Original Paper

Interpreting evidence from structural equation modeling in nursing practice

Journal of Research in Nursing
15(3) 275–284

DOI: 10.1177/1744987109346666
jrn.sagepub.com
SAGE

Kathy Newman
University of Alabama in Huntsville College of Nursing, 325 Sparkman Drive, Huntsville, AL 35899, USA

David Vance
School of Nursing and Edward R. Roybal Center for Translational Research in Aging and Mobility
School of Nursing 456, 1701 University Boulevard, University of Alabama at Birmingham, Birmingham AL 35294-1210, USA

Linda Moneyham
School of Nursing G015, 1701 University Boulevard, University of Alabama at Birmingham, Birmingham AL 35294-1210, USA

Abstract
Structural equation modeling is a statistical technique that allows researchers to examine multiple hypotheses while simultaneously controlling for error. It can consist of a variety of observed and latent independent, mediator, and dependent variables. Owing to it being confirmatory in nature, this statistical approach can be used quite readily to test theoretical models. Likewise, it provides overall fit indices that determine whether the model tested actually fits the observed data. This article provides nurses with the basics on structural equation modeling so they can assimilate evidence from studies that use this statistical tool and be able to incorporate such findings into practice.

Keywords
structural equation modeling, path analysis, latent variable, indicator variable, reference variable, evidence-based practice

Introduction

Nursing research uses a variety of well-reasoned conceptual models to explain a number of phenomena ranging from agitation in nursing home residents (Vance *et al.*, 2003) to predicting falls (Vance *et al.*, 2006a), the effects of physical activity on cognition (Vance *et al.*, 2005), and driving cessation in community-dwelling older adults (Vance *et al.*, 2006b).

Corresponding author:
Kathy Newman, University of Alabama in Huntsville College of Nursing, 325 Sparkman Drive, Huntsville, AL 35899, USA.
E-mail: newmank@uah.edu

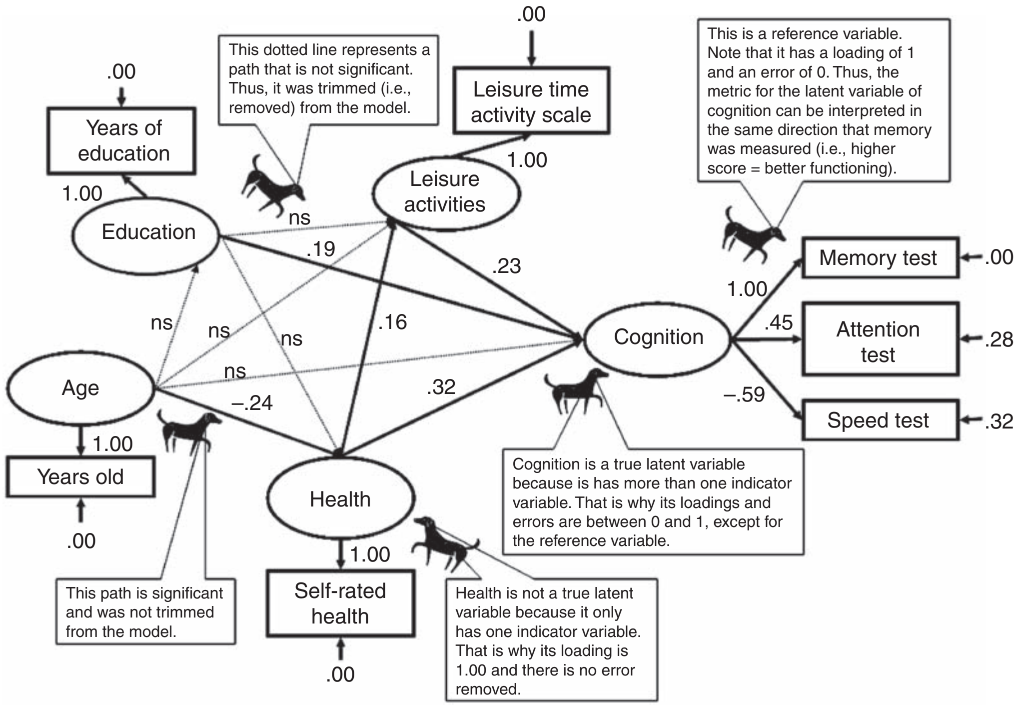

Figure 1. Fully trimmed structural equation model (ns = not significant)

Although conceptual models help to propel research, it is often difficult to test such models with conventional statistical approaches such as Student's *t*-tests, analysis of variance (ANOVA), multiple regressions, and chi-squared. One statistical approach that clearly stands out as an obvious choice for testing conceptual models is structural equation modeling (SEM).

