviate the need for a confirmatory factor analysis, which is an important step in any SEM analysis. A confirmatory factor analysis provides the basis for data validation through reliability and validity checks. These checks are necessary before a researcher can trust the results of a SEM analysis.

Figure 3. Example of selection of indicators based on anchor variable

A	B	C	D	E	F	G	H	I
	Quality			var02			var03	
Quality1(var01)	1		var01	0.636		var01	0.477	
Quality2(var02)	0.636		var02	1		var02	0.391	
var03	0.477		var03	0.391		var03	1	
var04	0.453		var04	0.379		var04	0.558	
var29	0.45		var05	0.16		var05	0.472	
var33	0.428		var06	0.227		var06	0.412	
var34	0.412		var07	0.298		var07	0.452	
var43	0.364		var08	-0.24		var08	-0.517	
var19	0.355		var09	0.265		var09	0.487	

Figure 4. A section of the extracted dataset used in the SEM analysis

A	B	C	D	E	F	G	H	I
Quality1(var01)	Quality2(var02)	TechOrient1(var23)	TechOrient2(var61)	Trust1(var39)	Trust2(var49)	Trust3(var15)	Trust4(var18)	Trust5(var46)
5	4	4	5	4	4	4	4	4
5	5	3	3	3	4	3	3	5
4	4	5	5	2	4	2	4	4
5	4	3	3	3	3	4	4	3
4	4	4	4	4	5	4	5	5
5	5	5	4	3	5	4	5	5
4	5	3	3	2	4	2	2	3
4	4	3	3	2	3	3	2	3
5	4	4	4	3	3	4	3	3
5	4	4	4	3	3	5	4	4
4	4	4	3	3	5	5	5	4
5	5	5	5	5	5	5	5	5
5	5	5	5	4	5	5	5	5
5	5	4	4	3	4	4	3	4
5	4	4	4	2	3	3	3	4
5	5	2	4	3	3	4	4	4
4	4	2	3	3	3	4	4	4

Table 1 shows the loadings and cross-loadings among indicators and latent variables, as well as reliability coefficients. The loadings and cross-loadings were obtained by an oblique rotation of the structure matrix, and thus make up what is known as the pattern matrix relating indicators to latent variables (Hair et al., 2009; Maruyama, 1998; Miller & Wichern, 1977). Loadings, which are shown within parentheses, should be greater than 0.5 for convergent validity to be considered acceptable (Hair et al., 2009; Kline, 1998). This is the case here. The two coefficients of reliability, namely the composite reliability (CR) and Cronbach alpha (CA) coefficients, should be greater than 0.7 for reliability to be considered acceptable (Fornell & Larcker, 1981; Nunnaly, 1978). This is also the case here.

We can conclude based on the discussion that the dataset extracted through the anchor variable approach has both acceptable convergent validity and reliability. This essentially means, respectively, that: (a) the indicators are semantically associated with the appropriate latent variables, and (b) the indicators "belong" with each other.

Beyond convergent validity and reliability tests, another test that is commonly conducted as part of a confirmatory factor analysis is that of discriminant validity. Table 2 shows the coefficients that are needed for this test. They are the correlations among latent variables, shown within the intersecting cells; and the square roots of the average variances extracted (AVEs) for each latent variable, shown across the diagonal within parentheses. For discriminant validity to be considered acceptable, the square root of the AVE for each latent variable should be greater than any of the correlations between the latent variable in question and the other latent variables in the model (Fornell & Larcker, 1981).

As can be inferred from the latent variable correlations and square-roots of AVEs, the extracted dataset presents acceptable discriminant validity. This essentially means that the indicators identified through the anchor variable approach proposed here do not "belong" to latent variables other than the ones with which they were associated.

One final test that can be added to the ones above, for completeness, is a collinearity test. This test checks for the existence of collinearity among predictor latent variables in each latent variable block in the model. That is, it assesses collinearity among latent variables that point at the same latent variable, or vertical collinearity. For that, variance inflation factors (VIFs) are calculated for each of the predictor variables. VIFs lower than 5 suggest no collinearity (Hair et al., 2009; Kline, 1998), which is the case here as indicated by Table 3.