SEM is a confirmatory statistical technique that is based on theory and has emerged as a reliable and useful method of examining relationships among multiple variables simultaneously. Succinctly, SEM allows one to draw a theoretical model with a variety of variables and paths among them (see Figure 1 as a finished example). Paths are the arrowed lines that reflect a hypothesized causal relationship between two variables. A path pointing away from a variable to another suggests that the variable being pointed to is being influenced by the preceding variable from which the arrow originated. These paths or relationships can be examined to determine whether they are significant; meanwhile, the entire model itself can be examined to determine whether the conceptual model proposed fits the observed data. Not only does SEM provide statistical indices to examine the data, the output actually includes a visual representation of the model being tested. This allows the researcher to visually inspect the conceptual model and determine whether the data confirm the conceptual model (the data fit the model, or the relationships among model variables occur in nature as predicted by the model) (Munro, 2005), which variable relationships are significant, and which effects of a variable on other variables are direct or indirect (a variable has an effect on a variable through its effect on a third variable).

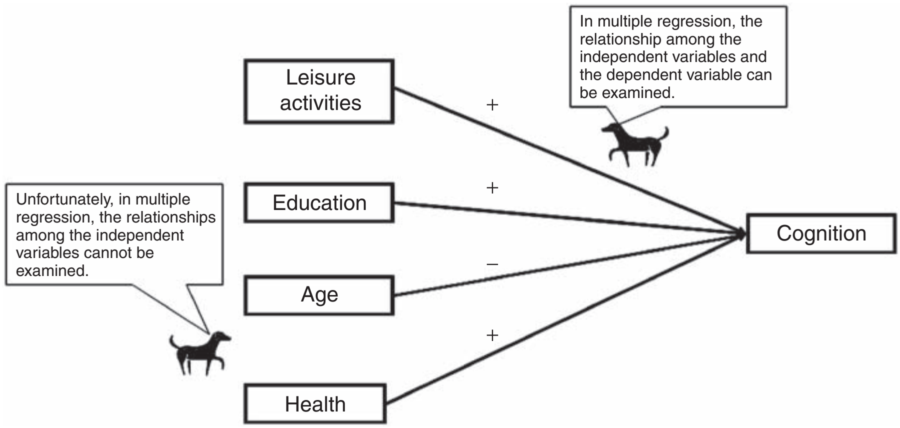

Figure 2. Conceptualization of multiple regression

The purpose of this article is to introduce the most basic principles of SEM including developing the research question, differentiating between path analysis and SEM, formulating the model, interpreting the models, and strengths and weaknesses of this approach. To facilitate this didactic process, several figures will be used along with a cartoon Irish pointer, which will explain some of the reasoning, rationale, and mechanics behind using SEM. Furthermore, a simple conceptual model relevant to nursing and successful cognitive aging will be used to explain the process of using SEM.

Developing the research question

Suppose that one wanted to identify what combination of independent variables are important in predicting successful cognitive aging in older adults. As depicted in Figure 2, one could use multiple regression to examine this problem. In this example, the independent variables are leisure activities, education, age, and health; the dependent variable is cognition. Assuming that each of these variables is measured on a continuous scale, multiple regression provides a simple and straightforward way of determining the amount of variance each independent variable accounts for in the dependent variable. Specifically, it is hypothesized that more leisure activities, higher educational attainment, and better health will have a positive relationship with better cognitive functioning; likewise, greater age will have a negative relationship with better cognitive functioning. Multiple regression allows one to determine which independent variables significantly account for the variance in cognitive functioning (Kutner *et al*., 2005).

Unfortunately, with multiple regression techniques, one can only identify predictors for a single dependent variable. However, more to the point, the independent variables of age, education, health, and leisure activities may have unique relationships among themselves that may also contribute to cognition. For example, age may be significantly related to health such that as age increases, health problems increase and age is only associated with cognition because of its relationship with health problems. However, multiple regression does not allow one to examine this explicitly, unless interactions are created. Even then, creating interactions becomes quite laborious, complex, cumbersome to interpret, and still

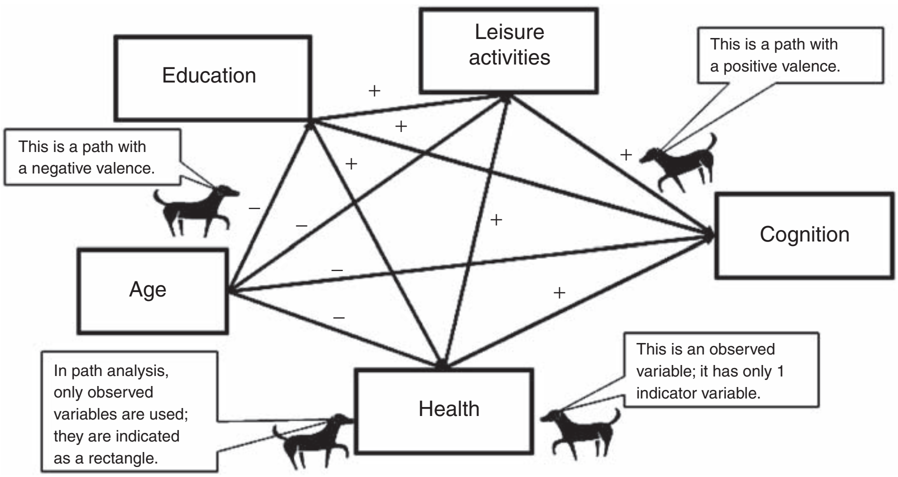

Figure 3. Conceptualization of path analysis

does not allow one to hypothesize on the causal influences between these independent variable as they impact cognition (Weston and Gore, 2006).

Fortunately, in order to examine and identify relationships among the independent variables, one could utilize path analysis, which is a simplified version of SEM (Figure 3). Therefore, in this example, one can also hypothesize on the relationships between age and health, age and education, health and cognition, and so forth. For example, one can infer that as people age, their health will become worse and in turn this will have a negative impact on cognition. Likewise, one can also infer that as someone's health declines, the ability to engage in leisure activities declines; since engaging in leisure activities has been shown to be important for maintaining cognitive abilities (Vance *et al.*, 2008), such a decline in leisure activities will lead to cognitive decline and so forth. Many of these causal hypotheses can be generated at the discretion of the researcher who is hypothesizing on the intricate relationships among the variables of interest based upon their theoretical model and knowledge of the literature. Thus, this statistical approach allows one the flexibility to examine a number of complex relationships simultaneously and infer causal influence in the statistical model, something that multiple regression cannot do (Munro, 2005).

Path analysis versus structural equation modeling

Path analysis includes an analysis of the combination of paths/relationships among the variables in the model, as has been presented so far. In the model (Figure 3), direct and indirect paths can be seen in the path analysis model; these are often referred to as direct and indirect effects. For instance, a direct path can be observed between independent and dependent variables by a direct line between the specified variables. For example in Figure 3, a direct effect can be hypothesized between age and cognition. An indirect effect can be noted between variables when there is not a direct line between a specified set of variables (Munro, 2005). For example, in Figure 3, note that age has an indirect effect with leisure activities based on the relationship of age with education and health. Also, note that age has both an indirect and direct effect on the dependent variable of cognition.

Moreover, in path analysis, the variables of interest only have one indicator variable; this is sometimes referred to as having only one observed variable. For example, in Figure 3, the indicator variable of education could be measured by asking the participant to indicate the grade last completed or could be measured by the Wide Range Achievement Test (WRAT), which is a measure of general education attainment (Wilkinson, 1993); however, only one observed measure of a conceptual variable is used in path analysis. Meanwhile, in SEM, multiple observed variables could be used to develop the conceptual variable. In this case, this is referred to as a latent variable. A true latent variable is the hallmark of SEM; it consists of two or more related measures that purport to measure the same concept such as education. In the example above, both observed variables (e.g., highest grade completed, WRAT) can be combined to form a unique composite measuring education. In SEM, observed variables and latent variables can be used. A latent variable is not measured directly by a tool or instrument. This is what distinguishes path analysis from structural equation modeling; otherwise, the statistical tools used to examine and interpret the hypothesized conceptual model in both path analysis and SEM are identical (Raykov *et al.*, 1991; Kleinbaum *et al.*, 1998).