The validity, reliability, and collinearity tests discussed above are fairly comprehensive. The results suggest that the dataset extracted from the free questionnaire data presents acceptable convergent validity, discriminant valid-

Table 1. Indicator loadings, cross-loadings and reliability measures (Quality: quality of inter-professional outcomes; TechOrient: electronic communication technology orientation; Trust: trust in other professionals; CR: composite reliability coefficient; CA: Cronbach alpha coefficient. Loadings and cross-loadings are oblique-rotated).

	Quality	***TechOrient***	***Trust***	**CR**	**CA**
Quality1	(0.896)	-0.010	0.025	0.900	0.778
Quality2	(0.913)	0.010	-0.025		
TechOrient1	-0.003	(0.813)	-0.001	0.885	0.804
TechOrient2	0.022	(0.869)	0.005		
TechOrient3	-0.020	(0.863)	-0.004		
Trust1	-0.021	0.057	(0.803)	0.887	0.841
Trust2	-0.025	-0.001	(0.841)		
Trust3	0.096	-0.067	(0.741)		
Trust4	0.030	-0.012	(0.757)		
Trust5	-0.083	0.023	(0.768)		

ity, and reliability. The results also suggest that the extracted dataset is free from vertical collinearity. These not only indicate that the extracted dataset passes a confirmation factor analysis, but also that the anchor variable approach proposed here may lead to datasets that can be trusted as a basis for variance-based SEM analyses.

MAIN RESULTS OF VARIANCE-BASED SEM ANALYSIS

Figure 5 shows the main results of the SEM analysis for models A and B. Beta coefficients are shown next to arrows, and are standardized partial regression coefficients. They reflect the

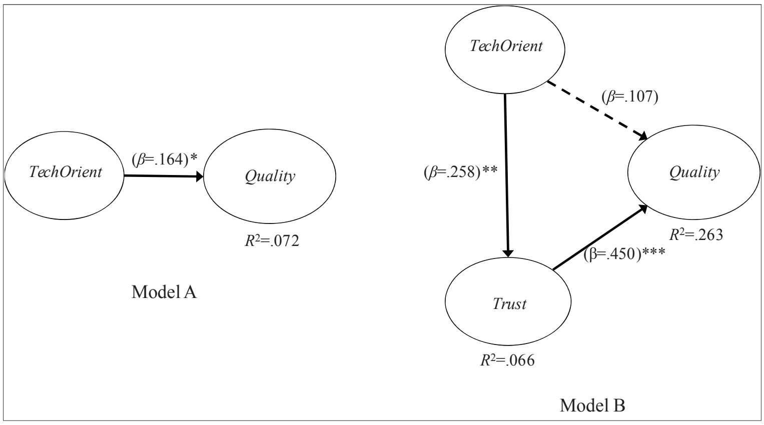
Table 2. Latent variable correlations and square-roots of AVEs (Quality: quality of inter-professional outcomes; TechOrient: electronic communication technology orientation; Trust: trust in other professionals; Square-roots of AVEs shown on diagonal within parentheses).

	Quality	***TechOrient***	***Trust***
Quality	(0.904)	0.129	0.353
TechOrient	0.129	(0.848)	0.204
Trust	0.353	0.204	(0.782)

Table 3. Vertical collinearity estimates (Quality: quality of inter-professional outcomes; TechOrient: electronic communication technology orientation; Trust: trust in other professionals; VIFs shown are for the only block where two or more latent variable predictors pointed a latent variable criterion in the model. VIFs lower than 5 suggest no collinearity).

	TechOrient	***Trust***
Quality	1.032	1.057

*Figure 5. Models with main results (*** P < .001; ** P < .01; * P < .05; Quality: quality of inter-professional outcomes; TechOrient: electronic communication technology orientation; Trust: trust in other professionals)*



strength of the associations between pairs of latent variables linked by arrows. Beta coefficients noted with an "***" are significant at the $P < .001$ level; "**" and "*" indicate beta coefficients significant at the $P < .01$ and $P < .05$ levels, respectively. R-squared values are shown under criteria latent variables to which predictor latent variables point to, and reflect the percentages of explained variance in the criteria by their respective predictors.

As predicted regarding Model A, one hypothesis was supported: *TechOrient* is positively and significantly associated with *Quality*. As predicted regarding Model B, three hypotheses were supported: *TechOrient* is positively and significantly associated with *Trust*; *Trust* is positively and significantly associated with *Quality*; and the association between *TechOrient* and *Quality* is non-significant when one controls for the effect of *Trust* on *Quality*. That is, the results for Models A and B, when taken together, provide general support for the meta-hypothesis that *Trust* is a "perfect" mediator of the relationship between *TechOrient* and *Quality* (Baron & Kenny, 1986).