Before discussing the formulation of the models, it is important to identify key terms specific to SEM. Variables are defined as indicator and latent variables. Indicator variables are directly measured either through self-report (e.g., age, gender), through a questionnaire (i.e., number of leisure activities), or through cognitive testing (e.g., score on the Mini-mental Status Exam or a neuropsychological memory test). Reference variables are specific types of indicator variables; they load (similar to a correlation but more reflective of the factor loadings in factor analysis) perfectly onto the latent variable. By doing this, the researcher can decipher the valence of the latent variables; in other words, do high scores mean the presence of more or less of that latent variable. In our example in Figure 1, a memory test is being used as the reference variable. Note the memory has a loading of 1.00 and an error of 0 being removed from it. Knowing that a higher score on the memory test means the participant has a better memory, therefore the latent variable of cognition will mean that a higher value on cognition will mean better cognitive functioning. This is important because the other cognitive tests that may be scored with a negative valence (e.g., lower scores on a speed test can mean the participant has faster reaction time and therefore better cognition) could make interpreting the results cumbersome.

As mentioned, true latent variables are composed of more than one indicator variable and often have a reference variable that helps with its interpretation within the model. Latent variables are 'constructs' not measured directly by the researcher. Latent variables are created by combining indicator variables, which are data collected by the researcher. They are unique but conceptually similar and are specified to load together similarly as in a factor analysis. In Figure 1, the latent variable of cognition has three loadings; memory is the reference variable and has a perfect loading of 1.00 while the attention and speed have a loading of 0.45 and −0.59 with error removed. In fact, one of the advantages of using a true latent variable is that such error is removed from the observed variables, making the latent variable a more stable measure of the concept being studied. These latent variables can be independent or dependent variables.

In SEM, independent and dependent variables have a different nomenclature but essentially mean the same thing. Exogenous variables are variables that have effects on other variables in the model and include the independent variables. Endogenous variables are variables that are affected by other variables in the model and include the dependent

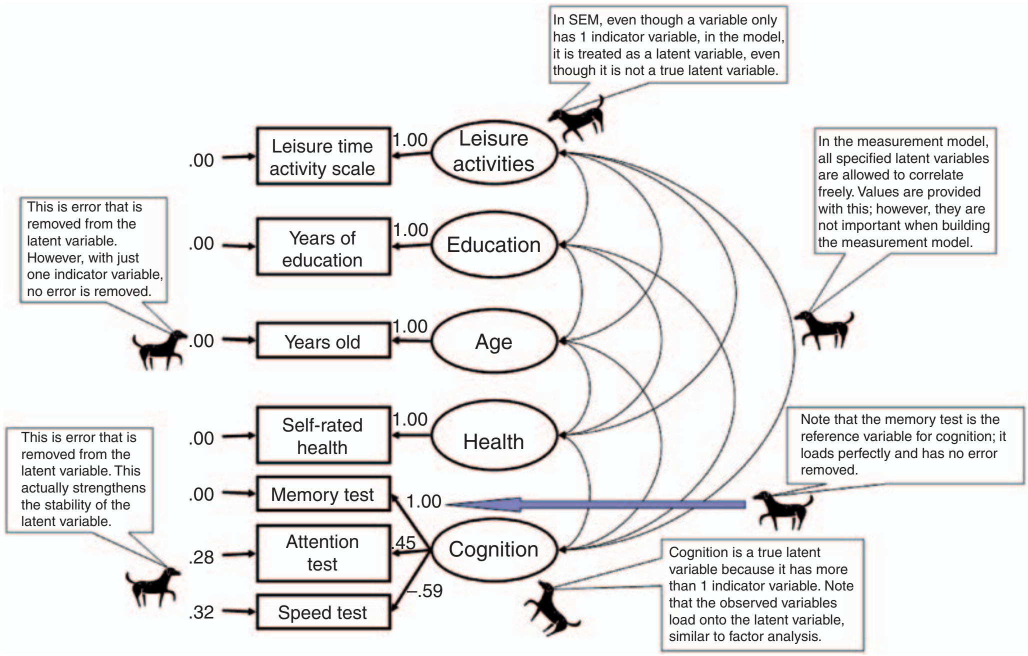

Figure 4. Baseline model

variables (Munro, 2005). In SEM, independent variables may function as exogenous or endogenous variables. For example, in Figure 1, health, education, and leisure activities are both exogenous and endogenous; however, only age is truly exogenous and only cognition is truly endogenous. Furthermore, a universal method for disclosing an indicator variable is to use a rectangle. The latent variable is represented in the model by an oval shape.

When utilizing SEM, a measurement model for each true latent variable must first be developed (Figure 4). Much like factor analysis, the measurement model construction is a process of evaluating the adequacy of the measures of the indicators of the latent variable. This is to ensure that the latent variables created for the model actually work. In other words, the factor loading for each of the indicators measures must be examined to determine whether they significantly load to the latent variable. If they do not, then the latent variable must be reassessed or that particular indicator variable that is not loading significantly must be removed. The researcher, as a rule, does not create a latent variable and use it in the model unless all of the indicators are significant; SEM computer programs provide them with this information (Grimm and Yarnold, 2008).

Formulation of the models

In building a SEM, three essential steps are needed: (1) building the baseline/measurement model; (2) specifying the full causal model; and (3) then trimming the model. The baseline model, also referred to as the measurement model, is comprised of the indicator and latent variables. The baseline model provides a means whereby the theoretical constructs are created and examined. For example, in Figure 4, the baseline model is depicted.

Note the indicator variables (rectangular in shape) of the leisure time activity scale, years of education, years old, self-rated health, memory test, attention test, and speed test. Latent variables are represented by oval shapes and include leisure activities, education, age, health, and cognition. In this model, the latent variable cognition is measured by three indicator variables (memory test, attention test, and speed test) and is thus a true latent variable. Therefore, one can hypothesize that the memory test, attention test, and speed test are indicators of cognition. Although leisure activities, education, age, and health are not true latent variables (e.g., they are measured by only one indicator variable), SEM allows for them to be treated as true latent variables. In essence, the baseline model is a confirmatory factor analysis model for each of the latent variables; in the baseline model, the relationships among the latent variables are not tested (Polit and Beck, 2004; Munro, 2005). Instead, the loadings of indicator variables on the latent variables are observed to determine whether they are significant. If so, then the latent variables are deemed an acceptable proxy for the conceptual variables that are being tested. As mentioned above, if one of the indicator variables does not load significantly onto the latent variables, then it is discarded or the latent variables are reconceptualized. The SEM programs determine whether these loadings are significant.

The primary purpose of the SEM is to build the full causal model that best reflects the relationships in the conceptual model and corresponds to the latent variable created in the baseline model. From this, SEM programs provide fit indices and information on whether the paths among the latent variables are significant or not. In fact, *t*-values are produced for each path, allowing one to visually see which paths are more significant and which are least significant (Kelloway, 1998).

Based upon the findings of the full causal model, the most non-significant path (below a *t*-value of 1.96) between the latent variables is removed one at a time. This process is referred to as model trimming. Each time a path is removed, the entire model is recalculated. If non-significant paths remain, the next least significant path is removed and the entire model is recalculated again. This process continues until there are no longer non-significant paths among the latent variables. Once this is achieved, the model is considered complete. As depicted in Figure 1, the non-significant paths are removed as indicated by a dotted line; only the significant paths remain and the model is interpreted from this trimmed state (Kelloway, 1998).

Interpretation of SEM

As seen in Table 1, a table of fit indices that are commonly reported in the literature is presented. In fact, there are a number of fit indices that help one determine whether the model that was specified and trimmed fits the observed data. There are more indices than listed here; even then, it may be difficult to judge whether a model fits the data by evaluating all of the fit indices. In truth, it is more than just looking at a *p*-value that determines significance. In general, those versed in SEM look at two primary fit indices: the goodness of fit index (GFI) and the root mean square error of approximation (RMSEA). A GFI ranges from 0 to 1 and a score of 0.90 or higher for the trimmed model is considered a good fit. A RMSEA ranges from 0 to 1 and a score of 0.10 or lower is considered a good fit (Polit and Beck, 2004; Munro, 2005). So, the trimmed model fit the data well; therefore, the trimmed model is accepted and then interpreted.