To add generality to the test of the model, three control variables were added to both models A and B. These are not shown on the figure. The control variables were gender, education level, and work experience. These variables stored the answers to the three questions at the end of the original free questionnaire, which is provided in Appendix A. The meaning of the addition of these three control variables, within the context of the analysis and its results, is that the results summarized above hold regardless of (or when one controls for) gender, education level, and work experience.

DISCUSSION

Statistical software tools that conduct exploratory factor analysis cannot make semantic judgments about variables. They can identify clusters of correlated variables, and thus factors that are correlated with the clustered variables. The anchor variable selection approach proposed here allows for the use of the semantic judgment ability of one or more researchers. As it can be seen from the study discussed here, that can lead to latent variable and indicator choices that are not only semantically sound, but also statistically sound.

Validity and reliability analyses of models containing reflective latent variables naturally rely heavily on correlations among indicators. In other words, once semantically sound choices are made, high correlations among indicators are critical for the model to pass validity and reliability tests.

The anchor variable selection approach proposed here for identification of latent variables and respective indicators, and more specifically the 0.5 correlation threshold criterion, seems to lead to choices that pass widely accepted validity and reliability criteria. This is demonstrated by the study's confirmatory factor analysis results.

In the study, loadings between indicators and respective latent variables were all greater than 0.5, suggesting good convergent validity (Hair et al., 2009; Kline, 1998). Composite reliability and Cronbach's alpha coefficients were greater than 0.7 for all latent variables, suggesting good reliability (Fornell & Larcker, 1981; Nunnaly, 1978). The square roots of the average variances extracted for all latent variables were greater than the correlations between the latent variables and the other latent variables in the model, suggesting good discriminant validity (Fornell & Larcker, 1981).

CONCLUSION

Survey research has been extensively used in the field of information systems and other fields that inform healthcare information management research, as well as in healthcare information management research itself. It enables investigators to study human-technology phenomena based on data that is both geographically distributed and builds on relatively large samples. In survey research typically questionnaires are used to collect data about a particular topic. They can be designed with a general topic in mind or, more specifically, with certain constructs in mind; the former are referred to here as free questionnaires, where the component questions are not tied to a particular set of constructs.

A novel methodological contribution is made here through the proposal of an anchor variable approach to the analysis of free questionnaires. The anchor variable approach proposed here relies on the researcher's semantic knowledge about the variables in a free questionnaire. It brings together two main elements: the analysis of coefficients of correlation among variables, ordered by strength; and the analysis of the meaning of the questions from which the variables were obtained.

The use of the approach is demonstrated in a study of electronic communication in healthcare organizations. The study suggests that a healthcare professional's propensity to use electronic communication technologies creates opportunities for interaction with other professionals; opportunities that would not otherwise be available only via face-to-face interaction. This eventually appears to lead to an improvement in the quality of the work outcomes generated by groups of healthcare professionals; an improvement that seems to be significantly mediated by an increase in mutual trust.

The study uses survey data and employs variance-based SEM techniques. Through the analysis of the data, it is shown that the anchor variable approach yields a measurement model that passes comprehensive validity, reliability, and collinearity tests. Moreover, the approach also appears to yield practically relevant and meaningful results.

Researchers in multidisciplinary fields like healthcare information management and information systems often make methodological contributions that are relevant for their specific fields and many other fields. A good example is the leadership role played by information systems researchers in the development of ideas, techniques, and software tools for variance-based SEM (Chin, 1998; Chin et al., 2003). The anchor variable approach proposed and discussed here aims at making an interdisciplinary contribution that is of importance for the fields of healthcare information management and information systems; and whose importance extends to other fields. As such, the contribution made here is likely to help solidify the position of healthcare information management and information systems as reference fields (Baskerville & Myers, 2002).