Table 1. Fit measures of baseline, causal, and trimmed models

	χ^2 (df)	GFI	AGFI	PGFI	RMR	RMSEA	NFI	PNFI	RFI
Baseline model	40.42 (22)	0.94	0.87	0.41	0.06	0.08	0.79	0.42	0.61
Full causal model	52.46 (33)	0.92	0.88	0.51	0.08	0.08	0.73	0.48	0.60
Trimmed model	54.45 (31)	0.91	0.89	0.56	0.08	0.07	0.72	0.51	0.62

Note: AGFI = Adjusted Goodness-of-Fit Index; df = degrees of freedom; GFI = Goodness-of-Fit Index; NFI = Normed Fit Index; PGFI = Parsimony Goodness-of-Fit Index; PNFI = Parsimony Normed Fit Index; RFI = Relative Fit Index; RMR = Standardized Root Mean Residual; RMSEA = Root Mean Square Error of Approximation.

At this point, once the trimmed model fits the data, the remaining significant paths between the latent variables are interpreted. For example, in this model, the older people are, the poorer their health. Better health predicts more involvement in leisure activities. Higher levels of education, more involvement in leisure activities, and better health predict better cognition. Meanwhile, there might be a significant indirect relationship between age and cognition mediated through health and leisure activities.

Strengths and weaknesses of SEM

SEM has several advantages. First, it is a confirmatory statistical approach that provides a method of examining a theoretical/conceptual model for determining significant paths among the variables of interest (exogenous and endogenous). Second, latent variables can function as exogenous or endogenous variables in the model. This is noteworthy when one considers these complex relationships in the real world. Third, it can incorporate latent variables composed of several indicator variables; using such true latent variables creates a more sound and reliable variable for analysis (Grimm and Yarnold, 2008; Munro, 2005).

Despite these strengths, there are weaknesses that hinder the use of SEM. First, depending on the actual number of exogenous and endogenous variables under study, developing the model can be very complex. The researcher may find it to be an arduous task from building the baseline model to trimming the final model (Kelloway, 1998). Second, as the baseline model is being built by theoretically constructing the latent variables, the researcher may find that the model does not converge. In other words, the model does not work and the SEM program cannot process the data. Hence, the importance of building the model with sound theoretical underpinnings is important in choosing the variables that will make up the latent variable. Utilizing stepwise techniques in model building with specific GFIs can assist the researcher in identifying the convergence of the model (Kelloway, 1998; Munro, 2005). Third, an additional consideration when using SEM relates to the sample size for the study. Some suggest 200–300 participants (Ullman, 1996). However, according to Bollen's rule, 5–10 participants are needed for every path observed in the model (Kelloway, 1998). Finally, SEM is a complex statistical approach that requires a great deal of time and effort to learn thoroughly. For those familiar with this approach, many may find that several key aspects were not covered. This was purposeful given the limited space, scope, and purpose of the article. Types of SEM programs, the difference between recursive and non-recursive pathways, an explanation of the 30 some fit indices, calculating and interpreting indirect pathways, using modification indices, correlations and covariance matrices, and matrix algebra were just some of the topics that were judged to only obfuscate a brief introduction into this technique.

Conclusion

SEM represents a powerful statistical approach that can be used to infer causation between observed and latent variables and test for complex hypotheses. To benefit from articles that use this approach, nurses do not need to know everything about SEM. However, it is important that nurses keep the following points in mind as they read the literature. First, the hypothesized model tested by SEM must be based on some combination of theory and findings in the literature. Second, SEM is strong if it has latent variables that are measured by multiple indicators. Third, only the significant paths in the trimmed model are considered important; and even this is only important if the GFI is 0.90 or higher and the RMSEA is 0.10 or lower. Finally, there are a number of computer programs that allow one to perform SEM; however, in reporting the results in the literature, generally they should all look similar. Other than that, there are many complexities of SEM that warrant discussion but are beyond the scope of this introductory paper. However, suggested readings are provided (e.g., Holye and Panter, 1995; Kelloway, 1998). The articles mentioned in the first paragraph are recommended readings that follow the template provided here and should serve as an addendum to this didactic on understanding the use of SEM in nursing research.