REFERENCES

Abraham, K. G., Helms, S. E., & Presser, S. (2008). *How social processes distort measurement: The impact of survey nonresponse on estimates of volunteer work*. Cambridge, MA: National Bureau of Economic Research. doi:10.3386/w14076

Adelman, I., & Lohmoller, J.-B. (1994). Institutions and development in the nineteenth century: A latent variable regression model. *Structural Change and Economic Dynamics*, *5*(2), 329–359. doi:10.1016/0954-349X(94)90008-6

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. doi:10.1037/0022-3514.51.6.1173

Baskerville, R. L., & Myers, M. D. (2002). Information systems as a reference discipline. *Management Information Systems Quarterly*, *26*(1), 1–14. doi:10.2307/4132338

Chin, W. W. (1998). Issues and opinion on structural equation modeling. *Management Information Systems Quarterly*, *22*(1), vii–xvi.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*(2), 189–218. doi:10.1287/isre.14.2.189.16018

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2000). *Handbook of qualitative research*. Thousand Oaks, CA: Sage.

Drew, C. J., & Hardman, M. L. (1985). *Designing and conducting behavioral research*. New York, NY: Pergamon Press.

Echambadi, R., & Hess, J. D. (2007). Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science*, *26*(3), 438–445. doi:10.1287/mksc.1060.0263

Ehremberg, A. S. C., & Goodhart, G. J. (1976). *Factor analysis: Limitations and alternatives*. Cambridge, MA: Marketing Science Institute.

Erstad, T. L. (2003). Analyzing computer based patient records: A review of literature. *Journal of Healthcare Information Management*, *17*(4), 51–57.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. doi:10.2307/3151312

Galliers, R. (Ed.). (1992). *Information systems research: Issues, methods and practical guidelines*. Boston, MA: Blackwell Scientific.

Galliers, R. D., Markus, M. L., & Newell, S. (Eds.). (2006). *Exploring information systems research approaches*. New York, NY: Routledge.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.

Hamilton, D. (1987). Sometimes $R^2 > r^2_{yx1} + r^2_{yx2}$: Correlated variables are not always redundant. *The American Statistician*, *41*(2), 129–132. doi:10.2307/2684224

Ju, T. L., Chen, Y.-Y., Sun, S.-Y., & Wu, C.-Y. (2006). Rigor in MIS survey research: In search of ideal survey methodological attributes. *Journal of Computer Information Systems*, *47*(2), 112–123.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

Kock, N. (Ed.). (2006). *Information systems action research: An applied view of emerging concepts and methods*. New York, NY: Springer.

Kock, N. (2011). *WarpPLS 2.0 User Manual*. Laredo, TX: ScriptWarp Systems.

Kock, N., & DeLuca, D. (2007). Improving business processes electronically: An action research study in New Zealand and the US. *Journal of Global Information Technology Management*, *10*(3), 6–27.

Kock, N., Lynn, G. S., Dow, K. E., & Akgün, A. E. (2006). Team adaptation to electronic communication media: Evidence of compensatory adaptation in new product development teams. *European Journal of Information Systems*, *15*(3), 331–341. doi:10.1057/palgrave.ejis.3000612

Lambert, D. M., & Harrington, T. C. (1990). Measuring nonresponse bias in customer service mail surveys. *Journal of Business Logistics*, *11*(2), 5–25.

Liang, H., Xue, Y., Ke, W., & Wei, K. K. (2010). Understanding the influence of team climate on IT use. *Journal of the Association for Information Systems*, *11*(8), 414–432.

Maruyama, G. M. (1998). *Basics of structural equation modeling*. Thousand Oaks, CA: Sage.

Meijer, A. J. (2008). E-mail in government: Not post-bureaucratic but late-bureaucratic organizations. *Government Information Quarterly*, *25*(3), 429–447. doi:10.1016/j.giq.2007.05.004

Miller, R. B., & Wichern, D. W. (1977). *Intermediate business statistics: Analysis of variance, regression and time series*. New York, NY: Holt, Rihehart and Winston.

Miller, R. H., Hillman, J. M., & Given, R. S. (2004). Physician use of IT: Results from the Deloitte research survey. *Journal of Healthcare Information Management*, *18*(1), 72–80.

Nunnaly, J. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, *2*(1), 1–28. doi:10.1287/isre.2.1.1

Robinson, J. C., & Casalino, L. P. (1996). Vertical integration and organizational networks in health care. *Health Affairs*, *15*(1), 7–22. doi:10.1377/hlthaff.15.1.7

Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis*. New York, NY: McGraw-Hill.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.

Thatcher, S. M. B., & Brown, S. A. (2010). Individual creativity in teams: The importance of communication media mix. *Decision Support Systems*, *49*(3), 290–300. doi:10.1016/j.dss.2010.03.004

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association. doi:10.1037/10694-000

Ned Kock is Professor of Information Systems and Director of the Collaborative for International Technology Studies at Texas A&M International University. He holds degrees in electronics engineering (B.E.E.), computer science (MS), and management information systems (PhD). Ned has authored and edited several books, including the bestselling Sage Publications book titled Systems Analysis and Design Fundamentals: A Business Process Redesign Approach. He has published his research in a number of high-impact journals including Communications of the ACM*,* Decision Support Systems*,* European Journal of Information Systems*,* European Journal of Operational Research*,* IEEE Transactions *(various),* Information & Management*,* Information Systems Journal*,* Journal of the Association for Information Systems*,* MIS Quarterly*, and* Organization Science*. He is the Founding Editor-in-Chief of the* International Journal of e-Collaboration*, Associate Editor for Information Systems of the journal* IEEE Transactions on Professional Communication*, and Associate Editor of the* Journal of Systems and Information Technology*. His main research interests are biological and cultural influences on human-technology interaction, nonlinear structural equation modeling, information overload, electronic communication and collaboration, action research, ethical and legal issues in technology research and management, and business process improvement.*

Jacques Verville is an Associate Professor of Information Systems (IS) and Department Head of Information Systems and Decision Sciences (ISDS) at the School of Business, Auburn University - Montgomery. He holds degrees in Linguistics (BA), International Relations (MA), Information Resource Management (MSc) and a PhD in Organizational Information Systems. Jacques has published in a number of journals including Decision Support Systems; International Business Review *(IBR)*; IEEE Transactions on Professional Communications; Long Range Planning; Decision Sciences Journal of Innovation in Education; Industrial Marketing Management; Journal of Enterprise Information Management; International Journal of Technology Management*. Dr. Verville is co-author of* Acquiring Enterprise Software: Beating the Vendors at their Own Game.

APPENDIX A

Full Free Questionnaire Used

In your opinion, please indicate to which you agree with the following statements. Specify your answer using a response scale of 1 to 5: 1) Strongly Disagree; 5) Strongly Agree.

1. Working with colleagues from other professions leads to outcomes that I could not achieve alone
2. Creative outcomes emerge from my work with colleagues from other professions that I could have not predicted
3. My colleagues from other professions and I collaborate regarding patient goals and plans
4. My colleagues from other professions work through conflicts with me in efforts to resolve them
5. Colleagues are as likely as I am to address obstacles to successful collaboration
6. Extensive efforts are taken to avoid conflicts concerning the sharing of tasks and responsibilities
7. As a decision is considered, colleagues openly represent their professional perspectives
8. I rarely communicate with colleagues from other professional disciplines
9. The exchange of information between myself and colleagues from other professional disciplines is frequent and informal
10. We have a formal communication structure among colleagues as well as between colleagues and patients
11. Interprofessional communications is emphasized for both patient understanding and the capacity to work with members of the teams
12. Often there is more information than can be interpreted immediately
13. Technology contributes to effective communication
14. I was able to put a human face on the people communicating with me through technology
15. My interactions with colleagues from other professions occurs in a climate where there is freedom to be different and to disagree
16. My colleagues from other disciplines and I discuss our professional similarities and differences (including role, competencies and stereotypes)
17. I discuss with professionals from other disciplines the degree to which each of us should be involved in a particular case
18. Open communication between colleagues takes place as decisions are made for a patient
19. I feel comfortable enough to ask questions during our team meetings
20. I can say what I mean without fear of repercussions or misunderstanding within the group as well as between groups
21. I communicate in writing with my colleagues from other disciplines to verify information shared verbally
22. I utilize informal methods of communication (i.e., social networks, lunchtime, impromptu meetings, etc.) to communicate with my colleagues from other disciplines
23. I use technology at every opportunity to communicate with my colleagues
24. Information pertaining to patient care is related promptly to other parties
25. I am more satisfied when communicating with technology rather than face-to-face
26. I am more satisfied with face-to-face communication since it allows me to communicate more effectively
27. Team members find more opportunities to express their opinions during face-to-face communication