References

Grimm L, Yarnold P (2008) Reading and Understanding Multivariate Statistics. Washington, DC: American Psychological Association.

Holye RH, Panter AT (1995) Writing about structural equation models. In: Holye RH (ed.) *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks, CA: Sage, 158–176.

Kelloway EK (1998) Using LISREL for Structural Equation Modeling: A Researcher's Guide. Thousand Oaks, CA: Sage.

Kleinbaum D, Kupper L, Muller K, and Nizam A (1998) Applied Regression Analysis and Other Multivariable Methods. Pacific Grove, CA: Brooks/Cole.

Kutner M, Nachtsheim C, Neter J, and William L (2005) Applied Linear Statistical Models, 5th edn. New York: McGraw-Hill Irwin.

Munro B (2005) Statistical Methods for Health Care Research, 5th edn. Philadelphia, PA: Lippincott Williams and Wilkins.

Polit D, Beck C (2004) Nursing Research: Principles and Methods, 7th edn. Philadelphia, PA: Lippincott Williams and Wilkins.

Raykov T, Torner A, and Nesselroade J (1991) Reporting structural equation modeling results in psychology and aging: Some proposed guidelines. *Psychol Aging* 6(4): 449–503.

Ullman JB (1996) Structural equation modeling. In: Tabachnick BG, Fidell LS (eds) *Using Multivariate Statistics*. New York: HarperCollins, 709–811.

Vance DE, Ball KK, Roenker DL, Wadley VG, Edwards JD, and Cissell GM (2006a) Predictors of falling among older drivers from the state of Maryland: A structural equation model. *J Aging Physical Activity* 14: 254–269.

Vance DE, Burgio LD, Roth DL, Stevens AB, and Kaci K (2003) The relationship of naturalistic factors of agitation in nursing home residents. *J Gerontol Psychol Sci* 58B(2): P59–P67.

Vance DE, Roenker DL, Cissell GM, Edwards JD, Wadley VG, and Ball KK (2006b) Predictors of driving exposure and avoidance in a field study of older drivers from the state of Maryland. *Accident Anal Prevent* 38: 823–831.

Vance DE, Viamonte S, Foote A, Webb N, Marceaux J, and Ball KK (2008) Mental stimulation, neural plasticity, and aging: Directions for nursing practice and research. *J Neurosci Nurs* 40(4): 241–249.

Vance DE, Wadley VG, Ball KK, Roenker DL, and Rizzo M (2005) The effects of physical activity and sedentary behavior on cognitive health in older adults. *J Physical Exercise Aging* 13: 294–313.

Weston R, Gore P (2006) A brief guide to structural equation modeling. *Counsel Psychol* 34(5): 719–151.

Wilkinson G (1993) WRAT-3: Wide Range Achievement Test, Administration Manual. Wilmington, DE: Wide Range, Inc.

Dr Kathy Newman (MSN, CRNP), Clinical Associate Professor, UA Huntsville, College of Nursing (Huntsville, Alabama). She has been an educator for 18 years and teaches across undergraduate and graduate programs in the college of nursing at the University of Alabama in Huntsville College of Nursing, USA.

Dr David Vance (PhD, MGS), Associate Professor, Family/Child Health and Caregiving Department, School of Nursing, University of Alabama at Birmingham (UAB; Birmingham, Alabama). David is a developmental psychologist studying cognitive aging and aging with HIV with a special focus on structural equation modeling. He also teaches research methods and statistics in nursing in the PhD and DNP programs at UAB.

Dr Linda Moneyham (DNS, RN, FAAN), Professor and Rachel Z. Booth Endowed Chair School of Nursing, The University of Alabama at Birmingham. Linda teaches research design in the PhD program in the School of Nursing at UAB, and has over 20 years of research focused on women with HIV disease. Her research expertise is in the area of development and testing of interventions.