28. Decisions about approaches to treatment are made unilaterally by professionals from other disciplines
29. Incorporating views of treatment held by my colleagues from other disciplines improves my ability to meet patients' needs
30. Formal procedures/mechanisms exist for facilitating dialogue between professionals from different disciplines
31. I utilize only formal procedures for problem-solving with colleagues from other disciplines
32. Having to report observations to my colleagues helps our interprofessional team members better understand the work of other health professionals
33. Effective and good leadership increase collaboration
34. Leaders have an effect on the communication style of members of the team
35. Cooperative work with colleagues from other disciplines is not a part of my job description
36. I am not willing to sacrifice a degree of autonomy to support cooperative problem solving
37. My colleagues from other professions are not committed to working together
38. I consistently give feedback to my colleagues
39. Professionals from different disciplines are straight forward when sharing information
40. My colleagues from other disciplines and I often discuss strategies to improve our working relationship
41. Colleagues from other disciplines do not usually ask for my opinion
42. The colleagues from other disciplines with whom I work have a good understanding of the distinction between my role and their role(s)
43. I view part of my professional role as supporting the role of my colleagues with whom I work
44. I am willing to take on tasks outside of my job description when it is necessary
45. My colleagues from other professions stick rigidly to their job descriptions
46. I have a good understanding with my colleagues from other professions about our respective responsibilities
47. Extensive efforts are done to avoid conflicts concerning the sharing of tasks and responsibilities
48. Conflicts concerning the sharing of responsibilities are resolved with difficulty
49. Mutual trust between my colleagues is high
50. My colleagues are expected to keep each other informed about events or changes that affect them
51. My patient care decisions are not always trusted/supported by other members of my team
52. Relationships with my colleagues from other professions sustain themselves despite external changes in the organization or outside environment
53. I work to create a positive climate in our interprofessional team
54. Disciplinary affiliation is harmful to interprofessional collaboration
55. Our current social structure inhibits team members' interaction
56. Our current organizational environment (i.e., location, etc.) inhibits team members' interaction
57. We have the current communication technology available in our work environment
58. My colleagues and I have a positive attitude towards using technology for communication
59. Technology makes the communication among team members easier and more efficient
60. I prefer using technology for communication
61. Using technology enhances the effectiveness of my work and communication
62. I see technology as a substitute environment to face-to-face communication
63. Technology enhances the collaboration among interdisciplinary team members
64. I feel more comfortable in asking questions when using technology

65. All available media types (i.e., email, PDA, etc.) and face-to-face communication provide the same amount of confidentiality
66. I prefer face-to-face communication as it provides a higher level of confidentiality compared to using technology
67. The emergency of the issue determines the choice of media used to communicate
68. I prefer using technology as a medium of choice for complex situations
69. I prefer face-to-face communication for complex situations
70. Comments:

    ____
    ____

71. What is your job title/area of responsibility?

    ____
    ____

72. What is your gender?
    1. Female
    2. Male
73. Education level
    1. Associate degree
    2. Undergraduate (BA, BSc.)
    3. Graduate (MA, MSc., or PhD)
    4. M.D.
    5. Other (i.e., certifications, diplomas, etc.)
       Please specify:________________
74. Work Experience (number of years)

    ____
    ____

APPENDIX B

Extracted Latent Variables and Indicators

TechOrient (electronic communication technology orientation)

- *TechOrient1*. I use technology at every opportunity to communicate with my colleagues (question 23)
- *TechOrient2*. Technology makes the communication among team members easier and more efficient (question 59)
- *TechOrient3*. Using technology enhances the effectiveness of my work and communication (question 61)

Trust (trust in other professionals)

- *Trust1*. Professionals from different disciplines are straight forward when sharing information (question 39)
- *Trust2*. Mutual trust between my colleagues is high (question 49)
- *Trust3*. My interactions with colleagues from other professions occurs in a climate where there is freedom to be different and to disagree (question 15)
- *Trust4*. Open communication between colleagues takes place as decisions are made for a patient (question 18)
- *Trust5*. I have a good understanding with my colleagues from other professions about our respective responsibilities (question 46)

Quality (quality of inter-professional outcomes)

- *Quality1*. Working with colleagues from other professions leads to outcomes that I could not achieve alone (question 1)
- *Quality2*. Creative outcomes emerge from my work with colleagues from other professions that I could have not predicted (question 2)